Abstract

The Danish government has digitalized more and more services in the last few years and thus forced citizens to use online resources such as e-boks.dk and borger.dk. Some older citizens have issues with using this technology and are therefore in a situation where they cannot use their civil rights. This thesis aims to investigate whether or not there is a difference between how the older citizens view and navigate these types of websites, compared to the young generation for whom this technology seems like second nature.

The thesis has inspiration in a theory by Prensky who claim that the younger generation thinks differently than the older one, as they have been exposed to technology in a way that the older have not. The younger generations have been born into a digitalized world and are thus digital natives, while their parents and earlier generations have had to adopt it and are digital immigrants. The questions are whether or not Prensky is right, a question this thesis tries to answer.

Previous research has found that the ability of the older generations is not based on age as a single factor, but on the user's prior knowledge of specific domains and the technology involved. Even though differences exist between different age groups it is found that it is not because of age itself.

The method used in this study for investigating this topic is eye tracking, a technique that allows for the tracking of gaze movement on an electronic screen. The data collected can answer the question of whether or not people of the old and young generation are different. This was supplemented by data collected through a questionnaire about participants' age and prior knowledge as research has found it to be the key difference between these age groups.

The data collected on prior knowledge and age showed that the median for the elderly age group was 2.75 on a scale from one to six, while the median for the young age group were 5.

Several measures were analyzed in regards to eye tracking, the study investigated click paths, time to complete, areas of interest and scan paths. These measures were analyzed in groups determined by demographics and prior knowledge to see if these factors made any difference in and between the groups.

The study concludes, with 31 participants, that age and prior knowledge does not have an effect on peoples' ability to use websites. Age does have an effect on how fast

users are, but the groups proved very similar in all other categories. The amount of errors by each group based on age and prior knowledge was similar. The scan paths taken was very similar as well. The study could not find any differences beside the time to complete and difference in prior knowledge, which led to the conclusion that the notion of digital natives and digital immigrants was unfounded in this thesis' context and that low prior knowledge did not seem to hamper users.

1 | Introduction

By November 2014, all public mail in Denmark was to be sent as digital mail in an effort, by the state, to digitalize it as is already done with other services such as doctors' appointments and tax registrations ("Digital Post" 2016). Even though Denmark is a first world country in a western society, there are still 10.6% of Danish citizens who are exempt from digital mail for several reasons; old age, mental or physical disability etc. ("Månedlig Statistik Om Digital Post" 2016). Those 10.6% are predominantly made up of people of 55 years old and up; 86.9% to be precise ("Månedlig Statistik Om Digital Post" 2016). This led us to wonder exactly why this age group seem to have more difficulties operating in a digital space and if they are different to other age groups entirely.

1.1 | Premise

We were originally inspired to write this thesis by observations in everyday life, seeing how different people of an older age group would react to modern technology in very different ways. Some would fear it and others would embrace it. It seemed that age did not affect this, as some of the fearful were younger than those who embraced it. We started to discuss the concept of digitalization and what issues might arise from it. Based on our observations we had the impression that some people could not use the websites they were required to. Our thought process was that these sites were maybe created for a different audience that were much more experienced with using websites; people in our own age group who grew up with the Internet. This sparked the idea that maybe there was a difference between young and old as website users. In particular, we were interested in finding out if these age groups would search for information differently due to our backgrounds in Information Architecture.

Marc Prensky, an American psychologist have coined the terms "digital native" and "digital immigrant", a reference to the difference between users who have grown up in a digital world versus the users who have adopted these technologies as adults. The premise for this way of thought is that young people, who have grown up in a digital world, are different to older people who learned to use computers when they were adults. Prensky's work has served as inspiration for this thesis as we are very interested in this supposed divide in ability between young and old.

As we digitalize more and more of the tasks and processes, which were previously done face-to-face, we are also giving more responsibility to the individual. In Denmark each citizen is required to check their digital mail with letters from public services and similar since they do not send physical letters anymore. Taxes are registered online, applications for different welfare benefits and more are done digitally, as we are digitalizing our society.

Digitalization is a part of a concept called E-Government which refers to public administration processes being conducted using information and communication technology. The European Union defines e-government as

"… the use of information and communication technologies in public administrations - combined with organisational change and new skills - to improve public services and democratic processes and to strengthen support to public policies." ("Electronic Government — eGov Conference" 2016)

Despite digitalizing administrative processes can have positive effects, the potential problems with e-government are numerous; users need to trust the sites available for them to be successful (Business Wire 2010), users need to feel like they get an advantage from using the sites (Billestrup and Stage 2014), poor design can cause user dissatisfaction (Meuter et al. 2000). Furthermore, Meuter et al. argues that designing a system like a self-service system only works if users were accounted for during the design process. Stage and Billestrup investigated and concluded that the software providers that developed a self-service solution for the Danish government did not involve any citizens in the development process(Billestrup and Stage 2014). The lack of user experience research for this e-government project is very problematic as citizens are forced to use it and simply cannot go anywhere else for their information.

In order for us to investigate the potential issues with elderly and their ability to use websites and the notion of a younger age group being more capable, we have to conduct tests. In these tests we will track users eye movement to find out where they look and in what order they look at different parts of the website. Doing an eye tracking study is more appropriate than doing usability tests as we are not interested in testing the site but the user themselves. Furthermore, conducting an eye tracking study also allows us to have information on how our participants tries to solve a task rather than finding out whether or not they can do it at all.

The data we collect from both groups can be contrasted to previously created data on those same age groups and ultimately aid in a comparison of the two.

1.2 | Problem statement

We would like to investigate if young and elderly users look at websites differently, when searching for information. To do this we conduct tests, to find any patterns that might indicate differences. The test participants will be selected by looking at commonly used groupings by age, already used today. The results will be used to either confirm or disprove the notion of 'digital immigrants' and 'digital natives' in this context. These terms are our adaptions of the terms coined by Mark Prensky.

What characterizes and differentiates the elderly and young people in their approach and ability to navigate and observe a website's information architecture?

- How is it possible to conduct tests on two age groups in regards to their ability to navigate information architecture?
- To what degree is the notion of natives vs. immigrants true in this context?
- What characterizes the two groups and to what degree are they natives and immigrants respectively?
- To what extent does prior knowledge and demographics affect the user's ability to scan a website?

2 | Theory

In the following chapter we describe the theory, which forms the basis for our methodology, the development of our test plan and discussion of our test results. We present the framework of information architecture and how we position this thesis in information architecture. Then we discuss the foundation of our inspiration for the thesis with the terminology of Marc Prensky, the different parameters of the users we want to test with Demographics and Prior Knowledge. We present information on Eye tracking and its uses. Our tools for acquiring the data; OGAMA and The Eye Tribe will be described and the chapter will end with theory on ethics in relation to research.

