Bachelor's Thesis

Website Navigation Structures:

Eliciting Mental Models of Dutch E-Government Websites Using Card Sorting

Jona Sabrina Kunert
Email: j.s.kunert@student.utwente.nl
Student Number: 2307820
July 4, 2022

Faculty of Behavioural, Management and Social Sciences
University of Twente, Enschede, Netherlands

1st Supervisor: Marlise Westerhof
2nd Supervisor: Dr. Martin Schmettow

Abstract

The purpose of this research is to give design recommendations for a navigation structure for Dutch e-government websites based on the GGD. A navigation structure that is congruent to the mental model of the average user enables them to find information on websites easily and intuitively (Schall, 2014). Therefore, the mental model of the average user had to be elicited which was done by using a remote, open card sort. Card sort studies give an insight into the way participants would expect to find information on a website. After applying participant exclusion criteria to the data of initially forty-five participants, thirty-seven participants were included in the study, of which nine were Dutch and twenty-eight were from other countries. In general, no large difference between the proposed and the existing navigation structure could be determined. Instead, minor changes are suggested which could positively influence the browsing performances of users. In addition, the study aimed to analyse the influence of the cultural background on the mental model of the user. Although one could argue that the small sample size makes the results less representative, intriguing differences in the mental model were detected, which are valuable for further research.

Contents

Introduction

Daily, enormous numbers of web users are confronted with large amounts of information when visiting websites, making it difficult to find the desired information. This is especially relevant during situations like the current COVID-19 crisis in which large quantities of information and contrasting opinions are published every day. To find crucial information about keeping themselves safe, users can access e-government websites. These provide users with information like quarantine regulations or the COVID-19 vaccine. The institution in the Netherlands that is responsible for sharing this health-related information is the GGD (Dutch: *Gemeentelijke Gezonheidsdienst*), the Municipal Health Service in the Netherlands. Municipal health service websites are a specified form of e-government websites that focus on health. There are 25 GGD websites in total, which was determined by looking at the official ggd.nl website. Their tasks include health promotion, medical screenings, and STD testing. Since there are many different GGD websites which vary in structure and design, it is even more crucial to propose a common navigation structure to have consistency between each website that makes browsing for information as easy as possible.

Enabling users to efficiently navigate through websites and intuitively gather information is a vital element of user experience (Schall, 2014). For that, the common user's mental model has to be considered when designing the navigation structure (Schmettow & Sommer, 2016). While the study by Branaghan et al. (2011) indicates that considering the users' expectations can improve a user interface, Schmettow and Sommer (2016) could not confirm these results in their study. That is especially relevant on e-government websites since users are browsing for health-related behaviour indicating that they might be under stress. Although it can be expected that web users would not search for information on the internet when they are dealing with a health emergency, but rather consult a doctor immediately, not finding the desired information can risk the health of users. The current study tries to investigate several research questions related to the beliefs of users visiting Dutch e-government websites allowing web designers and researchers to gain a deeper insight into the navigation behaviour of different users. This includes eliciting the mental model and based on that, making design recommendations for a user-centered website structure for the mentioned domain. In the following paragraphs, background information and exact details of the research question will be elaborated on.

Mental Models

When visiting a website, users have certain expectations of where to find specific information guiding them while browsing, referred to as the mental model of the user. These expectations are formed from past experiences that help people to understand how things work in

the external world (Roth et al., 2010; Schall, 2014). To be more specific, the mental model of a user describes where the user believes to find information on a website (Nielsen, 2010b). This information has been used by many researchers to make design recommendations regarding the navigation structure of a website (Faiks et al., 2000; Roth et al., 2010). Card sorting is one method that can be used to elicit the mental model. In a card sorting experiment, participants are given cards that contain items from a website. The participants are then asked to sort these cards into groups how they would expect to find them on a website and to label them accordingly. With this method, researchers and web designers can analyse where on the website the average user would expect to find the desired information. A detailed description of the different types and the exact procedure of card sorting will be given later throughout the paper.

Congruency between a website's navigation structure and the users' mental model enhances efficiency by facilitating orientation, helping to memorise the location of navigation components and by preventing errors (Oulasvirta, 2004). Since every individual has their own mental model, it is challenging to design a navigation structure that is perfect for each individual (Nielsen, 2010b). Thus, the aim is to recommend a navigation structure that represents the average user's mental model. Errors while browsing for information are more likely to happen if no attention is paid to congruency, which can lead to users being frustrated and disoriented (Otter & Johnson, 2000), making it difficult to find the information needed. That is especially problematic when looking for health-related information since the users are looking for crucial information and not simply, for example, for leisure-related information. To avoid this, eliciting mental models has become an essential step for efficiently designing the navigation structure. Domains for which the mental model has been elicited include municipal websites (Schmettow & Sommer, 2016), e-commerce websites (Bernard & Sheshadri, 2004), and digital libraries (Faiks et al., 2000). No data could be found on mental models specifically for municipal health service websites. For this reason, a goal of this research is to suggest an improved navigation structure for Dutch municipal health service websites.

Prior research has investigated the effect of demographic factors on mental models. Kralisch et al. (2005) demonstrated that cultural backgrounds influence navigation behaviour (i.e., time, amount of accessed information, and (non)linearity of information access). Similarly, Dam et al. (2003) reported that there are differences in user experience and behaviour on websites, which indicates that the mental model differs based on cultural background. Alexander et al. (2017) hypothesized that a navigation structure that is in line with the mental models of users from different cultures could enhance the performance of navigation behaviour. This is in line with the research by Nawaz (2013) who claims considering demographic diversity is essential

when designing a navigation structure. This paper focuses on Dutch websites because the Netherlands is known for its cultural diversity (Dam et al., 2003), and website users from different cultures that reside in the Netherlands should also be able to efficiently browse for information to stay up to date about current situations. Thus, this research aims to investigate the relation of cultural differences in relation to mental models of websites.

Influence of Navigation Structure on Usability

The e-government has been persistently making an effort to increase the poor usability of their websites - without success (Kotamraju et al., 2012; Lyzara et al., 2019). Usability can be defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO, 1998). A reason for their lack of improvement is that the users' needs are often not considered when designing a website (Kotamraju et al., 2012). E-government websites are obligated to design websites that provide a large amount of information which needs to be presented in the best way possible for all possible users, which includes, for example, users speaking different languages and users with visual impairments (Kotamraju et al., 2012), making it challenging to design one website optimally designed for everyone. This paper tries to investigate the expectations users of e-government websites have when browsing for information using a user-centred approach to resolve this issue.

There are different strategies for searching for information on a website. The majority of users choose to browse the navigation structure over using the search function to find the desired information (Katz & Byrne, 2003). Previous research has shown that the usability of a website is strongly influenced by the navigation structure (Rosenfeld & Morville, 2002): While a good navigation structure can increase the usability of users (Rosenfeld & Morville, 2002), an inadequate structuring, and labelling of concepts on a website negatively impact the usability (Nawaz, 2013). The results by Oulasvirta (2004) showed similar findings indicating that placing navigational regions at locations where users expect them facilitates the memorability of web pages. A negative impact that can result from an unstructured navigation structure can lead to ineffective (i.e., not finding desired information) and inefficient (i.e., taking longer to find desired information) browsing behaviour. Otter and Johnson (2000) specified these impacts by stating that poor usability leads to frustration, lostness and disorientation. Thus, attention was paid to investigating the navigation structure of e-government websites.

Card Sorting

Card sorting is a user-centred method that can be used to understand the mental models of users. It is frequently used in various areas of expertise, for example, web design, software

development, and designing navigation structures to understand the clustering of information from the users' point of view (Schmettow & Sommer, 2016). The main principle of card sorting is that participants receive an unsorted pack of cards which they have to group. Furthermore, the participants are asked to label the groups they formed themselves in the previous step. This is the most common type of card sorting and is referred to as *open card sorting. Closed card sorting* is another main type of card sorting in which participants match the cards to groups that were predefined by the researchers (Schmettow & Sommer, 2016).

Traditionally, card sorting was performed on paper, referred to as *paper card sorting* (Bussolon et al., 2006). Especially when many cards are used for the study, errors occur easily and negatively influence the results of the study (Petrie et al., 2011). Several studies compare the traditional card sorting version to *digital card sorting* and could not determine any differences in performance and reliability (Bussolon et al., 2006; Petrie et al., 2011). In their study, Bussolon et al. (2006) mention numerous advantages of digital card sorting in contrast to paper card sorting, including faster recruitment of participants which enables researchers to increase the number of participants and therefore improve the quality of the card sorting. Furthermore, digital card sorting allows for better classifications due to the possibility to test a greater number of participants with reduced costs. A downside of digital card sorting mentioned by Sherwin (2018) is that the usability of the card sorting program can influence the completion of the experiment. For example, technology problems can lead to frustration or hinder the performance of the participant (Sherwin, 2018). This research focuses on digital open card sorting

Summarised Research Aims

This research paper aimed to propose a navigation structure for Dutch e-government websites to make browsing for information more efficient. To make a design recommendation, a set of research questions was answered. First, the mental model of young adults visiting Dutch e-government websites was elicited to determine which navigation structure is the best for the mentioned domain. This was done by conducting a remote, open card sort. Second, the researchers aimed to find out if it can be confirmed that the mental model differs between users with different cultural backgrounds. To answer this research question, the mental model of Dutch participants was examined in contrast to the one of other participants creating heatmaps. Finally, it was tested if the mental model applies to other websites by comparing it to websites of the e-government domain. This was done by applying the mental model to the existing navigation structure of a GGD website.

Methods

Participants

Forty-five participants were recruited, out of which 59.46% signed up for the study through SONA systems and received a course credit. SONA systems is a psychology test subject pool of the Faculty of Behavioural, Management and Social Sciences (BMS) of the University of Twente. Convenience sampling was used to recruit the other 40.54% of the participants who accessed the study through an online link that the researchers shared with their relatives.

Participant exclusion was used to exclude certain participants to ensure that the results are reliable since an uncleaned dataset can negatively affect the results of a study (Righi et al., 2013). It is a challenging task to exclude data from certain participants because it might be unclear if participants did not take the experiment seriously, or if their mental models simply deviate from that of other participants. The following exclusion criteria were created: First, closer attention was paid to participants that created an amount of categories that strongly varies from the average amount of categories created, especially when very few categories were created. For this criterium, it was important to determine why many items were grouped into one category. In their paper, Righi et al. (2013) explained that using a few categories can be a sign of expert knowledge, but it could also indicate that participants lost motivation throughout the experiment and randomly assigned items to groups. Second, attention was paid to the quality of the labels assigned to categories. When it was obvious that meaningless or random labels were created, the data of that participant had to be excluded from the study.

The final set of participants included 37 participants, of which 25 were female, 11 were men, and one identified as non-binary. Of these participants, 78,38% were aged 18 – 24 years, 16.22% were aged 25 – 34 years, and 5.41% were aged 45 – 54 years. Nine of the participants were Dutch, 16 were German, and 12 originated from other countries. To be able to take part in the study, participants had to be above 18 years old and be able to understand English. The research was approved by the University of Twente BMS Ethics Committee of the University. All participants received and agreed to the informed consent.

Item Selection

The cards for the card sorting experiment were created by completing a series of steps. First, a selection of GGD websites was made on which the set of cards should be based. To select GGD websites that were going to be assessed, the website ggd.nl was visited which shows a map with all municipalities that have a GGD website. The websites were chosen based on the number of residents each municipality has, as well as how well they are known (i.e., well-known websites with a high number of residents were included). Moreover, after an initial look at the

navigation structure of each, it was clear that these vary across each website. Since this research aims to determine which navigation structure would be most efficient for users, websites with a variety of different navigation structures were chosen. For instance, while the GGD Groningen website has a first level that contains broad categories (i.e., '*Child and upbringing'*, '*Health*', '*Healthy while travelling*', '*Sexuality*', '*Healthy living in Groningen*'), the homepage of GGD Amsterdam has a large list of very specific categories (e.g., '*Care for refugees from Ukraine*', '*Travelers*', '*Hygiene*', '*Lead waterpipes*', '*Tuberculosis*'). A selection of ten GGD websites was made (see Appendix A) that were analysed in the current paper.

