Master Thesis Pre-Study
A cyclic risk-assessment visualization of habits in the health sector

TIMMY NIELSEN

Master in Computer Science
Date: June 22, 2020
Supervisor: Somayeh Aghanavesi
Examiner: Joakim Gustafsson
School of Electrical Engineering and Computer Science
Host company: Health Support AB
Swedish title: Svensk Titel

Abstract

To be written

Sammanfattning

To be written på svenska

Contents

Chapter 1

Introduction

1.1 Background

Not done

Habits

Habits control our everyday life. From morning to evening habits developed throughout the years are encountered, having both negative and positive impact on the daily life of a person. Habits can be developed unconsciously [1] and often have harmful consequences, some examples being drug-abuse, alcoholism and over-eating [2]. Measuring and evaluating habits posing as risk for a healthy and comfortable lifestyle is a field not thoroughly researched, perhaps because of its arguably vague definition.

Visualization

Visualization as a scientific method fills the purpose of interpreting and translating data to make it better perceivable by a human subject [3]. Visualization can be thought of as a bridge between raw data and the human eye, with the main goal of making the data structure intuitive through graphical visuals, often accompanied with user interaction to iterate and filter the data.

1.2 Problem

Bad habits can have detrimental effect on health and lifestyle. Changing habits is known to be very hard and often require professional assistance to analyze and counteract. The methods used to evaluate these habits remain primitive compared to the technological advances of recent years, using traditional analysis methods conducted by a professional paid by the hour. This, and the fact that technology such as information visualization is not as advanced within health care compared to other scientific disciplines [4] indicate there is room for technological improvement on habit evaluation, especially in the field of health care.

Using visualization methods to improve habit evaluation has not been extensively researched, although it has been tested and shown potential in convincing patients to stop smoking [4]. However, the question of what impact modern visualization techniques can have on habit evaluation is not yet answered. This thesis puts focus into examining the possibilities of habit visualization through implementing and testing a solution within the area of childhood obesity.

1.3 Objective and Purpose

The objective of the project is to streamline the patient evaluation process as well as decrease evaluation data complexity for both employees and patients in a childhood obesity clinic. This is done through the creation of a questionnaire concerning diet habits, followed by a visualization of the results. Among the 300 childhood obesity patients, a portion will adapt the solution and undergo user testing to conclude the success of the project. There are two main reasons for narrowing the solution of this paper to such a specific area:

1. To provide a concrete measure of success for the concept within a particular field
2. To counteract the current lack of information visualization technology within the health care industry compared to other fields of science [4]

The purpose is to conceptualize a general method of habit visualization adaptable towards other fields of interest, whilst proving its worth through testing it in the particular scope of execution described above. The result will serve an indication of the value and potential extension to other organizations and branches of research.

1.4 Research Question

- Will a state-of-the-art cyclic habit risk-assessment visualization improve the treatment of childhood obesity through streamlining the evaluation process?

Divided into three sub-questions:

- Does the visualization improve the patient evaluation experience compared to current evaluation methods?
- Does the visualization decrease time spent on patient evaluation compared to current evaluation methods?
- Does the visualization convey enough information to relieve the evaluator from workload compared to current evaluation methods?

Chapter 2

Background

2.1 Habits

Habits can be defined as satisfactorily repeated behaviour. Habits often operate outside of our own awareness, and they are usually a means to obtaining certain goals or states [1]. Habits control our everyday lives, whether they are helpful or harmful in nature.

They can range from harmless, or even beneficial, every-day routines such as brushing your teeth to harmful and even life-threatening behaviour like drug-abuse, alcoholism or food addiction [2]. These habits have a negative effect on a person's health and environment.

2.2 Patient evaluation and its importance

Patient evaluation is a big part of a healthcare professional's career. Before treating a patient, a study must be performed to understand the specific symptoms and determine the appropriate treatment accordingly. This can take many forms, ranging from physical evaluation techniques to asking the patient questions through an interview or a questionnaire. This being such an important part of the treatment process, much time is spent speaking to and evaluating each patient before the actual treatment can begin [5, 6]. Streamlining this process would provide healthcare employees with more time for the actual treatment of a patient.

As a small (but relevant) side-note, waiting time is also a source of stress for the patient and known to cause aggravation [6]. Decreasing waiting time through streamlining the evaluation process would therefore not only serve as a valuable service to the healthcare professionals and their industry, but also potentially increase patient happiness and comfort as well as reduce tension between the patient and the employee.

2.3 Surveys and questionnaires

2.3.1 History

Surveys as we know them today has been around for approximately 70-90 years [7]. A survey can be conducted in many ways including interviews, questionnaires and other techniques for analytical evaluation [8]. The survey in this project will solely focus on the evaluation and visualization of a questionnaire.

Questionnaires in terms of the psychology and educational form it is today has existed since at least the early 1800's [9]. This is not to be confused with statistical research in general, which can be traced back as far as ancient Chinese and Roman times [10]. The questionnaire method has throughout the years proven to be a scientifically valid and appreciated tool for statistical research.

2.3.2 Questionnaire definition

The term "questionnaire" is often associated with a high quantity distribution of questionnaires evaluated into one joint statistical presentation. This fills the purpose of evaluating the average statistics among a group of people based on a high quantity of data. This project however will focus its attention towards a different form of questionnaire analysis, namely that of analyzing one questionnaire instance at a time. In other words, the questionnaire developed in this paper will be evaluated based only on a single response from one single respondent, not influenced by other responses to that same questionnaire distribution.