2.1 | Information Architecture

Information architecture is a field that can be hard de give a clear-cut and precise definition of. It is discussed how much it encompasses and where to draw the line for what it covers (Dillon 2002). Is it just the structuring and labelling of data that Dillon calls little-IA or is it big-IA, which includes usability and user experience design? We find that the big-IA is the best description, as it includes more aspects, such as user experience, which supports the design of a well-structured and labelled information system. Morville and Rosenfeld gives the following definition of information architecture, in their book 'Information Architecture for the World Wide Web' from 2007, which we find to be a good definition:

1. The structural design of shared information environments.
2. The combination of organization, labeling, search, and navigation systems within web sites and intranets.
3. The art and science of shaping information products and experiences to support usability and findability.
4. An emerging discipline and community of practice focused on bringing principles of design and architecture to the digital landscape.

(Morville and Rosenfeld 2007)

This definition is broad and covers not only the act of organizing the content, but also how it can be evaluated and designed to the user's needs. The reason it is hard to make a single definition that everyone can agree on, is because information architecture is a field that is in-between other fields and overlapping these as well. It includes aspects of knowledge management and data management, but it is neither of those (Morville and Rosenfeld 2007). Instead it is about many different types of information and how it can be presented and structured. Resmini and Rosati defines information architecture as *"… a professional practice and field of studies focused on* *solving the basic problems of accessing, and using, the vast amounts of information available today."*(Resmini and Rosati 2012). Evernden and Evernden defines information architecture as *"a foundation discipline describing the theory, principles, guidelines, standards conventions and factors for managing information as a resource."*(Evernden and Evernden 2003). These definitions all support Morville and Rosenfeld as they deal with both problem solving and managing of information.

In the following we will discuss information architecture aspects relevant to answering our problem statement.

2.1.1 | Users

As information architecture, according to our chosen definition, includes the users, their needs and sees them as one of the most important factors, we will discuss their place in the field of information architecture and what aspects about them that are relevant to this thesis.

It is a common mistake to over-simplify users' information seeking behavior, according to Morville and Rosenfeld, which does not take into account all the different reasons for them to visit a website (Morville and Rosenfeld 2007). It is not just a user visiting a website with a very specific information need, but sometimes they are just looking around, want to learn something, but are not sure what it is called and so. To facilitate these different reasons for them to visit a website, it is important to understand them and include them in the design process.

To understand the users and participants in our study, we will go through the important aspects of the users in the following and how we define them.

2.1.1.1 | Digital natives and immigrants

We use the work of Mark Prensky as inspiration together with our own observations of elderly relatives that inspired us. Mark Prensky provides a definition of the young and elderly, which we will use and describe in more detail below. This definition is useful to distinguish between the different groups and can be used to describe the differences.

Marc Prensky is an American writer and speaker. His written work is focused on education and learning, while he has a background in teaching at the primary, secondary and college level("Meet Marc Prensky" 2016).

In 2001, he wrote the article "Digital Natives, Digital Immigrants" in which he defined two categories of technology users; digital natives and digital immigrants. He wrote a follow up article later that year in which he explored the supposed difference between the minds of teachers and students in the year of 2001. In the following, both of these articles will be presented shortly accompanied by criticism of the material, alongside an elaboration on why Prensky's work is part of this thesis.

"Digital Natives, Digital Immigrants" was published in the September/October issue of "On the Horizon: The Strategic Planning Resource for Education Professionals" in 2001. In five pages Marc Prensky describes a divide between technology users. A divide, which he claims, is determined by when an individual have been introduced to technology. Prensky argues that users who have experienced modern technology after reaching a certain stage in life, have a different way of interacting with said technology than users who have been around the technology since a very young age. In short; digital natives are users who have been using modern technology since an early age and therefore thinks it to be second nature and an integral part of their lifestyle. On the other hand, digital immigrants are users who have adopted technology at a later stage in life as a necessity to keep up with modern society. In Prensky's view, technology is not approached nor understood very well by these people as they have adopted it not from childhood, but later in life. As Prensky puts it: *"The importance of the distinction is this: As Digital Immigrants learn – like all immigrants, some better than others – to adapt to their environment, they always retain, to some degree, their "accent", that is their foot in the past."*
(Prensky 2001a)

Prensky's point of view is from that of a teacher and it is also the point of the article. Prensky makes the argument, in both this article and the next that students think and act differently around technology than their teachers, causing issues when the tool for learning is said technology. In the article he states that teachers must integrate newer learning methods in their teaching, if they want to get through to students as these students have particular learning preferences.

"Digital Natives, Digital Immigrants Part 2: Do They Really Think Differently?" was published in the November/December issue of "On the Horizon: The Strategic Planning Resource for Education Professionals" in 2001.

This article focuses on explaining through neurobiology, social psychology and studies of children playing learning games, how the two groups think differently about technology. In his article Prensky writes:

"We now know that brains that undergo different developmental experiences develop differently, and that people who undergo different inputs from the culture that surrounds them think differently."

(Prensky 2001b)

He argues that because of the social and technological environment, young people think differently because their minds have been affected differently than the generations before them.

This second article aims to prove what Prensky said in his first article by including psychology and biology. To quote the article:

"Based on the latest research in neurobiology, there is no longer any question that stimulation of various kinds actually changes brain structures and affects the way people think, and that these transformations go on throughout life."

(Prensky 2001b)

These articles form a whole in which Prensky seeks to make a postulate and later try to support this postulate through research done by others. This has led to criticism of the postulate for lacking any empirical evidence to support its claim. Bennett, Maton and Kervin writes in a critical review of the subject that Prensky's postulates lacks empirical weight, as he had done no empirical testing or research to support his claim and instead had based a lot of it on research done by others, which did not take into account the amount of "natives" who had access to technology (Bennett, Maton, and Kervin 2008).

Marc Prensky's articles and his concept of natives and immigrants have been presented in this thesis because it serves as our main inspiration for the thesis content. It is important to note that we will not delve into the issues that Prensky foresaw when he wrote these articles 15 years ago. We will not delve into learning theory as well, as our main focus lies strictly within trying to discover similarities or differences between age groups. Prensky's work has served as an inspiration to us as we see a compelling argument in his postulation and the notion of a difference in abilities, based on whether a user is born into a world with the technology or having to learn how to use it. It should also be noted that his writing is dated at this point but is still relevant. We live in a digital age now so more than ever the need for people to be capable of performing tasks with modern technology is becoming a necessity. The digitalization of several parts of the Danish government has put every citizen's ability to use these services into question.

Needless to say, the concept of different levels of capability amongst users has not been ignored since Prensky's articles in 2001. A lot of articles and research have been conducted both on performance and understanding. A lot of this research has been in the field of age, trying to discover whether or not age has an impact in understanding and using modern technology. In the next section the concept of age and other demographics will be explored.