In the following step, relevant items for the cards were chosen. The chosen ten websites were divided between the researchers. Each researcher scanned five websites for the most relevant items and wrote these down in a list. One reason to consider an item as relevant was when it appeared on several websites. Sometimes, an item only occurred on a few websites but was still written down as relevant when it seemed that this topic was not part of any other item yet, for example, '*Analysing drug content*' and '*Homelessness*' only appeared on very few websites. When creating the list, special attention was paid to the site map of the website if one was provided. Site maps are used to give web designers and visitors an overview of the content of the website (Alexander et al., 2017). This simplified the process of looking through each level of the navigation structure of the website.

Next, the researchers compared the lists they have created individually. Items were included if they occurred on both researchers' lists or if they only appeared on one list but seemed relevant. An example of an item that only occurred on one list is '*Domestic violence*', which was included although it did come up on a minority of the GGD websites. It was considered a relevant item since it could be assumed that many people are affected by domestic violence and would search for help on an e-government website. Problems resulted from the website's main language being Dutch. Since the researchers were not native Dutch speakers and the research was going to be conducted in English, each item had to be translated. Due to the translation process, the researchers often had different words for similar items. Therefore, the titles of items had to be combined so that they encompassed several words of the same category. Another challenge that was encountered is that some topics were sorted under different categories on different websites. To illustrate, '*Corona vaccine*' was sometimes found under '*Vaccines*' and sometimes under '*Corona*'. To see how participants would group these items, they were included as separate items. This enables us to find out, which website was closer to the mental model of the users.

In the last step, the set of cards had to be finalized. For each item, a definition was written down. The goal was to make the definitions as detailed as possible so that participants have a common understanding of the items. The final set of cards included 40 cards, including titles and definitions (see Appendix B). While previous research has shown that it is best to use 30 – 40 cards, keeping in mind the fatigue and the motivation of the participants (*Card Sorting*, 2013), other research stated that 40 – 80 cards should be chosen to have a set of cards that presents the main content of the website (Sherwin, 2018). Based on this insight, a set of 40 cards was created to ensure that the final set of cards leads to an accurate representation of the mental model, the cards were revised several times and went through multiple rounds of adjustments.

Materials

To conduct this study, a remote, online card sorting website called "KardSort" was used. The types of card sorting supported by the website are open-, closed-, and hybrid-card sorts. Moreover, a pre- and post-questionnaire, as well as a welcome page, an instruction page, and an acknowledgement page are included. An example of what the card sorting tool and the presentation of the cards look like can be found in Appendix C. After purchasing the premium version, there were no limits on the number of questions and cards that could be used. Moreover, it was possible to choose between an open or closed card sorting design. For this study, an open card sorting design is used. The collected data could be externally analysed with a variety of programs, for instance, SynCaps, Casolysis, and R.

Procedure

A pre-study questionnaire was prepared for the study. In the first question, the participants were asked if they agreed to the informed consent that they read on the previous page. The following five questions were about the participants' demographic information and website navigation proficiency. The last question asked for the SONA participant number. Answering this question was only required when participants accessed the study through SONA so that the researchers were able to trace back which participants completed the study and could therefore receive their credit.

The study took place remotely on the participants' own devices. Participants were able to access the website "KardSort" through a link that was shared. The online card sort study started with the informed consent, followed by the questionnaire and the instruction page (Appendix D). Afterwards, the card sorting itself was conducted. During the task, participants were presented with 40 cards, each of them containing a title and a description. Participants were able to create new categories and name them themselves. Next, the participants placed the 40 cards in these

self-made categories. At the end of the study, participants were shown an acknowledgement page thanking them for their participation.

Data Analysis

In the following step, the cleaned dataset was uploaded into SynCapsV3 to automatically create heatmaps, setting the foundation to elicit the mental models of users. To perform cluster analysis, a similarity matrix was created. In a similarity matrix, also referred to as the *items x items matrix*, similarity scores between any two items are calculated, representing the mutual semantic proximity of these two items (Schmettow & Sommer, 2016). A heatmap is a graphic depiction of a similarity matrix (i.e., the stronger the similarity, the stronger the colour) (Stroes, 2018). While the sorting procedure, conducted by participants, can be described as qualitative, the process of creating a similarity matrix by calculating similarity scores is quantitative.

After the heatmap was created, the mental model was obtained by analysing the clusters. When looking at a heatmap, darker colour represents stronger similarity and are accumulated on the diagonal of the heatmap. Lighter colours represent a weaker similarity between the two items. A benefit of heatmaps, in contrast to other card sorting analysis methods (e.g., dendrograms), is that it visualises off-diagonal clusters where an item can also fit, called *ambiguity*. These ambiguities can be manually identified since their colour deviates from the surrounding colours. It should be noted that this method is time-consuming and possibly inaccurate (Stroes, 2018). Various *cluster identification techniques* are applied to determine the clusters. The first technique that was applied is the visual proximity of dark blue cells. A group of darker cells implies that these items were grouped most frequently. Another technique that was used to identify clusters, especially to identify sub-clusters within main clusters with low similarity scores, is the researchers' semantic reasoning for the best fit.

Results

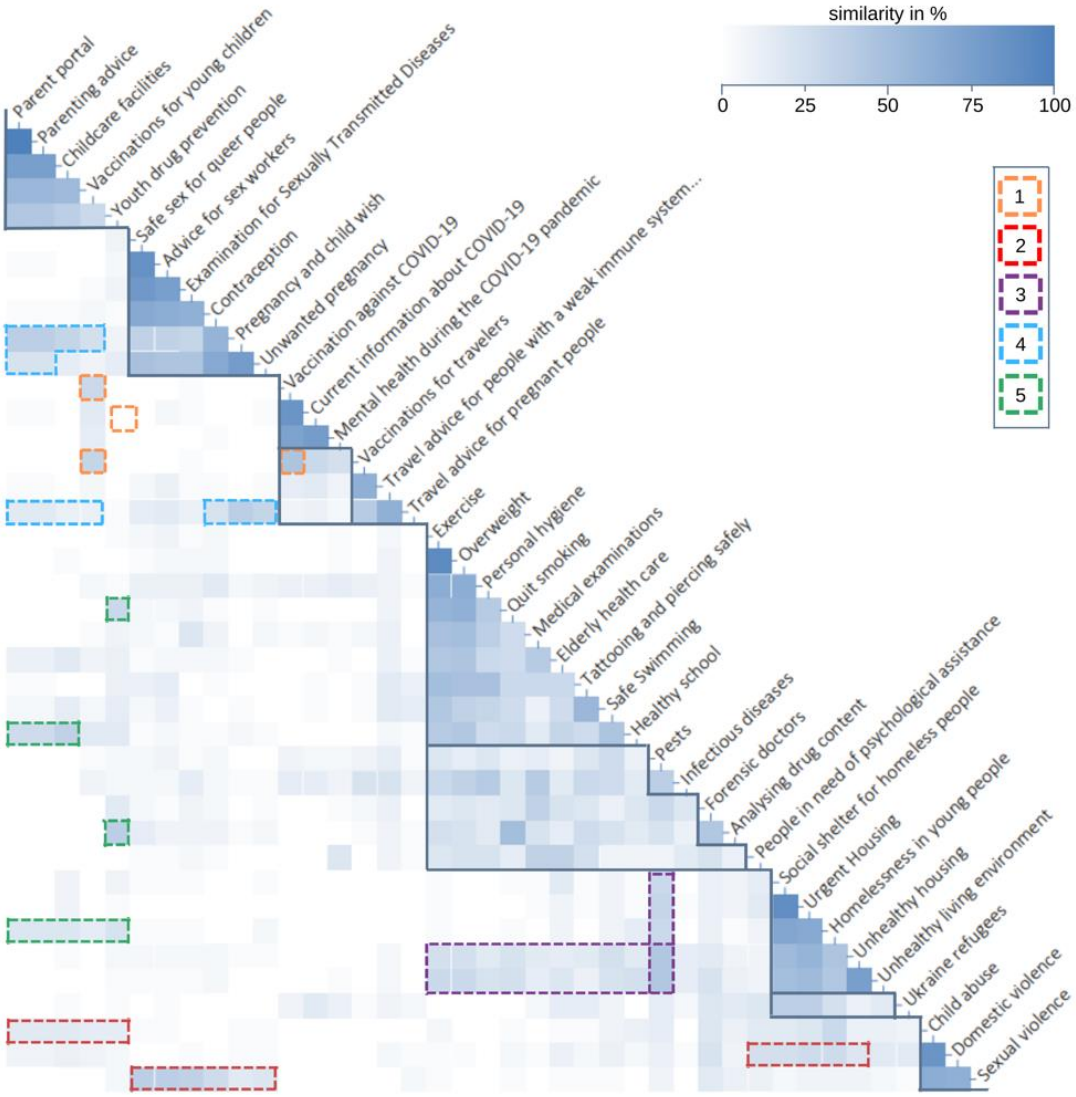
Cluster Analysis of General Heatmap

The first aim of this research paper is to elicit the mental model of users that visit e-government websites. To obtain an overview of the navigation structure user expect to see when visiting the website, a heatmap was created based on the card sorting results of all participants (see Figure 1), representing the average mental model of all participants. Based on the cluster identification techniques, a total of six main clusters were found in the heatmap. An overview of these clusters can be found in Table 1. Three out of these six clusters were further divided into smaller clusters. While five of these clusters were similar in size, containing between three and six items, Cluster 4 stood out due to its large size with 14 items. After closely examining the

cluster and applying semantic reasoning, the cluster could be subdivided into four smaller clusters. A more detailed description of each cluster will be given in the following paragraphs.

Figure 1

Mental model heatmap of all participants (n = 37) for the navigation structure of e-government websites.



Note. The continuous, dark blue outlines, which can be found along the diagonal, define the clusters of the heatmap, representing the mental model of the average user. The broken outlines define ambiguities and can be found off-diagonal. The different colours describe different ambiguity groups which are explained in more detail in the text. The stronger the blue colour of the squares, the stronger the similarity. A stronger similarity means that two items were grouped together more frequently.

Table 1

Summarised clusters of all card sorts.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
• Parent portal • Parenting advice • Childcare facilities • Vaccinations for young children • Youth drug prevention	• Safe sex for queer people • Advice for sex workers • Examination for STDs •Contraception • Pregnancy and child wish • Unwanted pregnancy	• Vaccination against COVID-19 • Current information about COVID-19 • Mental health during the COVID-19 pandemic • Vaccinations for travellers • Travel advice for people with a weak immune system • Travel advice for pregnant people	• Exercise • Overweight • Personal hygiene • Quit smoking • Medical examinations • Elderly health care • Tattooing and piercing safety • Safe swimming • Healthy school • Pests • Infectious diseases • Forensic doctors • Analysing drug content • People in need of psychological assistance	• Social shelter for homeless people • Urgent housing •Homelessness in young people • Unhealthy housing • Unhealthy living environment • Ukraine refugees	• Child abuse • Domestic violence • Sexual violence

Note. Lines in between item groups indicate that there are sub-clusters within a cluster.

Since an open card sort was conducted, participants were asked to label categories themselves. To specify, Table 2 shows some examples of labels the participants created during the card sorting study out of 371 labels that were created in total. The inclusion of labels in the table is based on the following rules:

(1) Similar labels which only differed in grammar or order of words were merged. For example, to category labels mentioned are '*Healthy Lifestyle*' and '*Health & Lifestyle*'. In this case, one of these labels was included to represent both of them.

(2) Additionally, extra labels were not mentioned in the table if there were semantically highly similar to other labels.

The final label for each cluster will be indicated in Table 2, as well as in the detailed description of each cluster in the next section.

Table 2

Participant's proposed category labels.

Cluster 1 '*Youth Health*'	Cluster 2 '*Safe Sex and Pregnancy*'	Cluster 3 '*COVID-19*' /'*Travel Advice*'	Cluster 4 '*Healthy Lifestyle*'	Cluster 5 '*Housing Issues*'	Cluster 6 '*Abuse and Violence*'
• Parenting • Youth Health • Childcare • Pregnancy and Childcare	• Pregnancy • Sex • Sex and Pregnancy • Sexual Health • Reproductive Health • Safe Sex and Diseases • Sexuality	• COVID-19 • Infections, Vaccinations & Covid-19 • Travelling • Travel Advice	• Examinations & Preventions • Healthcare • Healthy Living • Healthy Life • Healthy Advice • General Healthcare • Hygiene • Health & Lifestyle • Other	• Housing Issues • Living Situation	• Violence • Abuse and Violence • Threatening Situations • Trauma • Victims of Abuse • Help for Victims

Note. Cluster 3 is divided into 2 sub-clusters, indicated by the line in between both item groups. Cluster 4 is an exception: Since the results of this cluster are overall inconclusive, not every sub-cluster was labelled separately. Instead, one label was given for the main cluster as a whole.