Note that in the field of questionnaire research, the difference between *qualitative* and *quantitative* questionnaires do not refer to the number of distributed questionnaires analyzed, but rather to the kind of answering method used. A qualitative questionnaire refers to using open-ended questions answered by free-text while a quantitative questionnaire uses multiple-choice or yes/no questions [11]. This is not to be confused with analyzing a high *quantity* of questionnaires versus just one at a time. Since this study will be restricted to using only multiple-choice questions it is according to scientific terminology a quantitative questionnaire regardless of the number of instances used in evaluation. Since the difference between the number of questionnaires included in the evaluation process is discussed in several sections of this paper, the terms **"single-response"**- vs **"multiple-response"** questionnaire is hereby defined as a general distinction between analyzing one single response at a time or all together. This distinction is made for the sake of simplicity and to avoid further confusion for the reader.

2.4 Single-response questionnaire scoring

How to properly score a questionnaire is a non-trivial problem, as there are many different point systems serving different purposes. In this section, the most popular scoring techniques will be presented, as well as general research in the subject of scoring.

In the 1920's a theory aiding in creating, analyzing and scoring questionnaires was developed. This theory is known as *Classical test theory* and was replaced by an improved modern approach in the 1980's called *Item response theory (IRT)*. IRT has the advantage of analyzing one item at a time, compared to the now outdated classical test theory that analyzes the full response score [12]. Using the IRT thus give more detailed information although it deploys similar mathematical models in its evaluation. IRT is used in a wide variety of areas, such as test publishing, departments of education, credentialing agencies, armed forces etc. [13].

Because of its long history and complex mathematical depth, the science of IRT can become overwhelming for the less conversant. A simpler evaluation technique is provided by IBM in their *Questionnaire Template Scoring Formulas* [14]. The proposed scoring algorithm concerns giving weight values to items, sections, and subsections of the questionnaire. Here, the weight defines the importance of the item or section in the evaluation process. This weight is combined with an answer score for each possible answer, defining the amount of impact that particular answer has on the final evaluation.

It is not uncommon to apply even more trivial scoring techniques such as the Patient Health Questionnaire *(PHQ)*. There are several editions of this questionnaire, ranging from 1 to 15 questions in total. They all adopt the same evaluation principles based on a simple point system. Each answer has a hidden value attached to it which is simply summed up in the evaluation process to reach a conclusion based on the total number of points. PHQ-9 has been validated in several studies showing it provides significant evaluation insight despite its simplicity [15, 16].

2.4.1 Likert scale

A re-occurring scale of scoring in survey development is called the *Likert scale*. It was developed in 1932 by psychologist Rensis Likert [17, 18] and it is the most common attitude-measurement scale within workplace studies, considered reliable and functional in its simplicity. The original scale uses a 5-point scale, allowing the following responses to a question:

1. strongly approve
2. approve
3. undecided
4. disapprove
5. strongly disapprove

although other variations of the scale are considered appropriate, depending on the type of survey [19].

2.5 Factor analysis

Another concept concerning questionnaire development is the factor analysis, an approach toward factoring questionnaire items into groups of concept. Applied within personality analysis, items similar to each other in the sense that they give insight into the same kind of personality traits are grouped together to give an understanding of how many different traits are measured in the questionnaire [20]. Although this is an appreciated method to classify items and evaluate questionnaire structure, it will not be applied to this project since the factors and their associated answers will be predefined in the questionnaire design phase, choosing specifically which questions will refer to which part of the 24-hour period beforehand. The factor analysis would then only serve as a confirmation of the initial grouping, which for the purposes of this study is considered too big of a commitment for an otiose reward.

2.6 Visualization

2.6.1 Purpose and efficacy

As described in the 1987 National Science Foundation's Visualization in Scientific Computing Workshop report, *"The goal of visualization is to leverage existing scientific methods by providing new scientific insight through visual methods"* [3]. Visualization is, in other words, a way of mapping computed data to a perceptual state adapted for the human eye to understand. If properly visualized, it is known to increase human comprehension of the data structure significantly [21]. Scientific visualization uses graphical design based on the science of human perception and interpretation to provide the most intuitive representation possible from the underlying data.

Research have shown that the human brains memory systems favor visual storage over unaided presentations, with results indicating that graphical messages are up to 43 percent more persuasive than conveying information through sound (speaking) [21]. Brotherstone et al. (2006) also explored the effectiveness of visual illustrations through a test conducted on 318 people, randomly allocated to receive written information alone or with visual aid. The aim was to teach the subjects about Flexible Sigmoidoscopy, and the results confirmed that visualizing the information provided a significantly better understanding among the subjects [22].

2.6.2 Designing with science

A central part of designing a visualization is color. The science of color shows that many dimensions are to be considered when choosing an appropriate color for a visualization. One interesting paper discusses the fact that native language affects perception of color [23]. This indicates that the language used by the target audience majority might have an impact on the user experience regarding the colors of the visualization. Other research shows that there are certain color interpretations common for the general public, such as the color red indicating danger/hazard while blue or green is typically associated with peace/tranquility/openness [24].

2.7 Health care

2.7.1 History

Health care is an ancient practice that has been around for longer than mankind. Taking care of each others health is a vital part of survival for both human and animals alike, and throughout history man has been pushing the limits to improve health care with experiments and technology. Although there is no real consensus about when modern medicine first appeared, there are records of surgical attempts dating back thousands of years. The first known trepanning operation is set around 5000 BC [25] as well as an old attempt at dentistry, curing teeth ache by drilling crowns into the teeth of patients between 6000-7000 BC [26]. Health care is of no doubt an old and very important profession.

Health care as exercised today fills the purpose of keeping people close to normal functioning and preserves for people the ability to participate in the political, social, and economic life of their society [27].

2.7.2 Visualization in health care

The health care industry, despite its scientific importance to man, is still quite primitive in terms of visualization technology. The US Institute for Medicine's 2011 Report sharply noted *"Information visualization is not as advanced in parts of clinical medicine as compared with other scientific disciplines"* [4]. This raises the question of ***why*** the health sector advances slower in this type of technology, followed by the question of ***how*** to counteract this lack of advancements. This thesis aims to take society one step closer to solving the question of ***how***.