2.1.1.2 | Demographics

Demographics is a parameter this thesis cannot ignore in its aim to investigate the problem statement. Even though age is the primary focus, other parameters needs to be acknowledged as well. Similar to prior knowledge these things can have an effect on the subject at hand. Demographics is a term that encompasses several characteristics which ties to a person's age, gender, country of origin etc. Demographics have been used in wide array of studies in regards to information searching behavior. Williamson used demographics to investigate how elders search behavior in Australia was influenced by their physical, social and cultural environments(Williamson 1997). Dunne examined the information behavior of battered women and accounted for several barriers these women face when searching for information(Dunne 2002). Hamer studied the information needs of young men coming out and forming a gay identity(Hamer 2003). All of these studies have in common that they use demographics as a tool for creating a group of people for studying. Applying demographic parameters, they were able to take a large group and size it down to a more focused one; instead of Hamer investigating how gay people search for information when trying to come out he instead decided to narrow it down to young men, using bot age and gender.

Using demographics can provide more nuance on data and allowing for comparisons. Case states that gender as a demographic is frequently used in analyses of results(Case 2012). Our aim is to use demographics as a way of structuring our data and open up the possibility of seeing patterns and tendencies. Our usage of demographics will be presented in 3.1.4 | Questionnaire on page 58.

2.1.1.3 | Prior Knowledge

Searching online for prior knowledge provides a lot of results and many of those are about prior knowledge in learning. This focuses on how the teacher activates the students prior knowledge, to aid them in their learning ("Prior Knowledge | Learning Sciences" 2016). This is the basis of what prior knowledge is, as it is all about how people use their prior knowledge to solve new problems. In the following we will go into more detail and give a more thorough explanation of it.

Russell-Rose and Tate discusses their notion of users being divided into two groups; novices and experts, in their book 'Designing the Search Experience' from 2013 (Russell-Rose and Tate 2013). This is a very crude way to view users, but can provide a simple framework to view the users. Other definitions and notions can be used to expand on this framework, which we will do in this chapter.

The difference between the two groups is that they each choose the tool most suitable to the skill, when they try to solve a problem. This of course affects the way the users approach a problem and for us it is relevant, as any difference between our two age groups, might affect their approach to solving tasks. Russell-Rose and Tate broadens their notion by adding a distinction between domain expertise and technical expertise. Domain expertise being the knowledge about a certain subject such as taxes and technical expertise being knowledge about the tools used on the domain, such as using the Internet (Russell-Rose and Tate 2013).

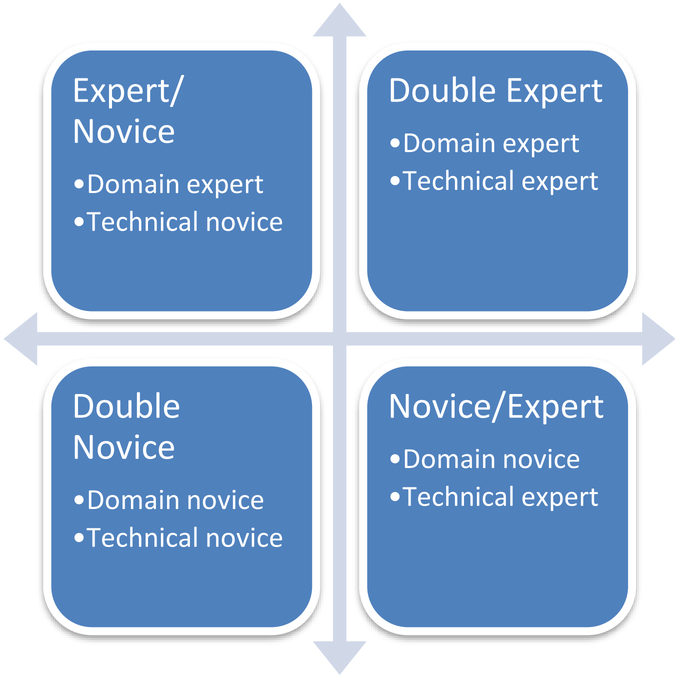


Figure 1 – Expert/Novice Matrix

Using these distinctions, they have made a matrix of two dimensions shown in Figure 1 – Expert/Novice Matrix. The vertical axis indicates the level of domain expertise, where the bottom is novice level and top is expert level. Horizontally it is technical expertise with novice level on the left and expert level on the right (Russell-Rose and Tate 2013). These distinctions can be used to identify, what the user's strengths and weaknesses are, when they search for information online.

In the following we will go into more detail on how the expertise, or knowledge as it is described below, is acquired and how it influences the user's search for information. This can be used to expand on Russell-Rose and Tate's matrix.

It has been shown as early as 1997 that prior knowledge increases the ability of the elderly to navigate and find information online (Meyer et al. 1997). This was shown by comparing two groups testing the same website, but with one group having received training in navigation. The two groups consisted of older and younger participants, but it is not clear what constitutes an older or younger participant. The test was conducted on a website, made up of 19 pages from the Scientific American website. The participants was divided into two groups, where one received an introduction to the navigational systems and the other received the same, but with hands-on training (Meyer et al. 1997). The results showed a reduction of the steps taken to accomplish each task, from an average of 9.3 to 7.8 (Meyer et al. 1997) which clearly indicates the importance of prior knowledge. The study also found a difference between the two age groups, but that prior knowledge reduced these differences(Meyer et al. 1997). Furthermore, it has been shown that prior knowledge increases the users effectiveness in interacting with IT-systems (Crabb and Hanson 2014).

These proves that prior knowledge regarding the technical expertise, described by Russel-Rose and Tate, is important, as it reduces the differences between the age groups and the necessary steps a user takes, to complete a task.

To acquire the technical expertise, Hurtienne and Langdon identified four levels of prior knowledge acquisition, in their 'Prior knowledge in inclusive design: The older, the more intuitive?' study. These four levels of knowledge acquisition is based on previous research done by Jörn Hurtienne (Hurtienne and Langdon 2009).

The acquisition of prior knowledge can be done on four levels:

1. Innate (Genes and parental state)
2. Sensorimotor (Learned through childhood)
3. Culture (What is common in the person's culture)
4. Expertise (Specific knowledge from hobbies and profession)

(Hurtienne and Langdon 2009)

Innate learning has to do with how we are formed and what genes are passed on to us. Sensorimotor are what we learn during childhood, such as recognizing faces. What is interesting to us is the last two levels. Culture being about what the people around you do and accept. This can for instance be about the attitudes toward certain types of technology. The last one, Expertise, can be about the persons use of IT systems in their hobbies or professional life (Hurtienne and Langdon 2009). These different types of knowledge acquisition influence the technical expertise of the user, described by Russell-Rose and Tate earlier.