Detailed Description of Clusters

The following paragraphs focus on a more detailed description of the clusters. Cluster 1 includes five items, namely '*Parent portal*', '*Parenting advice*', '*Childcare facilities*', '*Vaccinations for young children*', and '*Youth drug prevention*'. The similarity of these items gradually gets smaller starting at the first-mentioned item (Parent portal x Parenting advice = 100%; Vaccinations for young children x Youth drug prevention = 30%). Based on the labels proposed by the participants, a fitting name for this group would be '*Parenting*', '*Youth Health*', '*Childcare*', or '*Pregnancy and Childcare*'. Thus, as a final label, '*Youth Health*' was chosen since it most accurately describes the items of this category.

Cluster 2 consists of six items and includes the words '*Safe sex for queer people*', '*Advice for sex workers*', '*Examination for STDs*', '*Contraception*', '*Pregnancy and Child wish*', and '*Unwanted Pregnancy*'. These items have a high similarity score of approximately 60%, indicating that many participants frequently grouped those items. Proposed labels for this category were for instance '*Sex and Pregnancy*', '*Reproductive Health*', and '*Safe Sex and Diseases*'. The final label for this category is '*Safe Sex and Pregnancy*' since it includes all previously mentioned items.

Cluster 3 has a value of six items, which were further sub-divided into two clusters of three items. The first sub-cluster includes the items '*Vaccination against COVID-19*', '*Current information about COVID-19*', and '*Mental health during the COVID-19 pandemic*'. The second sub-cluster includes the items '*Vaccinations for travellers'*, '*Travel advice for people with a weak*

immune system', and '*Travel advice for pregnant people*'. The items of the first sub-cluster have a high mean similarity score of 78.33%. The mean similarity score of the second sub-cluster is 55%. A reason for the grouping of COVID-19- and travelling-related items could be that during the pandemic special travel regulations applied and participants, therefore, relate travelling to COVID-19. The division of the main cluster was not immediately clear since visual proximity indicated that the item '*Vaccination for travellers*' was often grouped with items of both sub-clusters. The final decision was made since the similarity score of the item was higher for the item '*Travel advice for people with a weak immune system*' (62%) than for the item '*Vaccination against COVID-19*' (30%) from the first sub-cluster. Furthermore, this sub-division was made due to semantic reasoning. The first sub-cluster exclusively contains items of the group '*COVID-19*', while the second cluster only contains items of the semantic group '*Travel Advice*'. A possible reason for the creation of subclusters is that due to the similar label, participants grouped these items without thinking about their actual concept (Faiks et al., 2000; Sherwin, 2018). For example, every item of the sub-cluster contains the word *COVID-19* and each item from the second sub-cluster contains the word *travel*. Because both subclusters are semantically different from each other, it was decided to create two sub-clusters. The final label of the group is based on proposed category names and demonstrates the final label.

Cluster 4 has 14 items and is, therefore, the largest cluster. The similarity scores within this cluster are rather low. Words in this cluster are for example '*Exercise*', '*Personal hygiene*', '*Medical examinations*', '*Safe swimming*',' *Pests*', and '*Forensic doctors*'. Since some similarity scores are noticeably higher than other similarity scores within these clusters, four sub-clusters were created. The items within the first sub-cluster seem to not have a high semantic relationship. For instance, '*Elderly health care*', '*Tattooing and piercing* safety', and '*Safe swimming*' are semantically unrelated and were still grouped. The low similarity scores support this observation indicating that these items have not been grouped frequently. This could be because the instructions invited participants to create a group labelled '*Other*' in case they are not sure about the meaning of an item. The fact that participants often proposed the label '*Other*' for this sub-cluster supports this idea. The second sub-cluster (Pests x Infectious diseases = 38%), as well as the third sub-cluster (Forensic doctors x Analysing drug content = 43%) include two items each creating rather small clusters. The fourth and last sub-cluster of this main cluster is a single item cluster only including the item '*People in need of psychological assistance*'. The highest similarity score of this item is 35% with the items '*Elderly healthcare*' and '*Medical examinations*', showing that the highest similarity score of this item is low in comparison to other similarity scores. Small sub-clusters like the previous occur since they were not grouped often

with any other items. Also, these clusters were labelled '*Other*', indicating that participants were not sure in what group to sort these items. As explained earlier, a reason for that could be that it was mentioned in the instructions that it is accepted to create a category labelled '*Other*' if it is believed that this item does not belong anywhere else. Another reason could be that the description or the title of the card is vague, so the participant did not understand the correct meaning of the item. In general, it can be said that the results for this cluster are overall inconclusive. The similarity scores are overall low, semantic reasoning can often not be applied, and proposed labels indicate that the results for this cluster are overall inconclusive. Due to the variety in semantic meanings of all items, it is challenging to decide on a label for the main cluster. Since all items are related to health and lifestyle, the final label '*Healthy Lifestyle*' was chosen for this cluster.

Cluster 5 contains six items, namely '*Social shelter for homeless people*', '*Urgent housing*', '*Homelessness in young people*', '*Unhealthy housing*', '*Unhealthy living*' and '*Ukraine refugees*'. The average similarity score of this cluster is relatively high with a score of 49.13%. The item '*Ukraine refugees*' stands out since it does not have a high similarity score with any of the other items. It was included in the main cluster due to the visual proximity of the dark blue cells, but it was sorted into a single item sub-category. When removing the item '*Ukraine shelter*' from the main cluster, the main cluster has a similarity score of 61% which is noticeably higher. The named item has a similarity score of 38% with the items '*Social shelter*' and '*Urgent housing*', which is the highest similarity score it has with any item. A reason for that is that due to the Russia-Ukraine war many refugees are seeking shelter. Therefore, the participant might have grouped the items since they currently seem to belong to the same category. The item '*Ukraine shelter*' was sorted into a single item sub-category because it currently has a higher relevance than the other items from this cluster. Proposed labels for the main category are '*Housing Issues*' and '*Living Situations*'. As a final, it was decided on '*Housing Issues*'.

Cluster 6 consists of three items, namely '*Child abuse*', '*Domestic violence*', and '*Sexual violence*', that all have a high mean similarity score of 70.33%. An explanation for the grouping since they are semantically highly related: All these items are related to abuse. Besides, the description of each of the items includes the information that cases of violence can be reported on the website which might have also led to participants grouping these items. Several different labels were proposed for this category, for example '*Abuse and Violence*', '*Threatening Situations*', '*Trauma*', and '*Help for Victims*'. The final label for this cluster is '*Abuse and Violence*'.

Detailed Description of Ambiguities

Five ambiguity groups can be determined when analysing the heatmap in Figure 1, but only those will be mentioned that are beyond a similarity score threshold of approximately 16%. These ambiguities could be identified since their colour is stronger than the colour of the surrounding blocks. This is due to a higher similarity score of these items in comparison to adjacent blocks. The found ambiguities are located off-diagonally since two items have a high similarity score and are not part of the main clusters. Ambiguities that are part of the same group have the same colour (see Figure 1). In total, five ambiguity groups were found which will be described in the following paragraph in more detail.

The first cluster of ambiguities that will be discussed is the Ambiguity Group 1. Items that contain the word *vaccination* have a high similarity score, although they are not part of one main cluster. This indicates that while some participants grouped '*Vaccination against COVID-19*' with the COVID-19-group, *'Vaccination for travellers*' with the travel advice group, and '*Vaccinations for young children*' in the youth health group, items of the semantic group *vaccination* were also frequently grouped. This is an intriguing observation because there were purposely three items created that are highly semantically related. Although each of these items is about vaccinations, they all are about vaccinations for different groups. The goal was to see if one group was going to be created that is labelled '*Vaccinations*', or if each item is sorted into a different, semantically related group. The results of this study show that for most participants the latter is true. A minority of participants did decide to create an extra group for these items. This could indicate that there are users that would look for a separate group for vaccines. It can be expected that if the users cannot find that category after an initial look, they would automatically start searching for the information in the semantically related group.

Another group of ambiguities that will be analysed in more detail are items that are often matched with other items from the category '*Abuse and Violence*', marked as Ambiguity Group 2. Items included in this ambiguity are '*Child abuse*', '*Domestic violence*' and '*Sexual violence*'. The similarity score between the three items of this main group is very high (70.33%). Nevertheless, those three items were also sorted into other groups by many participants. The ambiguity indicates that '*Child abuse*' was often sorted into the category '*Youth Health*'; '*Domestic violence*' was often sorted into the category '*Housing Issues*', and '*Sexual violence*' was often sorted into the category '*Sexual Health*'. This is because all these items are also semantically related to the other groups. To specify, while the item '*Sexual violence*' was grouped most often with the items '*Child abuse*' and *'Domestic violence*' since they are all

related to the semantic group *'Violence'*, the item is also semantically related to the group '*Safe Sex and Pregnancy*' in that sense that '*Sexual violence*' can be related to sex being unsafe.

Ambiguity Group 3 makes it obvious that the items '*Unhealthy housing*' and '*Unhealthy living environment*' were often grouped with items of the category '*Healthy Lifestyle*'. Additionally, the item '*Pests*' was repeatedly sorted into the category '*Housing Issues*'. This makes sense since all these items have to do with the semantic group '*Living Environment*'. It demonstrates that there is an overlap in the groupings of the two previously mentioned groups since several single items of both groups are often grouped. For a possible navigation structure, it means that participants might have trouble locating the in this paragraph mentioned items since they could be sorted into different categories. A way to make it easier for participants to find the desired information is by giving the categories a clear label and putting both categories in close proximity on the navigation bar so that it is immediately clear in which category the information can be found. To specify, if there is a category labelled '*Healthy Lifestyle*' and another category labelled '*Housing Issues*', it will be clear that for example the items '*Unhealthy housing*' or *'Pests*' will not be found in the former mentioned category, but rather in the latter one. Another way to increase the usability is by creating a main category labelled '*Health*' and then creating sub-categories that specify if it is about the healthy or unhealthy living environment.

Additionally, Ambiguity Group 4 contains items related to pregnancy. Those items are spread over different main clusters. The ambiguities indicate that those items were often grouped with other items that are not part of the main cluster. Items this applies to are '*Pregnancy and child wish*' and '*Unwanted pregnancy*' from the category '*Safe Sex and Diseases*', and '*Travel advice for pregnant people*' from the category '*Travel Advice*'. This group of ambiguities indicates that there are users that would probably have trouble finding the mentioned items. Since the similarity scores of these ambiguities are rather low, it cannot be expected that severe problems will be caused due to this unclarity.