Recent studies indicate that visualization serves as a powerful mean for conveying health information. Vivid information combined with visual elements seems to affect both affective and cognitive processes to maximize comprehension. The use of images, information graphics, animations etc. helps patients absorb information and gain situational awareness of medical information [21, 28]. According to Shneiderman et al. (2013), research on how to design persuasive visualization for patients who must be convinced to change habits such as smoking has considerable potential to bring large benefits to the industry [4]. This brings significant value to this study considering habit treatment being the main focus.

Some progress has been made throughout the years concerning visualization within health care, despite the fact that other scientific areas has adapted quicker. Examples include 3D volume visualization in x-ray and CT/MRI scans, expression analyses and tele-surgery etc [4]. These methods, among others, have increased treatment quality greatly. However, there are two notes to make regarding these methods. Firstly, they are heavily focused on on-site patient physical treatments. This does not cover the vast amount of long-term treatments where patients need to treat themselves outside of medical appointments, such as destructive habits and faulty diets. Secondly, the visualizations are typically adapted towards health care professionals to analyze. Considering visualization technology being a strong tool for abstracting complex information and amplifying cognition [29], why is it not being explored in terms of presenting data to the patients directly?

This study tackles this challenge through providing a brand-new, state-of-the-art visualization design with the purpose of providing both the patient and the health care professional with an intuitive graphical risk assessment of patient habits.

2.8 Related work

2.8.1 Questionnaire evaluation

A field of great success concerning the single-response questionnaire is the field of personality type evaluation, most known probably being the Myers-Briggs Type Indicator (MBTI). This test has grown steadily in popularity since it was first introduced around 70 years ago [30]. It is a questionnaire aimed to measure which personality type the subject possesses based on the famous psychological types described by Swiss psychiatrist C. G. Jung [31]. Although widely criticized and not officially validated, there is some research supporting the MBTI test being reliable in certain situations [30]. In whichever case, exploring the subject provides helpful information towards concluding how scoring is implemented in this paper.

A more extensive personality measurement is the Minnesota Multiphasic Personality Inventory (MMPI), first published in 1943 by the University of Minnesota Press. The revised version of the questionnaire (MMPI-2) contains up to 567 items, considerably more than the general questionnaire size [32], including the MBTI mentioned above. Another edition of the test was developed for adolescents (MMPI-A), holding 478 items and is the most widely used objective self-report personality and psychopathology questionnaire in existence [33].

A major field of research regarding the subject is the previously mentioned Item Response Theory [12]. The idea is based on mathematical models and rules aiding in reliably evaluating each test item in regards to the focus group. The mathematical models used are complex and take many variables into account, such as an *error of measurement*, calculated to eschew the problem of a person giving the wrong answer by mistake or checking a multi-choice answer box randomly [13]. A concrete example of IRT employment is the Graduate Record Examination *(GRE)* questionnaire - a proven reliable test of skill conducted when applying for an education [34]. The theory gives some useful indications on how one might go about developing and testing an evaluation algorithm for the questionnaire of this paper.

The Pediatric Quality of Life Inventory *(PedsQL)* is a validated approach to measuring health-related quality of life *(HRQOL)* primarily in adolescents and children [35], providing a risk-assessment not so unlike the one proposed in this paper. The latest revision (PedsQL 4.0) evaluates the four factors *Physical, Emotional, Social,* and *School Functioning*, and its scoring technique is based on a simple likert scale. The questionnaire has achieved excellent reliability scores in testing, distinguishing healthy children and children with chronic health conditions [36].

2.8.2 Related attempts at health care visualization

Health 2.0

Hesse et al. (2010) provide a framework for analyzing health informatics technologies, going by the name of “Health 2.0”. Health 2.0 suggests web based, participatory, and mobile strategies to employ social media and other personal technology to integrate with advanced statistical methods. They seek to combine personal, clinical and public health information to inform the public about their own health practices as well as enable clinics and the public to make better decisions through the valuable insight Health 2.0 provides. Several forms of visualization techniques are used, as can be seen in figure 2.1.