The acquired knowledge can be further described, to provide more insight into what the domain and technical expertise are and how it is used.

Knowledge consists of two types of intelligence: Fluid and Crystalized intelligence (Beier and Ackerman 2005). Fluid intelligence is the logic each person uses to solve new problems and acquire new knowledge. Crystalized intelligence is the knowledge already acquired (Beier and Ackerman 2005). The difference in these two types of intelligence is that Fluid intelligence is used to solve problems logically, where the problem solver has no prior knowledge or cannot apply it. Crystalized intelligence is used to solve problems based on past experience and accumulated knowledge (Crabb and Hanson 2014). This means that the problem solver applies their experience from past problems to a new one and solves it by knowing a specific solution that worked the last time or on a similar problem (Hailikari and others 2010).

Prior knowledge are, according to Dochy, Segers and Buehl, a combination of explicit and tacit knowledge (Dochy, Segers, and Buehl 1999). This means that it comprises of both the knowledge, which we know we have and can clearly describe, and the subconscious knowledge we have but can be difficult to describe (Paavola, Lipponen, and Hakkarainen 2004). Basing our decisions on knowledge we can describe and what we "just know", helps us do tasks naturally, but can also mean we perform the tasks based on incorrect or non-compatible knowledge.

Having prior knowledge on the specific subject is equally important, as it can help the user locate relevant information by having a mental model of where to look (Beier and Ackerman 2005). This means that the user might not have a lot of prior knowledge on the use of the Internet for information seeking, but by having knowledge about a certain subject; they can more easily understand the structuring of data. This is in line with domain expertise, as described by Russell-Rose and Tate above.

What is important about prior knowledge is that it is correct and relevant to the current situation a person are in (Shapiro 2004). Incorrect prior knowledge can negatively influence how much and what a person can extract, when reading a new text (Shapiro 2004). This further supports our hypothesis that the difference in how the different generations have found and acquired new knowledge, during their lifetime might influence how they use the Internet.

We intend to use the participants level of prior knowledge to add more details to our comparison of the young and elderly participants, to investigate how big a part of the potential differences is from the age compared to prior knowledge.

Measuring the prior knowledge of people can be done in different ways, which we will discuss in more detail in 3.1.4 | Questionnaire on page 58.

2.1.1.4 | Search behavior in elderly users

Our focus is in how the older users use and view websites, so in the following we will describe existing research on older users' behavior and how the previously described theory influence the older users search behavior.

Statistics from the Danish national statistics center show that 85% of the 16-64 years old use a computer every day, while the same number for the 65-89 years old are only 57% (Lauterbach 2015). This shows that the older users are not as used to computers, but of those who uses computers, most are accessing the Internet daily. 91% of the 16-64 years old are online daily, while the number for the 65-89 years is 54%, which is only 3% point less than the number of daily computer users (Lauterbach 2015). This indicates that the older users of computers primarily use them for accessing the internet. The older users are primarily using the Internet for e-banking and emails, while the younger users use the Internet for more varied tasks, such e-shopping (Lauterbach 2015). This could indicate that the older users are using their computers out of necessity and because they have to. It is interesting to note that only 1% of the users between 16-64 never uses the Internet, while the number for those between 65 and 89 are 23%, or nearly a quarter (Lauterbach 2015). A study by Berner et al. on the differences in demographics for Internet use in Sweden, found that there were differences in how much the elderly uses the Internet, based on their educational level and whether they live in urban or rural areas (Berner et al. 2015). The study was conducted on a large sampling size of 7,181 adults between 59 and 100 years old (Berner et al. 2015). These studies show that some older users have not embraced the Internet in the same manner as the younger and can be a problem in a more and more digitalized world. This difference between the age groups affects the elderly users' technical expertise in relation to information searching online.

These statistics shows a difference between the elderly and young, but it does not say much on how the older users actually use the Internet. Below we will discuss how the elderly compares to the younger users, in how long it takes them to complete tasks online, how many errors they make, how they scan for information online and how important prior knowledge are to their performance.

The time it takes for users to complete tasks online, can show how proficient they are at navigation and using the Internet. It also indicates differences between groups if one takes longer to complete the tasks. To investigate the differences, we have a number of studies that investigates if any differences exists. We will go through these studies and discuss their findings.

Artis and Kleiner found that older users (63-83 years old) were slower, when compared to younger users (18-23 years old), in their study The Effects of Age and the Design of Web-Based Training on Computer Task Performance from 2006 (Artis and Kleiner 2006). In this study they explored the effect of age and training on the two groups performances. They found that the two age groups both benefitted from training, but that the older group where slower to complete the task, both before and after receiving training (Artis and Kleiner 2006). Even though the older users took longer to complete the tasks, it was found that they could learn new skills, just as well as the younger, but it took longer as well (Artis and Kleiner 2006).

In 2002 Kurniawan, Zaphiris and Ellis made a study to compare the time spent and amount of errors made by younger (mean age of 26.8) and older adults (mean age of 67.5) (Kurniawan, Zaphiris, and Ellis 2002). This study found, just as Artis and Kleiner, that the time spent by the older adults was longer than the younger adults and that they used more clicks to complete the tasks (Kurniawan, Zaphiris, and Ellis 2002).

In a study to find differences between young and old adults for entering data in an e-health program in 2007, by Biterman, Lerner and Bitterman, it was found that the older users (72.2 +/-5.5 years) were significantly slower than the younger users (30.4 +/-4.9 years) (Biterman, Lerner, and Bitterman 2007). The difference was so large that the older users actually used more than double the time compared to the younger users (Biterman, Lerner, and Bitterman 2007).

In a study from 1997 by Mead et al. it was found that the older users (64-81 years old) were less likely to complete tasks, than younger users (19-36 years old). This was, however, because they had a time limit, but it still showed that the older users were slower than the younger users (Mead et al. 1997). They also found that the older users used more clicks and scrolled more to complete the tasks. The same researchers published another study in 1997, were they investigated whether training could reduce the differences and in this study they confirmed their findings in the previous study (Meyer et al. 1997). The results of their study on training impact on the older user's abilities will be discussed later, when discussing prior knowledge.

In a study of 49 people between 20 and 82 years old, Chadwick-Dias, McNulty and Tullis found that the users over 55 took longer to complete tasks on a prototype web site than the younger users (Chadwick-Dias, McNulty, and Tullis 2003). The older users were in fact almost 50% slower than their younger counterparts (Chadwick-Dias, McNulty, and Tullis 2003).

Czaja et al. made a study in 2001, where they had 117 participants conduct complex information search for three days and compared the results, by dividing the participants into three age groups of 20-39, 40-59 and 60-75 years old (Czaja et al. 2001). In this study it was found the oldest group were the slowest and even though all groups improved during the three days, they remained the slowest group (Czaja et al. 2001).