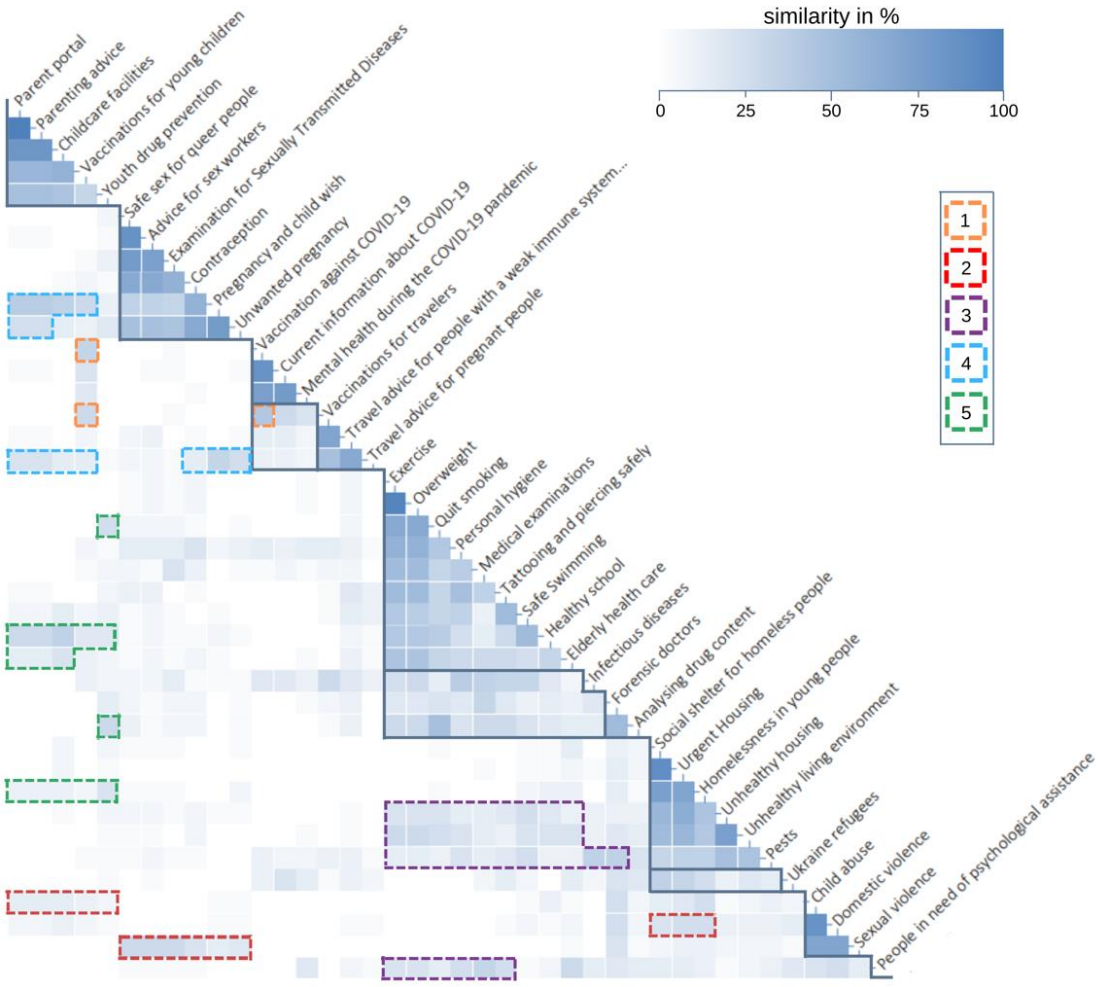
Last but not least, there is Ambiguity Group 5 which includes items that were frequently sorted into the '*Youth Health*' cluster. Those items are '*Quit smoking*', '*Healthy school*', '*Analysing drug content*', and *'Homelessness in young people*'. All these items are related to the semantic groups *'Youth'* or *'Drugs'*. This ambiguity shows that it is inconclusive in which category to sort these items because the similarity score of the items within the main clusters is very similar to the similarity scores of the ambiguities. This can cause inefficiency and ineffectiveness when users are browsing for the desired information.

Comparative Analysis of Dutch vs. Non-Dutch Participant Heatmap

To be able to answer the research question if mental models differ between users of different nationalities, two heatmaps were created, analysed and compared. This is relevant since the Netherlands is a country of cultural diversity and residents need to be able to find the desired information no matter what nationality they have. Figure 2 shows the heatmap representing the mental model of non-Dutch participants. Figure 3 shows the heatmap representing the mental model of only Dutch participants.

Figure 2

Mental model heatmap of non-Dutch participants (n=28) for the navigation structure of e-government websites.

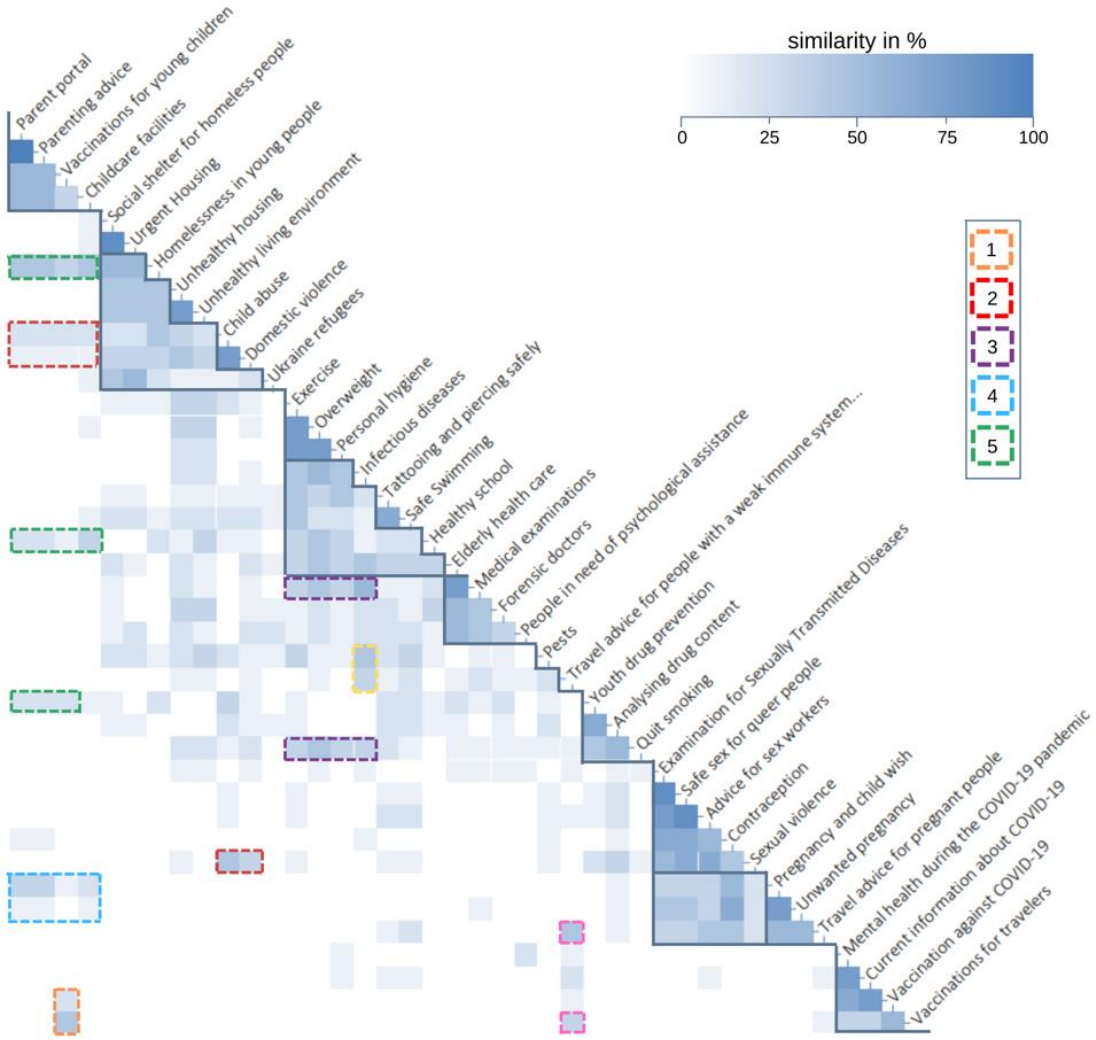


Note. The continuous, dark blue outlines, which can be found along the diagonal, define the clusters of the heatmap, representing the mental model of the average user. The broken outlines define ambiguities and can be found off-diagonal. The different colours describe different ambiguity groups which are explained in more detail in the text. The colours correspond with the colours in Figure 1. The stronger the blue colour of the squares, the stronger the similarity. A stronger similarity means that the two items were grouped more frequently.

After taking a closer look at the clusters found in the heatmap of non-Dutch participants it becomes clear that the clusters are nearly identical to the clusters of the general heatmap of all participants. The main difference that was found is that '*Pests'* was not sorted into the same cluster as it was in the general heatmap. In the general heatmap, the item was sorted into the '*Healthy Lifestyle'* cluster. In the heatmap of non-Dutch participants, it was sorted into the group '*Housing Issues*'. There are no dissimilarities regarding the ambiguities of the general and the non-Dutch heatmap. Merely the shape of some ambiguities is different.

Figure 3

Mental model heatmap of only Dutch participants (n=9) for the navigation structure of e-government service websites.



Note. The continuous, dark blue outlines, which can be found along the diagonal, define the clusters of the heatmap, representing the mental model of the average user. The broken outlines define ambiguities and can be found off-diagonal. The different colours describe different ambiguity groups which are explained in more detail in the text. The colours **do not** correspond with the colours in Figure 1. The stronger the blue colour of the squares, the stronger the similarity. A stronger similarity means that two items were grouped together more frequently.

The heatmap that was created based on all card sorting results from only Dutch participants shows noteworthy differences in comparison to the general heatmap. First, items from the category '*Abuse and Violence*' that was created in the general heatmap were divided into several categories in the heatmap about only Dutch participants. The items '*Domestic violence*' and '*Child abuse'* are now found in the category '*Housing Issues'* and the item '*Sexual* violence' was grouped with other items from the cluster '*Safe Sex and Pregnancy*'. The same goes for the cluster '*Travel Advice*' that was created in the general heatmap. In the heatmap demonstrating the results from exclusively Dutch participants, these items were spread over different categories. '*Travel advice for pregnant people'* is now found in the category '*Reproductive Health*', or more specifically in a sub-cluster of the semantic category '*Pregnancy*'. '*Vaccination for travellers*' is now grouped with COVID-19-related items, and '*Travel advice for people with a weak immune system*' is now part of a single-item cluster, indicating that there was not one item it was frequently grouped with. Thirdly, a new cluster was created with items that belong to the semantic group '*Substance abuse*'. The included items are '*Youth drug prevention*', '*Analysing drug content*', and '*Quit smoking*'. This is especially intriguing due to the Dutch open mind towards drugs and would indicate that there are significant cultural differences between the mental model of the users. These items were spread over several different groups when the general heatmap was analysed.

At a first glance, the ambiguities of the heatmap based on only Dutch participants seem to differ strongly from the other heatmap. When taking a closer look, it becomes clear that these ambiguities only look different because the categories are structured differently than they are in the general heatmap and the heatmap based on non-Dutch participants, as explained in the previous paragraph. Some clusters from the general heatmap were divided so that items were spread over several different clusters in the heatmap based on just Dutch participants. For example, in the general heatmap, the cluster-themed '*Travel Advice*' was created. This cluster does not exist in the heatmap representing results from only Dutch participants since all items of this semantic group are now found in different categories. Since the items are still related to the items of the previous cluster, ambiguities emerge on the heatmap, indicating that one item was grouped frequently with other items as well.

Selection and Data Analysis of GGD Websites

The second part of this research paper is used to find out if the elicited mental model can be applied to e-government websites. This is done to determine if the navigation structures of GGD websites are designed according to the average mental model of the user. This helps them to find the desired information faster. By applying the mental model to e-government websites, it

is possible to make design recommendations to enable users to navigate through websites more efficiently. Since the elicited mental models are based on Dutch e-government websites, namely GGDs, these are also chosen to see if the mental model is applicable.