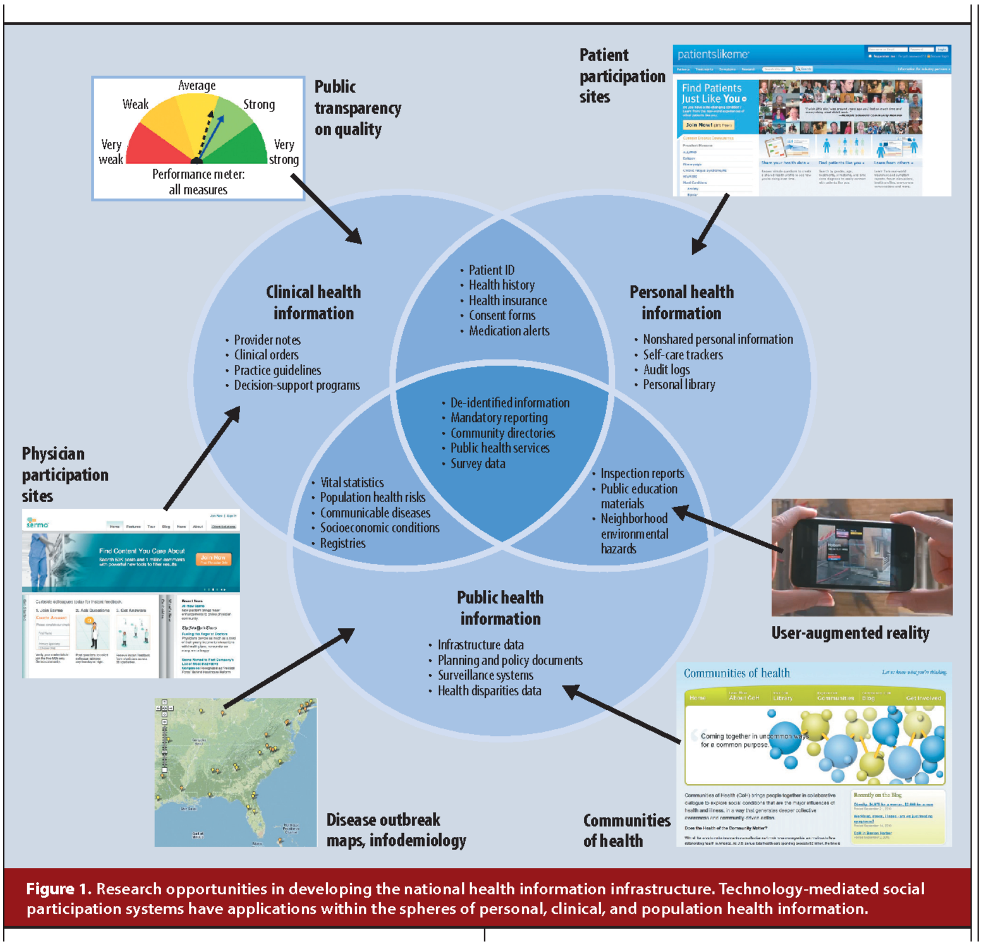

Figure 2.1: The three overlapping domains for Health 2.0. These three domains provide vastly different challenges for designers of interactive visualization tools. Clinical health information users are likely to be professionals who are frequent users with substantial training, but must make rapid life-critical decisions in busy distracting environments. Public health information users may be experienced statisticians, who process voluminous data to test complex hypotheses that could lead to massive national changes in public guidance or medical care [37].

Kurbo

Kurbo is a mobile application battling to solve a similar problem to that of this paper [38]. The app adapts a frequent use of visualization to make weight-loss more comprehensive to the children and adolescents using the app. An example is provided in figure 2.3.

Lifesum

The Lifesum app has a visualization keeping track of a subjects daily diet in the form of a circular timeline (See figure 2.4). The timeline is divided into sections ranging from breakfast to dinner. The design serves as a great inspiration to the solution of this project, as it is a popular app with a similar premise.

MBTI

MBTI also utilize information visualization in a customized manner when presenting their data (See figure 2.2).

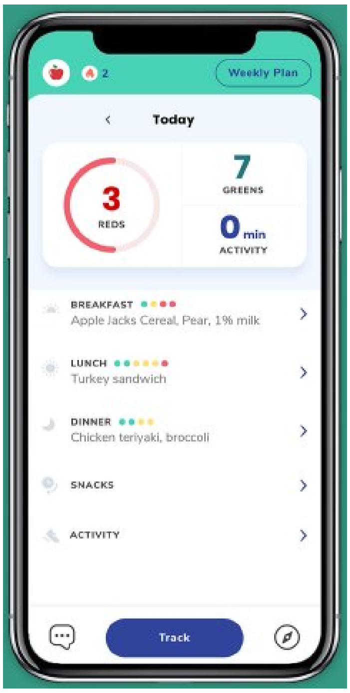

Figure 2.3: Visualization of diet habits in Kurbo

Figure 2.4: Lifesum daily consumption statistics

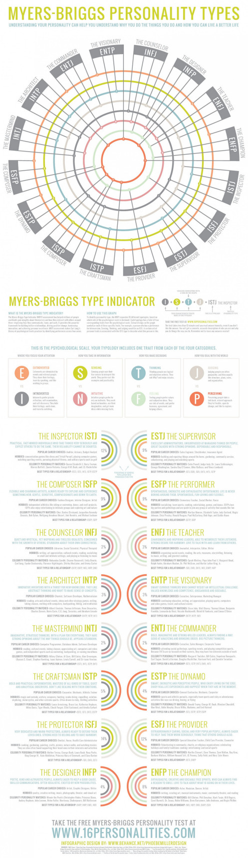

Figure 2.2: Visualization of MBTI personality traits

Chapter 3

Methodology

Although the main focus of the paper lies in applying visualization to relay information to patients more effectively, it requires an implementation of a questionnaire to gather the data to visualize. The project is thereby divided into two parts - the development of a survey in the form of a questionnaire and its automated answer evaluation algorithm, followed by a visualization presenting the result data. The questionnaire result will be presented in the form of a graphical visualization, perceptually pleasing and intuitive to understand by both patient and evaluator without any professional filtering or translation needed.

3.1 Software

- **JavaScript**
  Javascript was used as the main programming language for the frontend of the questionnaire as well as the visualization.

- **React.js**
  The react.js API was used within the JavaScript context for the questionnaire implementation

- **HTML Canvas**
  HTML Canvas was used for all the drawing computations of the visualization within the JavaScript context

- **Visual Studio Code *(VSC)***
  VSC was used as the integrated development environment throughout the thesis

3.2 Motivation

Analyzing a single individual based on a survey has been done within the health industry for a long time. When assessing a client's physical or mental health it is common to have them answer a form for the clinic in question to analyze and draw conclusions about their health from. Although this is a highly successful method, a flaw in the concept has been identified. The patient, not necessarily having experience in the field of diagnostics, has to have the survey results presented to them by a professional in a simplistic manner. This leads to an unnecessary time waste in the result distribution chain, where a simple set of questions is analytically interpreted by a professional to reach conclusions that then has to be simplified yet again to explain the results to the client, leading to an increased chance of misinterpretation and possible loss of detail in the process.

We aim to solve this issue by creating a dynamic survey algorithm that analyzes the answer data and outputs a visualization that is easily interpreted by both clinic and patient, hiding the more complex information only relevant to the clinic. This will give the clinic a direct overview of the patient without having to do any manual interpretation of the form, and furthermore it allows the patient to see its own survey results in a perceptually simple manner without the need of having a professional put time and effort into abstracting the results to an amateur level of understanding. Keep in mind that the possibility of more complex and customized analysis is not lost as all the data is still kept available to the clinic alongside the visualization.