In January 2012 Etcheverry, Terrier and Marquié. published a study comparing memory and recollection for younger (mean age of 22.31 years) and older users (mean age of 64.54 years), when seeking information online (Etcheverry, Terrier, and Marquié 2012). During this study they found that the older users were slower at completing tasks online, but that the difference was reduced when they conducted content-oriented searches compared to navigation-oriented searches (Etcheverry, Terrier, and Marquié 2012). Those findings inspired them to another study, which was published in

October 2012, by them, Baccino and Mojahid, where they investigated whether the difference really did exist and the reasons for these differences, using eye tracking (Etcheverry et al. 2012). Their study confirmed that a difference existed and they found that not only were the older participants not just better at content-oriented searches compared to navigation-oriented, they were even better than the younger participants, when conducting content-oriented searchers (Etcheverry et al. 2012). The differences in the study can be attributed the fact that then findings in their January study was a byproduct of their main study goals, while their October study was focused on exacerbating the findings to see if differences exist and why. This can mean that the results are not transferable to the real world, as they are the only ones to have results, where the elderly are the fastest. It does not, however, mean that there is no difference in how fast the elderly is, when searching for different types of information. Reasons for differences were that the older users tended to revisit previously seen menu items and spend more time looking at them and thus employ more verification strategies (Etcheverry et al. 2012). The reasons for the difference in which age group is the fastest will be discussed more in the section prior knowledge later.

It is clear that the elderly users are slower than young users, when it comes to complete tasks online. There seems to be some difference in how much slower they are, or as in one study actually faster. This has to do with the type of search conducted by the elderly. The question is why are there differences in how fast the elderly and young find information online. To answer this question, we look at how many errors the older users make, compared to the younger users.

Most of the studies comparing the time to complete tasks for young and elderly users, also recorded and compared the errors made by each group. Artis and Kleiner found that there were no differences in the amount of errors made by the 18-23 year olds and the 63-83 year olds (Artis and Kleiner 2006). This was the same for Kurniawan, Zaphiris and Ellis in their 2002 study (Kurniawan, Zaphiris, and Ellis 2002). Biterman, Lerner and Bitterman found that there were no differences in the number of error made by the 25 elderly and 25 young test participants, in their study (Biterman, Lerner, and Bitterman 2007). This is despite, as mentioned earlier, the elderly was twice as slow to complete the tasks in their study. Mead et al. found that in their tests in 1997, the elderly users tended to revisit already visited pages more often and some even got lost and did not know what page they were currently on (Meyer et al. 1997). They did, however, note that the elderly and young users all tended to start over, by visiting the home page equally often (Meyer et al. 1997). This could indicate that the two age groups handle making errors in similar manner. In the 2003 study by Chadwick-Diaz, McNulty and Tullis study, they gave the participants the option to give up, if they would normally, in a similar situation, give up or ask for help. Furthermore, they were asked to move on, if they made four errors (Chadwick-Dias, McNulty, and Tullis 2003). In this study they found that the 55 years old and older participants were more likely to fail at completing tasks, when compared to the younger participants. This might be contributed to not letting the elderly attempt to complete the tasks in their own time, as previous studies shows they require (Artis and Kleiner 2006; Kurniawan, Zaphiris, and Ellis 2002; Biterman, Lerner, and Bitterman 2007; Meyer et al. 1997; Chadwick-Dias, McNulty, and Tullis 2003; Czaja et al. 2001; Etcheverry et al. 2012; Etcheverry, Terrier, and Marquié 2012).

Seeing that the difference in completion time for the two age groups are not related to the older users making more errors than the younger and are just as proficient in completing the tasks, it might be because their prior knowledge or lack thereof slows them. We have found that there is a difference in how many older and younger users use the Internet daily, which can influence their technical expertise in relation to using the Internet to solve tasks and this lack of experience might be the reason for them being slower. We discuss this below.

Prior knowledge is, as mentioned in the previous chapter, the knowledge people use to solve new tasks, by applying what they learned doing the same task previously or something similar. The effects of prior knowledge on Internet users have been described in a number of studies. Artis and Kleiner attributed the differences in the age groups to the lack of repetition and training for the older participants, but found that they could be taught how to complete tasks online, but at a slower rate than the young users (Artis and Kleiner 2006). Etcheverry et al. argued that the crystallized intelligence, described earlier in chapter 2.1.1.3 | Prior Knowledge on page 14, was the reason for differences and could account for the fact that there were differences in the type of information search (Etcheverry et al. 2012). The prior knowledge of the elderly users could have helped them, when performing content-oriented searches, as they could draw on crystallized intelligence on domain expertise of content searching in newspapers, books and so on. This was speculated in the first study by Etcheverry, Terrier and Marquié (Etcheverry, Terrier, and Marquié 2012) and substantiated by their findings, when involving more researchers in their study later in 2012 (Etcheverry et al. 2012). In this later study they found that the participants prior knowledge helps them locate the correct area on the screen, but nor the correct link, which they have to look for more thoroughly (Etcheverry et al. 2012). Laberge and Scialfa conducted a study in 2005, to investigate the influences of age, prior knowledge and cognitive abilities on 41 elderly participants between 19 and 83 years old (Laberge and Scialfa 2005). They found that the time differences was due to the participants domain expertise on a tourism web site and technical expertise on Internet usage (Laberge and Scialfa 2005). They further states that the most important factor, when it comes to users performance in relation to solving tasks online, is the prior knowledge regarding computers and the Internet, while prior knowledge on the specific domain are important as well (Laberge and Scialfa 2005). They argue that the difference between content-oriented and navigation-oriented searching, which they found just as Etcheverry et al., can be explained by the same arguments as put forward by Etcheverry et al. They argue that experience using the Internet can help the user understand and orient themselves, when performing a navigation-oriented search. The prior knowledge on the domain aids the user in content-oriented searching (Laberge and Scialfa 2005). As we described earlier, it is the younger people who use the Internet the most and thus have more prior knowledge on navigation-oriented searching, while the older people can have the same or more prior knowledge on the domain and can employ it on content-oriented searches. This is further supported by Czaja et al. who found that navigation efficiency was influenced by prior knowledge and not age itself (Czaja et al. 2001). They further found that both older and younger users benefitted from training, as it provided them with more prior knowledge and found that the older users benefitted the most, as they started at a lower level (Czaja et al. 2001). This is supported by findings of Meyer et al. who concluded that all age groups benefitted from training (Meyer et al. 1997).