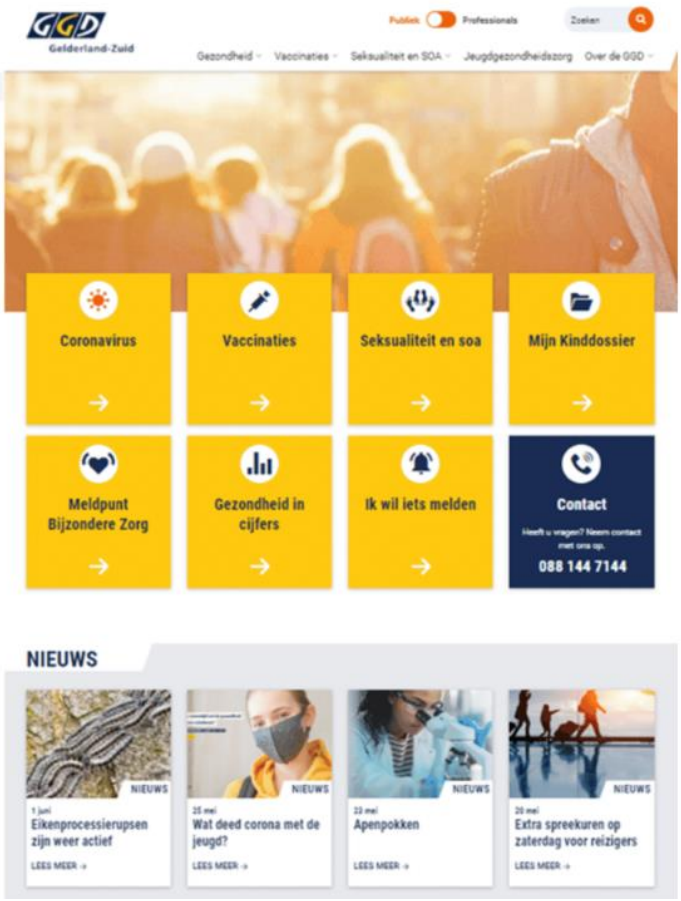
The aim was to randomly select two GGD websites, to determine the applicability of the mental model. One website represents a website the cards were based on (see Appendix A) to verify that the mental model is valid. The other one represents websites the cards were not based on, to find out if the mental model can be applied to other websites and therefore, universal design recommendations for all GGD websites can be made. The navigation structure of some websites is easier to determine than of other websites. Websites were chosen based on the clarity of the structure that could be seen at a first glance. While on some websites the structure seemed rather random, other websites had clear categories. That made it easier to determine the navigation structure. An example of varying structures is the GGD website of Amsterdam in comparison to the GGD website of Gelderland-Zuid (see Figure 4). As can be seen in the figure, the website of GGD Amsterdam has much information listed on the front page which makes it seem overwhelming at a first glance. Furthermore, the categories seem to be very specific and almost seem like single items were listed, for example ‘Tuberculosis’ and ‘Lead Waterpipes’. On the contrary, the website of GGD Gelderland-Zuid only has broader categories listed on the front page which are organised in a structured way. Thus, websites with a structured appearance at the first glance were chosen to determine the navigation structure more easily.

Figure 4

Comparison of varying home page structures of two GGD websites.

a) GGD Amsterdam

b) GGD Gelderland-Zuid



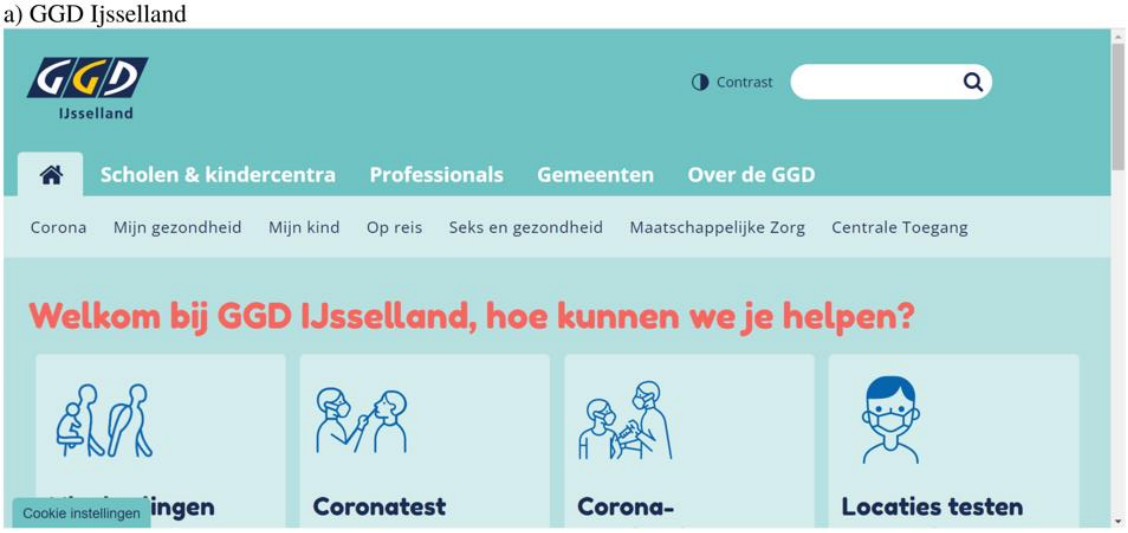
Note. While, at a first glance, GGD Amsterdam seems to have a more overwhelming home page setup, GGD Gelderland-Zuid is an example of a more structured website which makes it easier to determine the navigation structure.

To select two GGD websites, a list of all 25 GGD websites was created based on the mentioned criteria (i.e., clear-structure and navigation bar), divided into websites the cards are based on and websites the cards are not based on (see Appendix E). In the next step, each website was skimmed and marked if the structure seems clear at the first glance. A clear structure was for example if the website has a navigation bar since the research is based on the structure of the navigation bar. For the final decision, two random, clear-structured websites out of the websites marked as suitable for analysis were chosen. The final websites that were chosen are GGD Ijsselland representing the websites that were used to create the cards, and GGD Friesland representing the other websites (Figure 5).

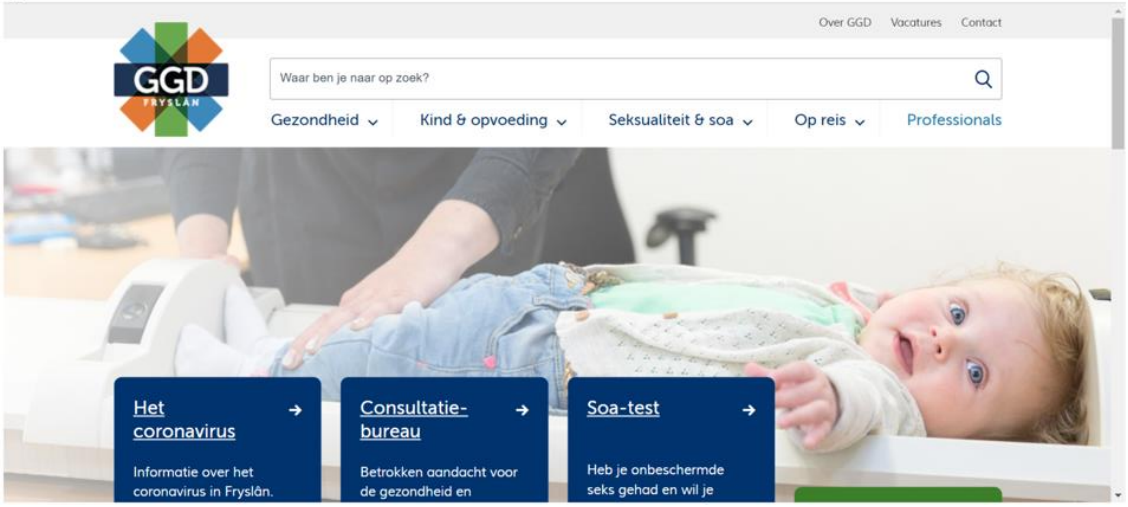
Figure 5

Final choice of GGD websites.

a) GGD Ijsselland



b) GGD Friesland



Note. a) GGD Ijsselland represents a website that was used to create the cards. b) GGD Friesland represents a website that was **not** used to create the cards. Both websites have a navigation bar that is clearly structured which makes determining the navigation structure easier.

The navigation structure of the chosen websites was determined before using SynCaps to determine if the mental model can be applied to those websites. There are different ways to access information on a website. Based on the two examples shown in Figure 4 it can be concluded that websites have a navigation bar at the top of the page, but there are also categories found below the navigation bar. Former research shows that the majority of users do not scroll past the page fold (Nielsen, 2010a). The page fold describes the first part of a website that a user sees without scrolling down (Nielsen, 2010a). For this reason, the main focus when determining the navigation structure of the websites will be on the navigation bar at the top of the page which is visible without scrolling. Not all items of the existing navigation structure could be

implemented into the heatmap since they were not found on each website. While seven items were not included in the heatmap of GGD Ijsselland, eleven items were left out for GGD Friesland. This led to gaps in the navigation structure in the heatmap. This will not have any effect on the data since enough items could be found to create a representative sample.

Using Syncaps, the function 'expert/references sort' was applied to the heatmap making the first participant of the dataset the reference for all other participants. The clusters and similarity scores within the heatmap are the same as before. The only difference is that the centres of the squares are now coloured white or black. While the white squares are not of relevance in this analysis, black squares show the navigation structure of the websites which was set to be the reference sort. This way it is possible to compare the mental model to the navigation structure: If the black squares are located within the clusters, the navigation structure is congruent with the mental model of the user. If the squares are located outside of the clusters, the navigation structure differs from the mental model, which can make navigating through the website more challenging since information is not found there where it is expected. Redesign recommendations can be made if the navigation structure differs from the mental model.

Applicability of the Mental Model to an Existing Navigation Structure

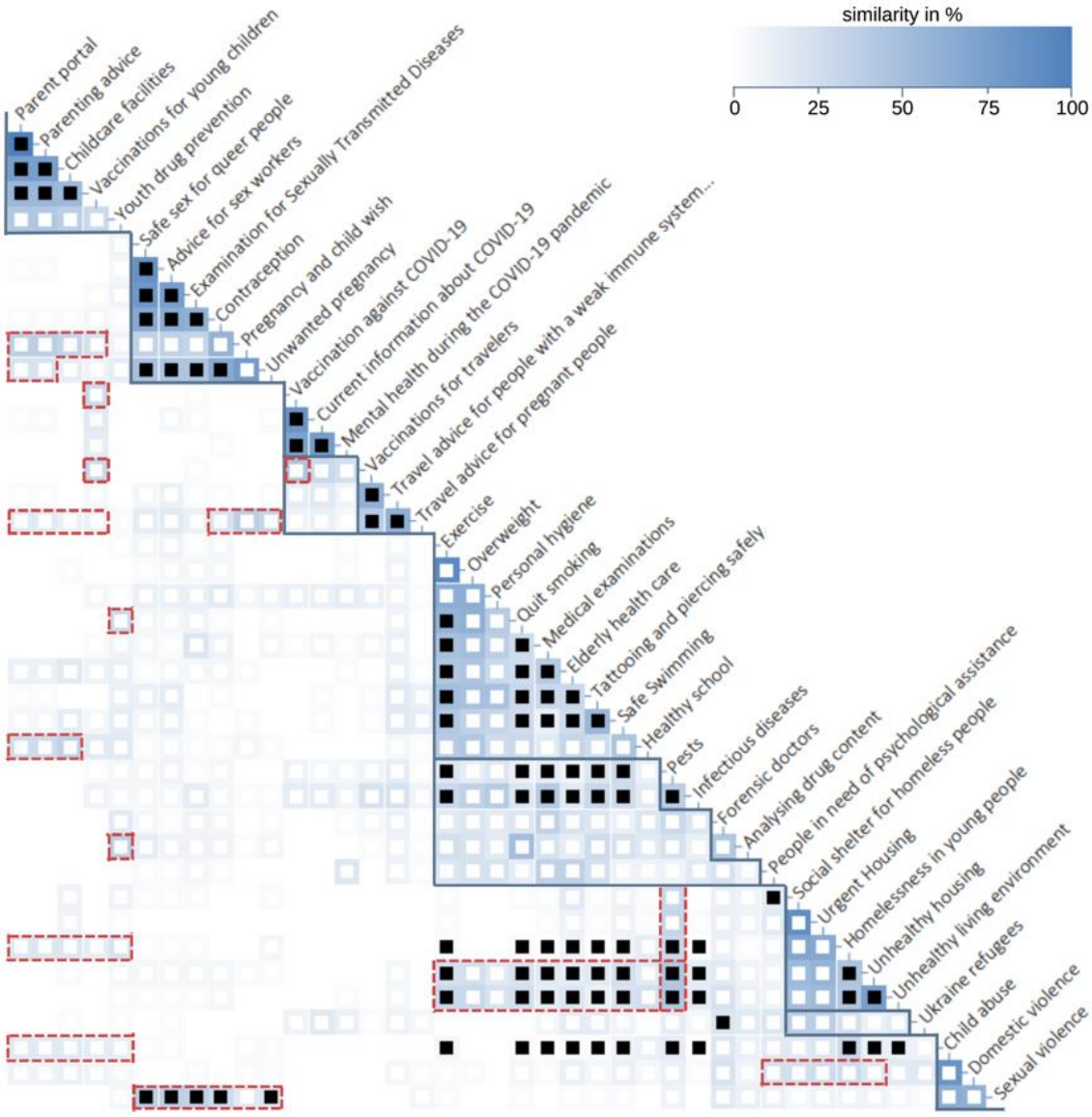
Heatmaps were created for both websites (i.e., GGD Ijsselland and GGD Friesland) to test the applicability of the mental model to the existing navigation structure of Dutch e-government websites. The existing navigation structure is noted in the heatmap with black squares. If two items are connected by a black square, it means that these items were also grouped on the website that is being analysed. There are gaps within the clusters of the existing navigation structure. These gaps are usually caused because an item was not found on the website and was therefore left out. Since the mental model is based on the clusters and the ambiguities, both were marked in the heatmaps.