3.3 The questionnaire

3.3.1 Overview

The questions are designed to reflect the daily habits concerning health (particularly weight) of a patient, such as food consumption and other diet based variables. Each question is mapped to a subset of a 24-hour time period, and the answers will be used as input data for the visualization to evaluate a risk assessment of where the most undesirable habits occur according to the principles of the treatment.

3.3.2 Evaluation

An approach similar to that of IBM's scoring method was initially implemented based on the simplicity of its implementation and potential of expanding to a more complex approach. Providing a weight value for each item grants the clinic a trivial approach for adding a broad spectra of questions with differing importance without having to adapt the answer scores accordingly. Consider the following:

A value interval of 0-3 is provided for scoring each answer. This means a 2-choice question would probably have one "preferred" answer with a score of 0, and one "risky" answer with a score of 3. A 4-choice question would then have answers ranging from 0-3 depending on the risk they convey. If one question concerns a more harmful habit than others, the question is given a higher weight value instead of increasing the answer scores for that particular question. The question weight is later multiplied by the answer score in the evaluation of the questionnaire. In this scenario, the goal of the treatment is to reduce the resulting total score to a minimum.

Although this scoring method is a good starting ground, it lacks the dimension of mapping items to a time interval. The questionnaire implementation is therefore extended to include a time interval variable to factor the questions into their respective groups. Each question regarded as a contributor to the visualization requires one or more defined time intervals at creation to determine its effect on the result.

The questions were formulated with the help of the clinic personnel. A communications network was put in place between the developer of the questionnaire and a professional dietitian as well as their colleagues to assure the best quality possible for each question of the form. Since no conclusive research could be found on mapping diet habits to a timeline, the best possible result was ensured through regular consultation with professionals in the field.

3.3.3 Implementation

The questionnaire was developed as an extension of the companies current platform for the clinic personnel. It was written in React.js using the formik library for dynamic questionnaire functionality. The questionnaire connects to a JavaScript backend which stores the data as JSON files, from where data can also be fetched to edit or view existing questionnaires using unique ID's.

Clinic side

Upon questionnaire creation, the evaluator may choose from 5 different kinds of questions, namely

1. Question Matrix
2. Graded Question
3. Yes/No Question
4. Free-text Question
5. Multiple-choice Question

Where only two are used for the visualization: *Yes/No* and *Multiple-choice* questions. The other question-types are not feasible for analysing in this scope based on the simple fact that free-text questions and matrices introduce too much complexity to both the algorithm and the evaluator. Evaluating free-text questions algorithmically would require complex methods of the natural language processing nature, most likely still not meeting the precision requirement of this thesis. Regarding the matrix question, this is a format not widely used overall, which deems it unworthy of putting effort into making compatible with the thesis. Furthermore - all questions can be transformed to a multiple-choice question with little effort and practically no loss of information.

Write about how the questions/answers are scored and graded and that it is also stored in the backend with the rest of the questionnaire data under the same ID.

When formulating a question, a question title and its answers are expected. The answers all have a drop-down attached that allows the evaluator to score each answer from 0-3 points. If no score is chosen, submission of the question is not allowed. This is to guarantee that the scoring of each question is done with some consideration, avoiding forgetting to put a score which might render the results invalid or inaccurate.

Apart from scoring the answers, another drop-down is provided below the answers that is used for defining the question weight. Choosing a weight is also required as it determines the impact that question will have on the final visualization. The drop-down allows for weighs ranging from 0-3, where 0 indicated that the question is not to be taken into account in the visualization, and 1-3 indicates the amount of impact this question has on the final result compared to others.

Finally, a tag-system is put in place to allow choosing the time-interval(s) a question belongs to. The tags are hard-coded to match the visualization intervals with human-readable text to simplify the creation process. They range from "Early morning" to "Night" and several tags are allowed to be chosen. Choosing none of these tags are not allowed unless the question is defined as not being used in the visualization, as the visualization algorithm needs information about which interval to score.

All the variables used in the visualization procedure is stored in the same database as the rest of the questionnaire.

Client side

The distribution of the questionnaire is done by mail where the respondent clicks a link redirecting to an answer-page. The answers are stored in the same backend database as the questionnaire and can be fetched with a user-ID to allow the user to edit or review their answers afterwards.

3.4 The visualization

3.4.1 Overview

Creating a custom visualization fit for the unique purpose of this study means no other visualization of this exact kind has been attempted before. The idea of visualizing data in a customized timeline, not as a standard graph or plot of measurements, has not been researched in great detail and mostly occur in sporadic graphical experiments scattered around the web. Some of the design choices are thereby based local research conducted such as a design preference survey.

The choices of design and implementation are all made with the intention of providing a visualization that maintains simplicity while relaying relevant information to the client. Local studies conducted are combined with official

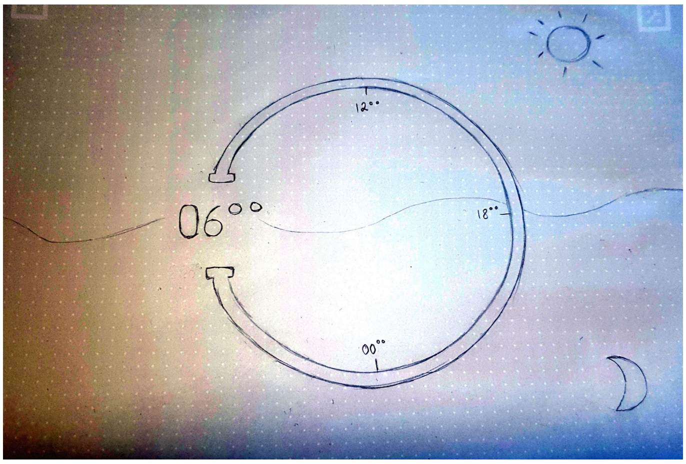

Figure 3.1: First sketch of potential design

research to find an optimal approach for the visualization to fulfill its purpose of being intuitively understandable from both the client and the evaluators perspective.