It is clear that the elderly Internet users are performing tasks online in different manner, when compared to the younger users. They take more time in completing the tasks, but does not, however, make more mistakes than the young users. It is instead related to their prior knowledge on the use of computers, the specific domain and Internet usage, meaning their technical and domain expertise. The elderly users spend more time conducting verification strategies, where they have found the correct link, but take the time to ensure that the previously visited or viewed pages are not better matches. The fact that the prior knowledge is so important are what accounts for the difference in timed needed to complete tasks, between the two age groups, as argued in the studies described above. These differences should be present in our own study as well and can be able to explain why other differences occur that we might observe.

2.1.2 | Information Architecture Systems

Information architecture consists of four blocks that, together, makes up a web sites information architecture. It is important to understand these four parts, as these affects how users search for and navigate an information architecture.

The four blocks of information architecture:

- Organization systems
- Navigation systems
- Labelling systems
- Search systems

(Morville and Rosenfeld 2007)

2.1.2.1 | Organization Systems

Organization systems are how the web site is structured and how the content are grouped (Morville and Rosenfeld 2007). Maloney and Bracke argues that organization on a site refers to the logical groupings and the content and services that are available to the user(Maloney and Bracke 2004). The organizational system of a web site directly affects how easy or difficult it is for the users to find the information they are looking for. To make as usable an organization system as possible, some challenges has to be addressed.

The classification of content can be difficult, as it can have multiple meanings and be ambiguous (Morville and Rosenfeld 2007). This is a problem, when trying to decide where to place content, as it might be fitting to place it in more than one place.

As the web contains many different types of content and are heterogeneous in its nature, it is not possible to make an organization system that fits all web sites and content (Morville and Rosenfeld 2007). Instead it is important to develop it, so it fits the content and the needs of the user.

Just as the web in itself is heterogeneous, so are the users and the way they structure content (Morville and Rosenfeld 2007). This means that an organizational system that fits one person perfectly can be extremely confusing and illogical to another person. This requires the organization system to be designed to fit many different users' perspective on how the organization should be. To do this, it is necessary to include them in the design process and attempt to create the organization system in a way that fits as many as possible. It can be necessary to provide more than one route to the content on the web site to accomplish it (Morville and Rosenfeld 2007).

Internal pressure to make an organization in certain ways, to fit to ideas by the management, marketing and more, will also influence the design process of the organization system and must be taken into account (Morville and Rosenfeld 2007). It can be difficult to balance the needs of e.g. the management on one hand and the users on the other. It is important to keep in mind that the system is being designed for the users and their needs should be the primary goal to fulfill.

Overcoming the challenges of designing the organization system, it is important to know that it consists of two parts: The organization scheme and organization structure (Morville and Rosenfeld 2007).

The organization scheme is how the content are grouped by shared characteristics and is related to the above mentioned challenges, while the organization structure is how the organization scheme is connected and relationships between the groups (Morville and Rosenfeld 2007). Maloney and Bracke writes that a sites organization scheme "…

will often serve as the foundation for the primary navigational choices on a site's main menu or primary navigational bar." (Maloney and Bracke 2004).

The organization scheme can be done differently, but two different types encompasses the various methods. The objective method is easy to design and maintain, as it focuses on clear distinctions between the groups. I can be alphabetical, chronological, geographical and so on (Morville and Rosenfeld 2007). It is, as stated, easy to design and maintain, but it requires the users to know exactly what they are looking for. The subjective method on the other hand can be difficult to design and maintain, but are easier for users, who does not know exactly where or what they are looking for (Morville and Rosenfeld 2007). The subjective method groups the content by topics, audience and more. This is done so the user can look for a specific topic, such as specific types of cars, so they do not have to scroll unending lists of vans, sports cars and trucks to find a station wagon. By doing it this way, the user does not need to know all the facts and exactly what they are looking for.

The organization structure is, as mentioned earlier, how the organization scheme is connected internally. This is done by structuring the content in a hierarchy, where the content is structured in the groups of the organization scheme. The challenge is to structure it in a manner that are easy to navigate through. Decisions has to be made on how wide and deep the structure should be, to help the users find their desired information (Morville and Rosenfeld 2007).

2.1.2.2 | Labelling systems

The purpose of labels is to, quickly and without taking up too much space, communicate what can be found by clicking them (Morville and Rosenfeld 2007). Labels make up the navigation, discussed in the next chapter, and shows the user what is in the organization schemes and what can be found in different parts of the organization structure (Morville and Rosenfeld 2007). Maloney and Bracke argues that labels should clearly communicate the meaning of the site and that they must target the user with a language they understand(Maloney and Bracke 2004).

As labels are what explains the contents of the different organization schemes to the users, it is important that they are as clear and explanatory as possible. The labels can be a number of different objects on a web site. It can be a heading for a chunk of text or a video, it can be contextual links tying related content together across the organization structure and it can be navigation options in the web site navigation (Morville and Rosenfeld 2007).

To have a successful labeling system it is important to be consistent, to help the users learn to use the system and accommodate their expectations to how they think it should perform. This is done by keeping the style the same and the language consistent (Morville and Rosenfeld 2007). It is also advisable to make the labels based on the target group, so a label is tailored to the type of users, who are interested in the specific organization scheme it labels (Morville and Rosenfeld 2007).

Designing the labeling system is done by understanding the content, the organization schemes, the users and the context how they are used. It requires careful consideration of all these aspects to design the best labeling solution (Morville and Rosenfeld 2007).

Etcheverry et al. described lexical matching, which is how users read labels and words and attempt to match them to a subject they are looking for (Etcheverry et al. 2012). This method of searching for information makes the labels incredibly important, as they must be able to hit as many matches as possible for as many users as possible.

2.1.2.3 | Navigation systems

Navigation on the web is not just how the user gets around in the organization system, but also how they orient themselves and understands where in the system they are. Maloney and Bracke identifies navigation systems as navigation structures which *"… define the relationships between content and service elements of a site, and between groupings in the site's organization."*(Maloney and Bracke 2004).

Navigation can be divided into two types of navigation: The embedded navigation and the supplemental navigation (Morville and Rosenfeld 2007). The embedded navigation is often wrapped around the top and sides, are inserted in the sites content and functions as the main navigation and to provide information on the location in the organization system. Maloney and Bracke calls these local relationship schemes and writes that they are used to define clear relationship between content on a website(Maloney and Bracke 2004). The supplemental navigation are extra navigational options, which are outside the main structure with content and include guides, indexes, search features and more (Morville and Rosenfeld 2007). The navigational features included in browsers support the features of the web sites, by providing methods for moving back and forward, bookmark pages and providing color coding to show which links have been visited (Morville and Rosenfeld 2007).

If we are to explain the two types of navigation, embedded and supplemental in more detail, we find that they do a wide variety of jobs.