GGD Ijsselland

First, the navigation structure of the GGD Ijsselland website was implemented (see Figure 6) into the mental model heatmap of all participants to see if the mental model applies to the websites the items were based on. At a first glance, it can be seen that most black squares are located within the predefined clusters and ambiguities. A total of seven items were missing on the website and were therefore left out in the heatmap. While many categories of the existing navigation structure were identical to the mental model of the users, some parts of the navigation structure differ from these categories.

Figure 6

Mental model heatmap of all participants (n = 37) for the navigation structure of e-government websites in comparison to the existing navigation structure of the GGD Ijsselland website.



Note. The continuous, dark blue outlines, which can be found along the diagonal, define the clusters of the heatmap. The stronger the blue colour of the squares, the stronger the similarity. A stronger similarity means that the two items were grouped more frequently. The red, broken outlines define ambiguities and can be found off-diagonal. The black centres within the squares define the existing navigation structure of the GGD Ijsselland website.

Predefined clusters that are almost identical to the mental model are '*Youth Health*', '*Safe Sex and Pregnancy*', '*COVID-19'*, and '*Travel Advice'*. The only difference within these clusters is that some items could not be found on the GGD Ijsselland website (i.e., '*Youth drug prevention*' and '*Pregnancy and child wish*'). The category '*Abuse and Violence*' was not created in the existing navigation structure. While the item '*Domestic violence*' could not be found anywhere on the website, the other two items (i.e., '*Child abuse*' and '*Sexual violence*') were located

outside of the main cluster. The item '*Sexual violence*' was sorted into the category '*Safe Sex and Pregnancy*' which was also identified as an ambiguity in the general heatmap. Thus, this item was grouped congruent with the mental model. The item '*Child abuse'* is the only item within the previously mentioned clusters that do not comply with the mental model and was grouped with items from the cluster '*Healthy Lifestyle*' and '*Housing Issues'*. A possible explanation is that the mentioned item was not found directly on the website. Instead, '*female circumcision*' was found which is a form of child abuse. Since female circumcision was found under the category '*Health*' on the GGD Ijsselland website, the item '*Child abuse'* was also sorted into the same category when creating the code.

Based on what was mentioned in the previous paragraph it can be said that the mental model of e-government websites is mostly congruent to the existing navigation structure. When looking at the clusters '*Healthy Lifestyle*' and '*Housing Issues*' it becomes obvious that many items are located outside of the main categories and form a large cluster of black squares. The named cluster includes the items '*Homelessness in young people*', '*Unhealthy housing*', and '*Unhealthy living environment*', which were previously sorted into the category '*Housing Issues*', and '*Child abuse*', previously sorted into the category '*Abuse and Violence*'. All these items were found in the category '*Health*' on the website of the GGD Ijsselland. The items '*Unhealthy housing'* and '*Unhealthy living environment*' are located within a previously defined ambiguity. On the contrary, the other two items are located outside of the ambiguity and have a low similarity score in the general heatmap, indicating that these items would have not been grouped by the average user. Overall, it seems as if the main clusters '*Healthy Lifestyle*' and '*Housing Issues*' from the mental model are subclusters of the main category '*Health*' from the GGD Ijsselland website. Hence, this once more implies congruency.

Although most of the relations between navigation structure and mental model have been described, some items that were paired on the website still have to be clarified. First, the items '*People in need of psychological assistance*' and '*Social shelter for homeless people*' were found in the same group on the website, namely '*Social Care*'. In the general heatmap, these items have a very low similarity score of 14% showing that users would have not expected to find both items in the same category. In the general heatmap, the item '*People in need of psychological assistance*' was not grouped frequently with any other item and was therefore considered a single item cluster. The item '*Social shelter for homeless people*' has a high similarity score of 89% with '*Urgent housing*'. Accordingly, a grouping of the items '*People in need of psychological assistance*' and '*Social shelter for homeless people*' was unexpected and probably leads to inefficient navigating through websites. Last but not least, the items '*Ukraine refugees*' and

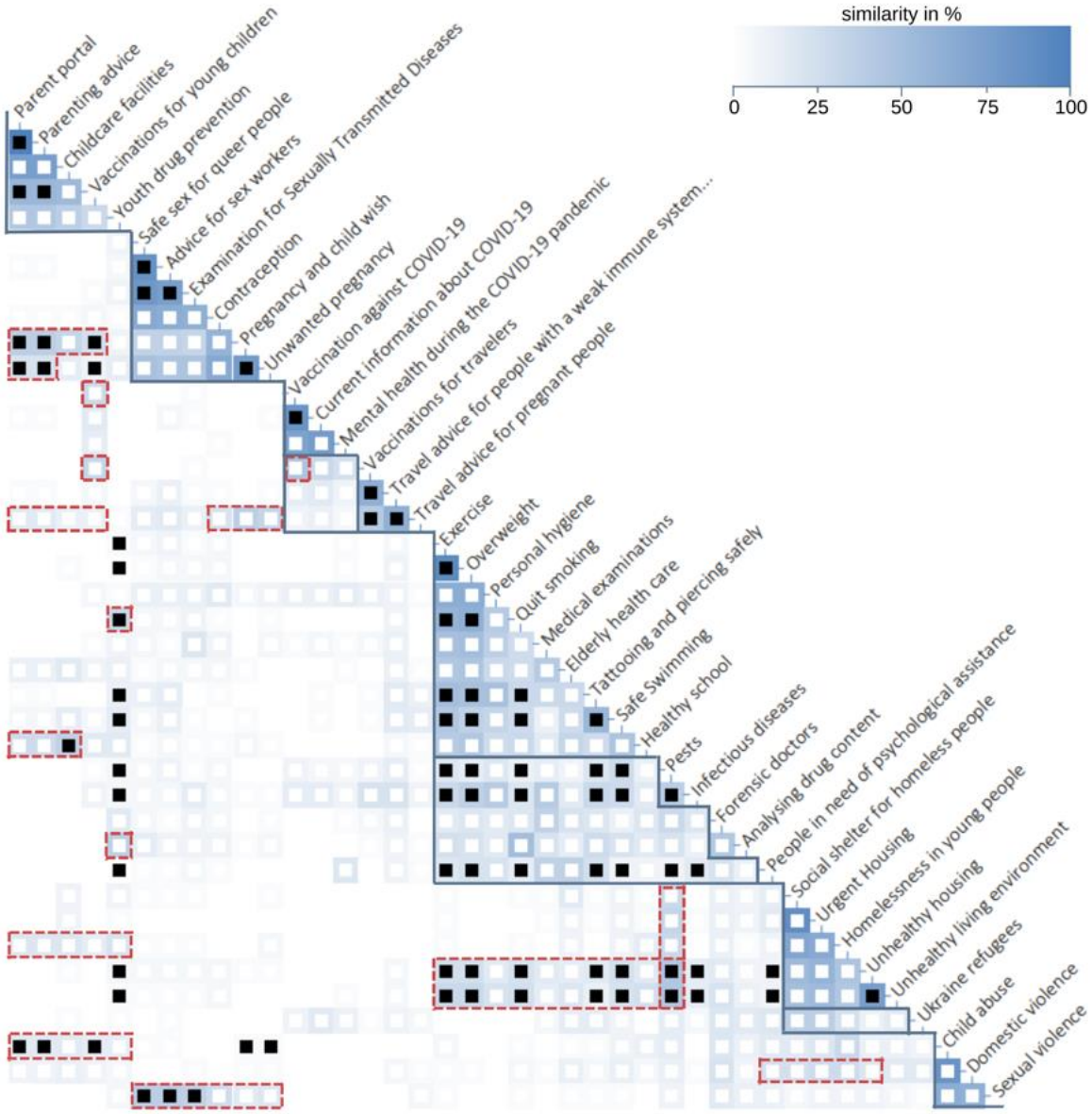
'Forensic doctors' were put into one category when creating the SynCaps code which can be confusing when looking at the heatmap. The reason behind this grouping is that both items were in located a section called '*Professionals*' on the GGD website, which is not meant to provide information for residents. Since the recruitment of participants was not aimed at professionals in this paper, the group in the SynCaps file was also named '*Professionals*' and is not of further importance for this research paper. In conclusion, it can be said that the mental model was successfully applied to the navigation structure of the GGD Ijsselland.

GGD Friesland

To determine if the mental model can also be applied to websites that were not used to create the cards, the navigation structure of GGD Friesland was implemented into the heatmap (see Figure 7). Noticeably more black squares are not located outside of the main clusters, but they are mostly still found within ambiguities. This indicates that at a first glance, the mental model seems to also apply to other e-government websites. For this website, eleven items in total could not be located on the website. Similarly, as it was sorted in the heatmap for the GGD Ijsselland, a large cluster of black squares becomes apparent that connects the main clusters '*Healthy Lifestyle*' and '*Housing Issues*'. This cluster includes the items '*Unhealthy housing'* and '*Unhealthy living environment*'. As with the GGD Ijsselland website, these two categories are also subcategories of an overarching category labelled '*Health*'. Similar to the GGD Ijsselland website, the category '*Abuse and Violence*' was also not created on the GGD Friesland website. Also here, the item '*Domestic violence*' could not be located on the website. The other two items from this cluster from the general heatmap (i.e., '*Sexual violence*' and '*Child abuse*') could be found within ambiguities and are therefore conform with the mental model. Finally, the item '*Ukraine refugees*' will be evaluated. Since it was not grouped with any items, at a first glance it seems like the item could not be located at all on the website. This is misleading because there are simply no black squares related to this item since it is its own category on the GGD Friesland website. A possible explanation for this is that refugees from Ukraine are a very recent topic that currently has a lot of attention due to the Russia-Ukraine war. For this reason, this item most likely has its own category on the homepage.

Figure 7

Mental model heatmap of all participants (n = 37) for the navigation structure of e-government websites in comparison to the existing navigation structure of the GGD Friesland website.



Note. The continuous, dark blue outlines, which can be found along the diagonal, define the clusters of the heatmap. The stronger the blue colour of the squares, the stronger the similarity. A stronger similarity means that the two items were grouped more frequently. The red, broken outlines define ambiguities and can be found off-diagonal. The black centres within the squares define the existing navigation structure of the GGD Friesland website.

Although this website overall seems to have more outliers, the conclusion can be drawn that the mental model is also applicable to the GGD website of the municipality of Friesland. Many outliers of this heatmap are related to one single item (i.e., '*Youth drug prevention'*) which does not make the mental model a bad fit for this website. All in all, the mental model seems to

apply to all Dutch e-government websites, namely the GGD, and is, therefore, an accurate representation of the mental model of the average user.

General Discussion

Design Recommendations

The purpose of this research paper was to elicit the mental model of e-government websites. It became clear that the mental model and the existing navigation structure do not differ much. Still, some minor changes could positively influence the efficiency of users while browsing the internet. In the next part, the applicability of the mental model was tested to validate the elicited mental model and make design recommendations for e-government websites. Instead of making specific redesign recommendations for one website, a general design recommendation for all GGD websites will be given. It became evident that there are clear differences between the structure and design of the websites, which would make it tedious work to give redesign recommendations for every single one of them.

The rationale for giving design recommendations is that GGD websites can be designed congruent to the mental model of the average user so that they can navigate through e-government websites more efficiently. A uniform design across all GGD websites would cause that users can adjust to the new website fast (Lynch et al., 1999) if they have to look up information on different websites. This can for example happen when someone moves from one GGD region to another and different regulations apply. Hannum (2001) adds to the statement by Lynch et al. (1999) that not having a consistent design can lead to confusion and feelings of frustration.