3.4.2 Design

Shape

A circular expression of the visualized timeline is preferred based on the intuitive loop-like structure of a 24-hour day. This design can be compared to that of an analogue clock - the last hour of one period is always followed by the first hour of another. This way, the client is expected to interpret the visualization as re-occurring days in their every day life intuitively. Circular expressions also seem to be a re-occurring theme in other visualization approaches in the field (see figure 2.2-2.4), indicating that the general public will most likely be accustomed to the notion. The first sketch of a potential design can be seen in figure 3.1.

Intervals

A decision was made to have the day split into 8 equal sized time intervals to maintain both questionnaire development simplicity and visualization accuracy (See figure 3.2).

While having more intervals provides more detailed information, utilizing this property forces one to use very specific questions mapped to small time-frames. This has two side-effects - it limits the kinds of questions usable in the visualization and it increases the likelihood of notifying the patient of the intent of the questionnaire. Giving questions a wider interval provides more wiggle room in the way you ask questions - eg *"Do you walk home from school?"* could be mapped to the afternoon hours 15-18. This poses less risk of alerting the patient that this question will directly impact their result, which could encourage lying to prevent the feeling of being exposed for their bad habits. Therefore, the more general and unsuspicious a question is, the better it is in terms of quality based on the project intent.

Although fewer intervals is considered better in the aspect of forming questions, having too few intervals disregards too much valuable information from the visualization. A compromise is therefore made to include as much information as possible without alerting the patients of the questionnaire's intent or limiting the freedom of formulating questions too harshly - namely 8 intervals.

3.4.3 Implementation

Concluding the appropriate approach

Figuring out the most appropriate libraries and API's to create the visualization is not an intuitive task. The key aspects considered when comparing visualization methods is artistic freedom (not being restricted to using predefined objects or figures) and simplicity (avoiding low-level programming interfaces such as WebGL/OpenGL that take a considerable amount of effort to fully grasp).

Considering the project aim of creating a web-based visualization, the JavaScript language is used based on the vast amount of helper libraries and API's available as well as its inherent web-based functionality. In regards to the purpose of maintaining simplicity in the design, most focus has been put on researching 2D graphical libraries as a 3D approach brings unnecessary complexity to the implementation as well as to the comprehension of the result.

After testing and evaluating potential approaches, some conclusions can be drawn about the implementation preferred for the visualization. Libraries such as Pixi.JS, D3.JS, and Victory.JS was explored among others, as well as the WebGL API. Considering the unique nature of this visualization, the most appropriate choice appears to be using HTML Canvas in the context of

JavaScript. This setup brings minimum design restriction while abstracting the implementation to a simple enough level for a programmer lacking the full graphics coding experience. HTML Canvas is known to work well with D3.JS, although this mainly applies to visualizations representing big data. Since this visualization only handles one questionnaire instance at a time, there is no need to introduce another library with the risk of further complicating things.

Retrieving the data

An API call fetches data based on a user-ID. The data is filtered and grouped into tags in the backend before returning the JSON. The data received therefore consist of tags mapped to questions, mean answer score, and mean total score for that tag. *See figure blabla for example*

Manipulating the data

Before drawing, the data traverses some further calculations:

- The data is sorted and indexed from morning to night (followed by full-day)
- The "full-day" tag value is spread out to affect the full active daytime as the tag is removed from the data
- The night-tag is duplicated to cover the full night (2 intervals from 00.00-06.00)
- The mean values are normalized to a number between 0-1 where 1 is the worst possible risk over all the intervals. This value is used to color the visualization

Drawing the visualization

The cyclic visualization is rendered through a series of drawn arcs, representing one time interval each (See figure 3.2). This was done through using the embedded arc() function provided by HTML Canvas, making the drawing of arcs trivial with the use of some simple arithmetic techniques based on the Cartesian coordinates provided by Canvas.

The numbers are drawn using vector calculations on the circle edge, placing them a few coordinates from the circle edge along the reversed norm vector, towards the center of the circle. By using simple dynamic coding, a variable

Figure 3.2: Empty interval representation

controlling the number of intervals can be changed freely and the numbers will appear in the correct positions, as long as the number of intervals is a divisor of 24.

Coloring

Regarding the color of the design, a shade of red is used for areas of risk and green for the opposite based on research implying that red is affiliated with danger and therefore intuitive to avoid [24].

A simple gradient calculation is used for transitioning colors. The output color for an interval is put in the exact middle of that interval, and a gradient between each interval is produced to get a smooth color transition throughout the circle.

Glow

A glow is rendered to expand from the edge depending on the interval value. The more risky (red) an area is, the stronger is the glow. This is to alert the patient of relevant areas while also providing a smooth blur gradient to the visualization. Note that the glow is rendered before the circle itself to make sure the glow does not intrude on the coloring gradient of the circle *(Picture here of that mistake)*. This is avoided by making sure the circle overwrites the glow by drawing it secondary.

The glowing effect is controlled by incrementing/decrementing a global variable in the draw loop that switches direction at a certain threshold. The increments are divided by the fps to guarantee maintaining the same animation speed across environments with differing fps.

Figure 3.3: An example of the visualization with low standard deviation and tuning disabled

Tuning

To use the normalized weight values directly gives an overall bland result. There is a clear difference between a very risky and a very safe interval, but if several intervals have somewhat alike values the naked eye cannot differentiate which interval is the riskiest (See figure 3.3). This is especially true for values at each end of the risk spectra (high or low risk), as the gradient requires more deviation at the border-cases to show a noticeable difference.