Starting with embedded navigation we find the element global navigation, which is constant across all pages of the website. Local navigation changes from page to page and provide special navigation from the current position in the organization system. Finally we have contextual navigation which provides specific links to related content, which might anywhere in the organization system, even in other organization schemes (Morville and Rosenfeld 2007).

Supplemental navigation includes site maps, site indexes, guides and search (Morville and Rosenfeld 2007). All these provide extra complementary navigation methods to the user by showing the entire web site structure (site map), a list of all the content (site index), tips and tricks to navigating the specific web site (guide) and the ability to look for specific terms (search) (Morville and Rosenfeld 2007). Search is a complex navigation method, which in itself is a building block of information architecture, which will be described below.

James Kalbach divides navigation into three categories in his book Designing Web Navigation from 2007 (Kalbach 2007). He divides them into categories based on what and how they connect content. The three categories are: Structural, Associative and Utility. The main difference from Morville and Rosenfeld to Kalbach is that Morville and Rosenfeld divides it by whether the navigation is integrated into the content or not. Kalbach divides it by how the navigation links the content together. His structural navigation includes the global and local navigation and are grouped because they provide methods of moving up and down the organization structure (Kalbach 2007). This type of navigation are described as vertical navigation by Morville and Rosenfeld (Morville and Rosenfeld 2007). The associative navigation includes contextual navigation, adaptive navigation like most read articles and most popular product, quick links to the most popular content and footer navigation that provides pages with extra information that does not fit into any specific organization schemes (Kalbach 2007). Contextual, adaptive and quick links provide navigational options across the organization system, which Morville and Rosenfeld calls lateral navigation (Morville and Rosenfeld 2007). Kalbach's last type of navigation is utility navigation and are the navigational methods that help the user use the web site and are similar to Morville and Rosenfeld's supplemental navigation (Kalbach 2007). Utility navigation includes internal page navigation, search forms and help pages.

Furthermore, Kalbach includes the pages of the website in his navigational options as they can help facilitate navigation. He divides the pages into three types: content, functional and navigation pages (Kalbach 2007). The content pages are of course just pages of content. The functional pages allows the user to complete tasks, such as signing up for a newsletter and finally we have the navigation pages that facilitates further navigation by providing information on the organizational structure and further navigational options (Kalbach 2007).

Kalbach's navigational methods elaborates on those of Morville and Rosenfeld, but are very similar. They show that there are many different ways of providing navigation methods to tie the organization system of a web site together, which can help the users but maybe also confuse them, if there are many different methods or the web site uses a method not familiar to the user.

2.1.2.4 | Search systems

The final system of information architecture is search systems, which are a supplemental navigation method, but a complex and powerful one. Search lets the user circumvent the other navigation systems and go directly to the desired information, but only if it is well designed. Search is not only entering a search query, but it is also how the interface is displayed, the query language can be difficult to use and the algorithm making up the search engine can produce unanticipated results (Morville and Rosenfeld 2007). Furthermore, the user's query must be similar to how the labeling have been done and the terminology for the subject in general. The results page can confuse a user if it is not properly organized or cluttered with information (Morville and Rosenfeld 2007).

2.1.2.5 | Summation

The four systems, organization, labeling, navigation and search, are all integral parts of information architecture and neither of them can stand on its own. An organization system is needed for navigation to take place and labels are needed to give meaning to the navigation. Labels are necessary to describe the organization schemes and labels need the navigation systems to exist. Search is a part of the navigation and utilizes labels when searching. The difficulties in designing the organization system will often require a search system to help the user's when they do not understand the organization system. These four systems are directly influencing the users' abilities to find their desired information and are therefore important aspects to consider, when investigating how users use a web site.

The following section will provide a thorough elaboration on the concept of eye tracking, our chosen method of testing users in this thesis.

2.2 | Eye tracking

Testing a project, such as a website, can be done many different ways and with varying problems and upsides. Many suffer from distracting the test participant during the test, for instance during a think aloud test, or by relying on the participants ability to accurately remember and explain their actions (Busjahn et al. 2014). Eye tracking lets the user perform tasks without having to think about remembering, what and why they perform certain actions or having to explain them during the test.

Eye tracking is based on recording the fixations and saccades of the test participant (Busjahn et al. 2014). Fixations are data showing where the participant looks and saccades are the paths their eyes take between each fixation (Busjahn et al. 2014). Eyes tend to be moving many times each second and are rarely looking a one specific point for long at a time and these quick eye movements can show us how users scan information, where they start, where they end and so on (Ehmke and Wilson 2007). Ehmke and Wilson explains this by the eye-mind hypothesis, which states what a user is looking at is what they think about or are focused on (Ehmke and Wilson 2007).

Recording these fixations and saccades are generally done by shining infrared rays at the test participant's eyes, which are then reflected back from the pupils. The reflected rays are then captured by the eye tracking hardware and the software processes it, by analyzing the angles the light comes from, to calculate where the participant was looking (Zhu, Fujimura, and Ji 2002).

Being able to know how the user locates relevant information, how they scan for it and in what order is of course relevant, when trying to design the most usable information architecture. It has been found by looking at the divide between novices and experts, as described in chapter 2.1.1.3 | Prior Knowledge on page 14, that repetitive eye patterns are associated with novices (Busjahn et al. 2014). This shows that eye tracking can be used to identify differences between different groups.

Only recently have the technology become affordable, accurate and non-intrusive enough for it to become widely used in businesses and research, despite it being a research subject for a long time (Kurzhals et al. 2014). Now it is, however, easier and cheaper to use with the onset of freeware such as OGAMA and affordable hardware like The Eye Tribe, which will be described in more detail in 3.1.3.1 | Hardware and software on page 54.

Eye tracking technology are currently being used for many different tasks. It is used in neuroscience for research into vision and how we use it (Duchowski 2002). In psychology it is used for research into how humans read, perceive art, scene and films and how we perform visual searchers on a website for instance (Duchowski 2002). Eye tracking is used in evaluating human factors in designing work areas for humans. This include optimizing the design and layout of aircraft cockpits and cars. It is furthermore used in these areas to understand and prevent accidents, by gaining understanding of what we pay attention to in different situations (Duchowski 2002). Eye tracking are used in advertising and marketing as well, to evaluate the work done and provide information on the best design and placement (Duchowski 2002). It is used to evaluate what features in an advertisement gets the most attention from test participants, which allows the marketers to optimize their campaigns (Duchowski 2002). In computer science it is used in a variety of ways. It is used in usability to understand what the users focus on and where improvements can be made, but it is also used in more complex ways (Duchowski 2002). Eye tracking can be used to aid in the communication between people, such as people who have disabilities inhibiting their speech (Duchowski 2002). The famous physicist Stephen Hawking who suffers from ALS uses an eye tracker to communicate. He looks as letters and words on a screen, which is tracked and processed by the computer that in turn speaks for him ("The Computer - Stephen Hawking" 2016). It is useful in multiparty communication, such as a video conference, where it can help the users in knowing who each person are talking to, by tracking their eyes to find who they are looking at (Duchowski 2002).