The product of the previously conducted research is a navigation structure enabling users to easily find the desired information on an e-government website. Table 3 shows the proposed navigation structure which includes seven categories listed in a sorted manner illustrating the categories of the navigation bar on the home page of the website. Categories 1 and 2 are placed at the beginning of the navigation bar because '*Russia-Ukraine War*' and '*COVID-19*' are both topical information which is currently of great relevance and will therefore be looked up more often. Category 3, namely '*Abuse and Violence*', was sorted next since it closely linked to topics that require immediate attention. Category 4, '*Healthy Lifestyle*', could be subdivided into two subcategories (i.e., '*Housing Issues*' and *'Mental Health*'). This category was placed right after the current topics since it contains much information about many different fields. An example of the variety of themes in this category is '*Infectious diseases*', '*Elderly care*', and '*Quit smoking*'. '*Youth Health*' is the label of Category 5. This category was sorted after the Category '*Healthy Lifestyle*' because it is semantically closest out of all other categories. Category 6, namely '*Safe*

Sex and Pregnancy' is placed after '*Youth Health*' due to semantic reasoning as well. Since '*Travel Advice*' (i.e., Category 7) is related to leisure instead of essential medical advice, it was placed at the end of the list. All in all, it became obvious that the categories are sorted based on relevance and semantics.

Table 3

Proposed navigation structure for municipal health service websites.

	Label
Category 1	Ukrainian Refugees
Category 2	COVID-19
Category 3	Abuse and Violence
Category 4	Healthy Lifestyle
Subcategory 4.1	Housing Issues
Subcategory 4.2	Mental Health
Category 5	Youth Health
Category 6	Safe Sex and Pregnancy
Category 7	Travel Advice

In the research at hand, mental models of users from different countries were compared. This is relevant because the focus of this paper is on Dutch e-government websites. Additionally, since crucial information regarding current topics, for instance, COVID-19 and Ukraine refugees, can be found on websites from this domain and all residents in the Netherlands, no matter their cultural background, need to access this information, designing the website to be efficient to use for everyone is essential. The Netherlands is known to be a country of cultural diversity (Dam et al., 2003). Therefore, Dutch participants were compared to non-Dutch participants. Kralisch et al. (2005) claim in their paper that cultural backgrounds influence navigation behaviour. The findings of this paper are in line with these results. Intriguing differences were found between both participant groups. The clusters '*Abuse and Violence*' and '*Travel Advice*' were not created in the heatmap of only Dutch participants. Instead, the items from these clusters were grouped with other semantically related items. Furthermore, a new category has been created, namely '*Substance Abuse*', which is interesting due to the Netherlands being known for its open mind towards drugs.

Limitations

There are several limitations regarding the research paper at hand that have to be discussed. First, one could argue that the number of cards could create a limitation. With the tremendous amounts of information found on e-government websites and the diversity of

information throughout the different GGD websites, it is a challenging task to create cards that are representative of all information that is listed on a website. Although the content of the cards was carefully chosen, not all information could be included in the set of cards. Besides, the information found on GGD websites varies across each website. Thus, it is impossible to create a set of cards that is representative of all GGD websites. In addition, one had to be careful to not choose too many cards. Miller (2011) argues that having a sample that is too small could lead to missing content items that could be sorted into an additional category. It can be expected that a card sorting experiment with too many cards can lead to the participants losing motivation and not seriously completing the task.

Another limitation could be the labels of the card. Previous research has shown that attention should be paid to the wording of the items (Faiks et al., 2000): Labels of cards should vary and not contain the same words. If the labels of cards are too similar, participants might group these without thinking about the actual meaning of the labels. An example is the items '*Vaccination against COVID-19*', '*Current information about COVID-19*', and '*Mental health during the COVID-19 pandemic*'. In addition, Faiks et al. (2000) recommend instructing users to keep the concepts behind the words on the cards in mind rather than just considering the words themselves. The users in the study at hand were not given any instructions of that kind. To avoid grouping based on merely the words, it is advised to not use the same words for different labels but, for example, use a synonym or a paraphrase.

Last but not least, the representativeness of the participant sample is a limitation of this paper. This especially applies to the comparison of Dutch and non-Dutch participants. While 28 participants were not Dutch, only 9 out of 37 were Dutch. In their study, Sherwin (2018) clarifies that at least 15 participants should participate in a card sorting study to have a representative sample. Thus, there was not a representative Dutch sample which could lead to a different mental model than if there would have been more Dutch participants.

Future Research

Based on the findings of the study at hand, suggestions for further research can be made. Further research should explore how similar GGD websites are regarding their navigation structure. In the current research paper, the assumptions were made that the navigation structures across GGD websites vary, but it was not explicitly mentioned in what way they vary. The assumptions are only based on first-glance observations and are not based on scientific research. The similarity of the websites can be researched by, for example, comparing the site maps of the websites which represent the navigation structure.

Regarding the cultural differences, it is intriguing to find out if the usability is increased if varying navigation structures are created for users of different cultural backgrounds. In the results of several different research papers, it is mentioned that the cultural background influences the behaviour and experienced user experience on a website (Dam et al., 2003), as well as the navigation behaviour (Kralisch et al., 2005). Further research hypothesized that designing the navigation structure in line with the mental model of users with demographic diversity could enhance navigation behaviour (Alexander et al., 2017; Nawaz, 2013). Future research could go a step further and try to confirm that this hypothesis is true by, for example, having a sample of two cultures from whom the mental model is elicited and then conducting usability testing using two example navigation structures which are designed congruent to each of the mental models.

Moreover, future studies should aim at validating the card sorting method. Although the results by Schmettow and Sommer (2016) indicated that card sorting does not bring any benefits when designing a navigation structure, the study at hand is completely based on card sorting results. Other studies did find a positive effect of card sorting interfaces on user interfaces, so the future study should determine if these results can be confirmed and explicitly apply card sorting to elicit mental models of websites.

Conclusion

In this research paper, the mental model of the average user on e-government websites was elicited. This adds to existing knowledge since mental models of municipal health service websites have not been elicited before. Based on this insight it was possible to make recommendations regarding adjustment of the navigation structure for this domain. Although the proposed navigation structure is not too different from the existing navigation structure, there are crucial details that should be improved to positively influence the browsing experience of users. It is recommended to use the suggested navigation structure as a framework for GGD websites and add additional items that were not included in the item set. By implementing the recommendations, navigating through GGD websites will be easier and more intuitive.

References

Alexander, R., Murray, D., & Thompson, N. (2017). Cross-cultural web usability model. *International Conference on Web Information Systems Engineering* (pp. 75-89). Springer.

Bernard, M., & Sheshadri, A. J. U. N. (2004). Preliminary examination of global expectations of users' mental models for e-commerce web layouts. *6*(2), 1-9.

Branaghan, R. J., Covas-Smith, C. M., Jackson, K. D., & Eidman, C. J. A. e. (2011). Using knowledge structures to redesign an instructor–operator station. *42*(6), 934-940. https://doi.org/10.1016/j.apergo.2011.03.002

Bussolon, S., Russi, B., & Missier, F. D. (2006). Online card sorting: as good as the paper version. *Proceedings of the 13th Eurpoean conference on Cognitive ergonomics: trust and control in complex socio-technical systems*, 113-114. https://doi.org/10.1145/1274892.1274912

Card Sorting. (2013). Department of Health and Human Services. https://www.usability.gov/how-to-and-tools/methods/card-sorting.html#:~:text=Best%20Practices%20for%20Card%20Sorts,especially%20for%20an%20open%20sort.

Dam, N. v., Evers, V., & Arts, F. A. (2003). Cultural user experience issues in e-government: Designing for a multi-cultural society. I*nternational Digital Cities Workshop* (pp. 310-324). Springer.

Faiks, A., Hyland, N. J. C., & libraries, r. (2000). Gaining user insight: a case study illustrating the card sort technique. *61*(4), 349-357. https://doi.org/10.5860/crl.61.4.349

Hannum, W. J. W.-b. t., New Jersey, Educational Technology Publications. (2001). Web-based training: advantages and limitations. 13-20.

ISO. (1998). *Ergonomic requirements for office work with visual display terminals (vdts)–part 11: Guidance on usability. ISO Standard 9241-11: 1998* (Vol. 55).

Katz, M. A., & Byrne, M. D. J. A. T. o. C.-H. I. (2003). Effects of scent and breadth on use of site-specific search on e-commerce Web sites. *10*(3), 198-220. https://doi.org/https://doi.org/10.1145/937549.937551

Kotamraju, N. P., van der Geest, T. M. J. B., & Technology, I. (2012). The tension between user-centred design and e-government services. *31*(3), 261-273. https://doi.org/10.1080/0144929X.2011.563797

Kralisch, A., Eisend, M., & Berendt, B. (2005). The impact of culture on website navigation behaviour. *Proc. HCI-International* (pp. 1-9).

Lynch, P. J., Horton, S., & Horton, S. (1999). *Web style guide: Basic design principles for creating web sites*. Universities Press.

Lyzara, R., Purwandari, B., Zulfikar, M. F., Santoso, H. B., & Solichah, I. (2019). E-government usability evaluation: Insights from a systematic literature review. *Proceedings of the 2nd International Conference on Software Engineering and Information Management* (pp. 249-253). https://doi-org.vu-nl.idm.oclc.org/10.1007/978-3-319-68786-5_6

Miller, C. S. (2011, May). Item sampling for information architecture. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2211-2214). https://doi-org.vu-nl.idm.oclc.org/10.1145/1978942.1979264

Nawaz, A. (2013). *Website user experience: A cross-cultural study of the relation between users' cognitive style, context of use, and information architecture of local websites*. Frederiksberg: Copenhagen Business School (CBS). http://hdl.handle.net/10419/208876

Nielsen, J. (2010a). *Scrolling and Attention (Jakob Nielsen's Original Research Study)*. Nielsen Norman Group. https://www.nngroup.com/articles/scrolling-and-attention-original-research/?lm=scrolling-and-attention&pt=article

Nielsen, J. (2010b). *Mental Models*. https://www.nngroup.com/articles/mental-models/

Otter, M., & Johnson, H. J. I. w. c. (2000). Lost in hyperspace: metrics and mental models. *13*(1), 1-40. https://doi.org/10.1016/S0953-5438(00)00030-8

Oulasvirta, A. J. I. w. C. (2004). Task demands and memory in web interaction: a levels of processing approach. *16*(2), 217-241. https://doi.org/10.1016/j.intcom.2003.12.004

Petrie, H., Power, C., Cairns, P., & Seneler, C. (2011). Using card sorts for understanding website information architectures: technological, methodological and cultural issues. *IFIP Conference on Human-Computer Interaction*, 309-322. https://doi.org/10.1007/978-3-642-23768-3_26

Righi, C., James, J., Beasley, M., Day, D. L., Fox, J. E., Gieber, J., . . . Ruby, L. (2013). Card sort analysis best practices. *8*(3), 69-89.

Rosenfeld, L., & Morville, P. (2002). *Information architecture for the world wide web*. " O'Reilly Media, Inc.".

Roth, S. P., Schmutz, P., Pauwels, S. L., Bargas-Avila, J. A., & Opwis, K. J. I. w. c. (2010). Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages? , *22*(2), 140-152. https://doi.org/10.1016/j.intcom.2009.10.004

Schall, A. (2014). Information architecture and web navigation. In *Eye tracking in user experience design* (pp. 139-162). Elsevier. https://doi.org/10.1016/B978-0-12-408138-3.00006-6

Schmettow, M., & Sommer, J. (2016). Linking card sorting to browsing performance – are congruent municipal websites more efficient to use? *Behaviour & Information Technology*, *35*(6), 452-470. https://doi.org/10.1080/0144929x.2016.1157207

Sherwin, K. (2018). *Card Sorting: Uncover Users' Mental Models for Better Information Architecture*. Nielsen Norman Group. https://www.nngroup.com/articles/card-sorting-definition/

Stroes, J. D. (2018). *Adjusting the navigation structure of a faculty intranet to the average user with Card Sort and Q-sort.* University of Twente].