Furthermore, the questionnaire is meant to pinpoint weaknesses in dietal and lifestyle habits. Having any score above 0.1 on an interval is a sign of risk, and should be treated as such. Therefore, having someone get an average score of, say 0.5, is considered unhealthy and should not get a visualization that has no red areas, giving the impression that the clients habits are nothing to be concerned about (See figure 3.3). On the same note, a person with very harmful habits should not be dis-encouraged by being presented with a fully red glowing visualization as hope is considered an important factor in client motivation and has been shown to benefit treatment retention [39].

All of these considerations leads to some variables and algorithms whose sole purpose is to tune the visualization dynamically to provide a fair and relevant visualization in all cases, providing:

- **Hope**
  For the overwhelmingly harmful results

- **Clear signs of risk**
  For the mediocre results not reaching the red part of the color scale

- **Hills and valleys**
  For the low-variance results with no obvious change in original color values

- **Reliability**
  Despite all the tuning, healthier results should still provide a greener average color value than unhealthier ones. Reaching a fully green circle should, of course, be possible

Below, some examples of the visualizations with- and without tuning are presented for comparison.

Testing

Testing a solution of this sort is troublesome to say the least. The long-term effect of applying the visualization in treatment cannot be reliably measured in short-term testing, and the perceptive satisfaction or knowledge gained from the side of the client is subjective and could prove difficult to directly compare with current methods of treatment evaluation. Furthermore, expecting all 300 clinic patient families to partake in a user study concerning such a delicate subject as obesity of which they are affected by daily is not feasible. And if these circumstances are not bad enough - the COVID-19 pandemic hit the world by storm not long before the testing process was planned to take place, further halting the possibility of making qualified user test cases through interviews or live display of the visualization.

However, this does not stop a scientist of civil engineering caliber from squeezing out whatever valuable insights can be unraveled in this turmoil of opposition.

Considering the solution mainly being used in the evaluation process of treatment - if one could intercept the steady stream of new patients doing their first clinic evaluation, one could provide them with a personalized visualization based on their questionnaire results followed by an interview. As meeting with the clinic personnel is mandatory to begin treatment, having one more person attending within a safe distance for testing purposes will not be a noticeable concern for the clinic nor most of the patients. This provides quality user testing, although limited in the sheer quantity of subjects. Quantitative measurements were certainly needed to arrive at a sensible conclusion, widening the scope to the public. A general survey was made available and distributed to *"XXX"* people, having them rate and compare different methods of clinic evaluation on test data, including the visualization of this thesis.

Now, having both qualitative and quantitative data despite of the arguably dire circumstances, results could be provided and conclusions could be made.

Chapter 4

Results

Everything proved to go well. All the patients are happy now. I am also happy; this was a good thesis.

Chapter 5

Discussion

Why was this thesis so good? I think it was good because everything went well.

Chapter 6

Conclusion

Great job!

Chapter 7

Glossary

MBTI = Myers-Briggs Type Indicator
MMPI = Minnesota Multiphasic Personality Inventory
MMPI-2 = Minnesota Multiphasic Personality Inventory (revised)
MMPI-A = Minnesota Multiphasic Personality Inventory (adolescents)
IRT = Item Response Theory
Item = One question in a questionnaire
Quantitative questionnaire = Questionnaire consisting of multi-choice alternative questions Qualitative questionnaire = Questionnaire consisting of multi-choice and/or free-text questions
Single-response questionnaire = A questionnaire uniquely analyzed per subject response
Multi-response questionnaire = A questionnaire where many or all the responses are evaluated together to reach mean values and general statistics of a group
HRQOL = Health Related Quality Of Life
Evaluator = Professional at the clinic or suggested field, responsible for evaluating patients

Chapter 8

Appendix

Pictures of the questionnaire/questionnaires here etc

Chapter 9

Source criticism

The sources are gathered from renowned organizations such as IEEE and the Google Scholar online library. They consist of published and cited papers of which there is no significant reason to doubt the quality. While some papers might be cited more than others, the authors are all of a credible scientific caliber and have a history of research in their respective fields reaching further than the common student.

There are two exceptions, the IBM questionnaire template and the Swedish State's Committee For Medical and Social Evaluation article.

Even though IBM do not have the scientific credibility as the other sources, their methods are considered reliable enough for testing. This should be done with emphasis on the fact that there is no prior scientific validity in their method.

The Swedish State Committee is part of the Swedish state, and there is no well-founded reason to doubt their credibility as it is not a private company or organization. However, as in the case of the IBM scoring method this method should be thoroughly tested and evaluated before being used as a base for scientific study.

Bibliography

[1] Bas Verplanken and Suzanne Faes. "Good intentions, bad habits, and effects of forming implementation intentions on healthy eating". In: *European Journal of Social Psychology* 29.5-6 (1999), pp. 591–604.

[2] Jordi Cami and Magı Farré. "Drug addiction". In: *New England Journal of Medicine* 349.10 (2003), pp. 975–986.

[3] C.D. Hansen and C.R. Johnson. *Visualization Handbook*. Elsevier Science, 2011. ISBN: 9780080481647. URL: https://books.google.se/books?id=mA8ih1AieaYC.

[4] Ben Shneiderman, Catherine Plaisant, and Bradford W Hesse. "Improving healthcare with interactive visualization". In: *Computer* 46.5 (2013), pp. 58–66.

[5] Andrew Wilson and Susan Childs. "The relationship between consultation length, process and outcomes in general practice: a systematic review." In: *Br J Gen Pract* 52.485 (2002), pp. 1012–1020.

[6] Farnaza A Ahmad BA Khairatul K. *An assessment of patient waiting and consultation time in a primary healthcare clinic*. 2017.

[7] R.M. Groves et al. *Survey Methodology*. Wiley Series in Survey Methodology. Wiley, 2011. ISBN: 9781118211342. URL: https://books.google.se/books?id=ctow8zWdyFgC.

[8] R. Applegate. *Practical Evaluation Techniques for Librarians*. ABC-CLIO, 2013. ISBN: 9781610691604. URL: https://books.google.se/books?id=TdiOAQAAQBAJ.

[9] Robert H Gault. "A history of the questionnaire method of research in psychology". In: *The Pedagogical Seminary* 14.3 (1907), pp. 366–383.

[10] Robert H. Gault. *A History of the Questionnaire Method of Research in Psychology*. Sept. 1907. DOI: 10.1080/08919402.1907.10532551. URL: https://doi.org/10.1080/08919402.1907.10532551.

[11] I. Holloway and K. Galvin. *Qualitative Research in Nursing and Healthcare*. Wiley, 2016. ISBN: 9781118874479. URL: https://books.google.se/books?id=66PIDAAAQBAJ.

[12] Susan E Embretson and Steven P Reise. *Item response theory*. Psychology Press, 2013.

[13] R.K. Hambleton, H. Swaminathan, and H.J. Rogers. *Fundamentals of Item Response Theory*. Measurement Methods for the Social Science. SAGE Publications, 1991. ISBN: 9780803936478. URL: https://books.google.se/books?id=gW05DQAAQBAJ.

[14] *IBM Template Scoring Formulas*. https://www.ibm.com/support/knowledgecenter/SSFUEU_8.1.0/op_grc_user/c_op_qr_about_scoring.html.

[15] Laura Manea, Simon Gilbody, and Dean McMillan. “A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression”. In: *General hospital psychiatry* 37.1 (2015), pp. 67–75.

[16] “Patient Health Questionnaire-9 (PHQ-9) som stöd för diagnostik och bedömning av svårighetsgrad av depression”. In: ().

[17] I Elaine Allen and Christopher A Seaman. “Likert scales and data analyses”. In: *Quality progress* 40.7 (2007), pp. 64–65.

[18] Rensis Likert. “A technique for the measurement of attitudes.” In: *Archives of psychology* (1932).

[19] Harry N Boone and Deborah A Boone. “Analyzing likert data”. In: *Journal of extension* 50.2 (2012), pp. 1–5.

[20] H.H. Harman and University of Chicago. *Modern Factor Analysis*. University of Chicago Press, 1976. ISBN: 9780226316529. URL: https://books.google.se/books?id=e-vMN68C3M4C.

[21] Yah-Ling Hung and Catherine Stones. “Visual Design in Healthcare for Low-Literate Users–A Case Study of Healthcare Leaflets for New Immigrants in Taiwan”. In: *International Conference of Design, User Experience, and Usability*. Springer. 2014, pp. 44–55.

[22] Hannah Brotherstone et al. "The impact of illustrations on public understanding of the aim of cancer screening". In: *Patient education and counseling* 63.3 (2006), pp. 328–335.

[23] Guillaume Thierry et al. "Unconscious effects of language-specific terminology on preattentive color perception". In: *Proceedings of the National Academy of Sciences* 106.11 (2009), pp. 4567–4570. ISSN: 0027-8424. DOI: 10.1073/pnas.0811155106. eprint: https://www.pnas.org/content/106/11/4567.full.pdf. URL: https://www.pnas.org/content/106/11/4567.

[24] John R Anderson. *Cognitive psychology and its implications*. Worth publishers, 2000.

[25] Amélie A Walker. "Neolithic surgery." In: *Archaeology* 50.5 (1997), pp. 19–19.

[26] Massimo VIDALE and Roberto MACCHIARELLI. "Dental lesions on the permanent teeth at Neolithic Mehrgarh, Pakistan". In: ().

[27] Norman Daniels. "Justice, health, and healthcare". In: *American Journal of Bioethics* 1.2 (2001), pp. 2–16.

[28] Yair G Rajwan and George R Kim. "Medical information visualization conceptual model for patient-physician health communication". In: *Proceedings of the 1st ACM International Health Informatics Symposium*. 2010, pp. 512–516.

[29] Stuart Card, JD Mackinlay, and B Shneiderman. "Information visualization". In: *Human-computer interaction: Design issues, solutions, and applications* 181 (2009).

[30] Tammy L Bess, Robert J Harvey, and Dana Swartz. *Hierarchical confirmatory factor analysis of the Myers-Briggs Type Indicator*. 2003.

[31] *MBTI Basics*. https://www.myersbriggs.org/. 2020.

[32] James N Butcher. "Minnesota multiphasic personality inventory". In: *The Corsini Encyclopedia of Psychology* (2010), pp. 1–3.

[33] S. Whitcomb and K.W. Merrell. *Behavioral, Social, and Emotional Assessment of Children and Adolescents*. Taylor & Francis, 2013. ISBN: 9781136737251. URL: https://books.google.se/books?id=LUpEArMgQSsC.

[34] Nathan R Kuncel, Sarah A Hezlett, and Deniz S Ones. “A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance.” In: *Psychological bulletin* 127.1 (2001), p. 162.

[35] James W Varni et al. “The PedsQL™ family impact module: preliminary reliability and validity”. In: *Health and quality of life outcomes* 2.1 (2004), p. 55.

[36] James W Varni et al. “The PedsQL™* 4.0 as a pediatric population health measure: feasibility, reliability, and validity”. In: *Ambulatory pediatrics* 3.6 (2003), pp. 329–341.

[37] Bradford W Hesse et al. “Social participation in health 2.0”. In: *Computer* 43.11 (2010), pp. 45–52.

[38] Thea Runyan. *Kurbo: A Digital Health Solution for Overweight Youth*. 2018.

[39] Ashley S Hampton et al. “Pathways to treatment retention for individuals legally coerced to substance use treatment: The interaction of hope and treatment motivation”. In: *Drug and alcohol dependence* 118.2-3 (2011), pp. 400–407.