Vind een GGD bij u in de regio. (n.d.). GGD. https://www.ggd.nl/

Appendices

Appendix A – Final Selection of GGD Websites

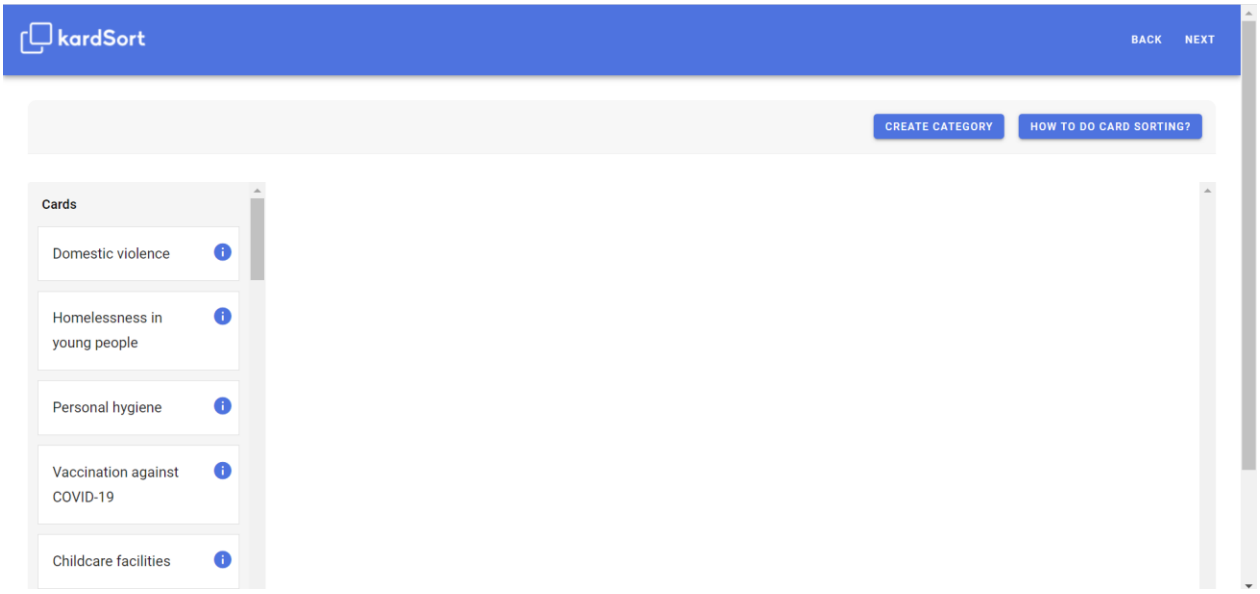
	Municipality	URL
1.	Leiden	https://www.ggdhm.nl/
2.	Rotterdam	https://www.Ggdrotterdamrijnmond.nl/
3.	Eindhoven	https://www.ggdbzo.nl/
4.	Groningen	https://www.ggd.groningen.nl/
5.	The Hague	https://www.ggdhaaglanden.nl/
6.	Amsterdam	https://www.ggd.amsterdam.nl/
7.	Enschede	https://www.ggdtwente.nl/
8.	Zwolle	https://www.ggdijsselland.nl/
9.	Nijmegen	https://www.ggdgelderlandzuid.nl/
10.	Utrecht	https://www.ggdru.nl/

Appendix B – Final Set of Cards

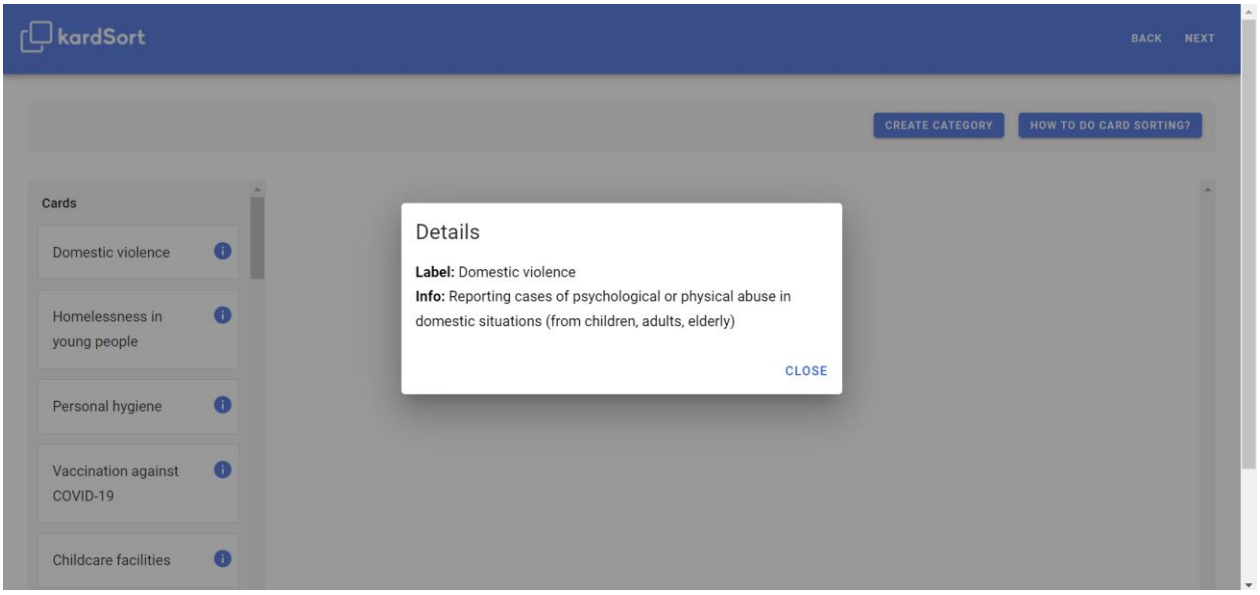
Card No.	Title	Description
1.	Pests	Reporting of infestations of rats, critters, caterpillars, oak processionaries etc. inside and outside the house
2.	Examination for Sexually Transmitted Diseases	Testing and consultation for preventing sexually transmitted diseases
3.	Exercise	Tips and advice on reaching the exercise norm to have a higher chance of ageing well
4.	Overweight	Information about being overweight and organized activities for people that suffer from obesity
5.	Elderly health care	Information about health care for elderly, for example, dental care and exercise
6.	Quit smoking	Information on the dangers of smoking and advice on how to stop smoking
7.	Youth drug prevention	Information for parents and caregivers on how to prevent teenagers from taking drugs
8.	Unhealthy housing	Reporting and information on asbestos, dirty housing, lead in water, increased moisture
9.	Contraception	Information about what contraception methods exist and options for consultation
10.	Pregnancy and child wish	Tips and advice before and during pregnancy
11.	Unwanted pregnancy	Anonymous decision-making advice for women that are pregnant unintentionally
12.	Vaccinations for travellers	Information about which vaccines are necessary when travelling to specific countries
13.	Vaccinations for young children	Information about the vaccination program for babies and children (What vaccines will my child get? When will they get them?...)
14.	Forensic doctors	Information, practicalities and reports related to police reports or investigations (e.g., reporting accidents, suspicious deaths, driving under influence etc...)
15.	Medical examinations	Medical examinations and general health monitoring with doctor appointments
16.	Infectious diseases	Information about a variety of infectious diseases (e.g., general information, insurance for hepatitis B, C, scabies, chicken pox, whooping cough etc...)
17.	Homelessness in young people	Reporting homeless youth and information on how organizations will care for homeless youth
18.	Tattooing and piercing safely	Monitoring of tattoo and piercing studios to make sure that they are safe and hygienic
19.	Sexual violence	Reporting, information, and help with sexual assault and rape
20.	Safe sex for queer people	Preventing the transmission of STDs through vaccination, medicine and consultations for men who have sex with men and trans people
21.	Analysing drug content	Legally getting content and risk of drugs checked before consumption
22.	Advice for sex workers	Offering sex workers to test for sexually transmitted

		diseases and consultation hours for giving information on how to work safely
23.	Parent portal	A website on which parents can access their children's digital health records (e.g., growth curve, development, registered vaccinations)
24.	Mental health during the COVID-19 pandemic	Information and advice about the negative impact of the coronavirus on mental health
25.	Healthy school	Information and advice about how to be a healthy school (e.g., fruit days, education about relationships and sexuality, smoke-free school grounds)
26.	Personal hygiene	For example, handwashing, cleaning, and food hygiene
27.	Childcare facilities	Childcare facilities that meet requirements of a healthy and safe environment for daycare
28.	Child abuse	Information about child abuse and possibility to report suspicions of child abuse
29.	Ukraine refugees	Information about how to help refugees from the Russia-Ukraine war
30.	Vaccination against COVID-19	Information about vaccination locations and frequent questions about the corona vaccine
31.	Travel advice for pregnant people	Information on how to prepare for travelling while being pregnant
32.	Travel advice for people with a weak immune system or an illness	Information about risks they can encounter, and advice related to travelling to certain countries for people with a weaker immune system
33.	Current information about COVID-19	Latest news about the COVID-19 pandemic, for example, corona-updates
34.	People in need of psychological assistance	Reporting or seeking help for people who are exhibiting worrying behaviour and might need psychological assistance
35.	Social shelter for homeless people	Access and guidance to various forms of social shelter and sheltered housing for homeless people
36.	Safe Swimming	Information and advice on diseases from swimming waters and clean hygienic swimming
37.	Urgent Housing	Information and advice on getting emergency shelter for people in need
38.	Domestic violence	Reporting cases of psychological or physical abuse in domestic situations (from children, adults, elderly)
39.	Unhealthy living environment	Information about bad air quality, soil pollution, noise pollution
40.	Parenting advice	Advice and information on raising children based on different ages

Appendix C – Main Experiment Screen in KardSort



Description of Cards



Appendix D – Pre- and Post-Experiment Screen in KardSort

Informed Consent

Please read the following information thoroughly:

You are being invited to participate in a research study titled Website Navigation Structures: Eliciting Mental Models Using Card Sorting. This study is being done by Jona Sabrina Kunert and Erin Mino as part of the B.Sc. Psychology Bachelor's thesis at the University of Twente.

The purpose of this research study is to investigate the mental model of users when visiting a municipal health care service website and will take you approximately 20 minutes to complete. The data will be analyzed and used for making design recommendations and finding out which websites are less or more congruent to the mental models of users.

Your participation in this study is entirely voluntary and you can withdraw at any time. You are free to omit any question.

We believe there are no known risks associated with this research study; however, as with any online-related activity, the risk of a breach of information is always possible. To the best of our ability, your answers in this study will remain confidential. We will minimize any risks by safely storing and confidentially handling the anonymized data.

On the next page, you will be asked if you read the above information and agree to continue and participate in this study. Please continue to the next page now by clicking on "NEXT" in the top-right" corner.

Instruction Page

Welcome to our study!

We are currently researching how information on municipal health service websites (e.g., GGD) can be best organized. This will help users of these websites to find the information they are looking for more easily.

To gain insight into how to best structure information on municipal health service websites, the method "card sorting" will be applied. On the next page, you will be presented with 40 cards. Each card contains an item that can be found on municipal health service websites, along with a description of it. Please sort the cards into groups as you would expect to find them on a municipal health service website. Keep in mind that there are no right or wrong answers! Please provide us with the grouping that you feel makes the most sense.

You can create a new group by clicking on the "CREATE CATEGORY" button in the top-right corner. After you have created a group, you can name it. How would you describe the cards in each of the categories? You can use a single word, phrase, or sentence.

Please try to put all cards into meaningful groups. If you feel there is a card that does not belong anywhere at all, you can create an extra group called "others" and sort this card into that group. Otherwise, please put the cards into the second most fitting group. If you feel that a particular card belongs in more than one location, please place the card in the best location you believe it fits in.

Have fun!

Acknowledgement Page

Thank you for participating in our study!

If you have any questions or comments, please send a mail to j.s.kunert@student.utwente.nl or e.mino@student.utwente.nl.

Appendix E – Overview of all GGD websites

Websites cards are based on	Websites cards are not based on
Amsterdam	Drenthe
Brabant Zuidoost	Flevoland
Gelderland-Zuid	Friesland
Groningen	Gelderland-Midden
Haagenlanden	Gooi en Vechtstreek
Hollands Midden	Hart voor Brabant
Ijsselland	Hollands Noorde
Rotterdam-Rijnmond	Kennemerland
Twente	Levoland
Utrecht	Limburg Noord
	Noord- en Oost Gelderland
	West-Brabant
	Zaanstreek-Waterland
	Zeeland
	Zuid Limburg