By looking at the different types of data collected, by the eye tracker, it is possible to uncover a wide variety of information on the web site and user.

The combination of the order of fixations and the saccades between them forms the scan paths, which can show how and where the users search for information (Goldberg et al. 2002). This can be used to find problems with the existing structuring of data or confirm the validity of the existing one. To analyze the similarity of scan paths, Levenshtein's algorithm of edit distances are used by some eye tracking software, such as OGAMA. This algorithm counts the number edits needed to change a string, so it is the same as another string (Guan et al. 2006). In eye tracking it is done by assigning letters to areas of the screen or areas of interest on assigning these letters to fixations. The strings of letters produced by this procedure are then compared and the edit distance calculated (Guan et al. 2006). This procedure provides a percentage of the similarity between test subjects, which is useful to understand how similar of different groups or test participants look through a website.

Looking at fixations it is possible to get information on a number of user behaviors. Search efficiency can be evaluated by looking at the spread of fixations. The number of fixations can also indicate how easy it is to locate information. Repeated fixations on the same target might indicate something that is difficult to extract meaning from. The duration of fixations can indicate ease of extracting data. How quickly the first fixation happens shows how effective elements are at getting the user's attention. Long scan paths can indicate inefficient searching, which is also the same for scan path durations (Ehmke and Wilson 2007).

The number of saccades shows how much searching has been needed on the product. Long saccade duration can indicate that it is difficult to see or read parts of the product (Ehmke and Wilson 2007).

Looking at scan paths it is possible to combine the fixations and saccades and extract more information. The duration and length of scan paths indicates the search efficiency, as longer length and duration can point to the user having difficulties locating relevant information. Scan paths going back and forth between the same areas can indicate a problem with how clear and understandable the information is (Ehmke and Wilson 2007).

Building on the scan paths, it is possible to record specific areas of interest (Goldberg et al. 2002)(referred to as AOI hereafter). These are used to record for how long a user looks within a certain area of the product which is being tested. This could be images, groups of links, individual text blocks, navigational elements and more. AOIs can be used to identify what are most looked at, the order of when they are being focused on and how many are looked at (Goldberg et al. 2002). Fixations in each AOI can also be used to compare the relative importance of each (Kurzhals et al. 2014).

Common methods of presenting the data from an eye tracking study is by attention maps, which uses colors to show what areas of the test product the participants have focused on and gaze plots which shows the routes taken between the fixations that make up the attention map (Kurzhals et al. 2014). Displaying gaze transition is another method, where it is visualized the order of transitions between areas of the product (Busjahn et al. 2014). This can be used to identify what the participant focuses on and in what order, which can identify where participants get lost in their information search.

A number of problems exist for the use of eye tracking, which is mostly regarding consistency and accuracy.

Fixations are based on the software recognizing when a user looks at certain points for a minimum duration of time, but this duration varies from study to study and software to software (Goldberg and Helfman 2010). Al Maqbali et al. states that the difference in studies range from 100ms to 200ms, but for familiar objects it can be less than 100ms (Al Maqbali et al. 2013). This inconsistency can give different results from the same data as it will be analyzed differently. This problem can be reduced by ensuring that the same definition of a fixation, is used in the entire study (Goldberg and Helfman 2010). We will address this issue by using the same software and settings for the entire test.

It is also possible to misinterpret the results of the data collected. Long fixations can be seen as confusion or interest, depending on the situation. Scan paths that goes back and forth between targets can be seen as it being difficult to extract meaning from the content, but it can also be that the user is just trying to decide between different options (Ehmke and Wilson 2007). This is something we need to be aware of, as our interpretation of the collected data will influence the conclusions based on it. Researchers must be careful to consider more than one possible reason for tendencies observed as described by Kurzhal et al. (Kurzhals et al. 2014).

As described earlier, it is possible to use AOIs to help in understanding what, when and where a user is looking. Setting up these areas are mostly based on what the moderators sees as appropriate and therefore has to be done after careful consideration (Goldberg and Helfman 2010). It is important to consider what needs to be defined as individual areas of interest and how much padding each of these needs (Goldberg and Helfman 2010) as it will influence how many fixations will be included in each area of interest.

Inaccuracies because of the test participant moving during the test or issues with the calibration can also effect the validity of the collected data (Goldberg and Helfman 2010). It is further important to remember that the user will be put into an unnatural situation, which means they might not act in their usual behavior (Kurzhals et al. 2014). More on this will be described in 3.1.3.2 | Test setting and testing on page 55. This means it is important to make the calibration as accurate as possible and ensuring the participant is as comfortable as possible before starting. This can reduce how much physically they move during the test.

In general there are issues with the fact that eye tracking lacks standardization in not only the terminology, but also in the algorithms used to track eye movements (Al Maqbali et al. 2013). As mentioned earlier, it is important how the duration of a fixation is defined.

3 | Methodology

In the following we will describe how we intend to answer our problem statement and research questions. It is divided into three sections: our theoretical approach, our data collection and data analysis. In data collection we describe how we intend to collect the data and what we base our chosen methods on. In data analysis we describe how we analyze the collected data and which measures we intend to use. However, firstly we will account for our theoretical approach to the thesis.

The theoretical approach to this thesis can be considered to relate to critical rationalism("Popper: Critical Rationalism | Internet Encyclopedia of Philosophy" 2016). The purpose of this thesis is to investigate whether or not Prensky's postulate regarding digital immigrants and digital natives is true in relation to eye tracking. His postulate can be seen as the thesis' hypothesis as he believes that natives are different to immigrants. To translate this hypothesis into our framework; Young people and old people view and use a website's navigation differently. To either falsify or verify this hypothesis we create our own empirical data through extensive testing on people from both the native and immigrant group to identify differences and similarities. This data is analyzed thoroughly to answer our chosen problem statement which is closely related to the perceived differences that Prensky believes to exist. Even though this thesis can be linked to critical rationalism it is essentially not as we do not seek to falsify the hypothesis stated, but rather to seek information with the sole purpose of investigation.

Furthermore, we employ a social constructivist approach(Atwater 1996) as we maintain a critical view of the chosen theory and method and the created data throughout the thesis. The basis of the thesis is a critical view of the work of Marc Prensky. This critical approach encroaches on the concept of phenomenology(Merleau-Ponty and Smith 1996); that everything is because someone says it is. In relation to this, we test whether or not the notion of digital natives and immigrants is true, an idea which is conceived in the mind of Prensky.

With the theoretical approach accounted for, the following chapter will cover the data collection section.

3.1 | Data collection

We collect our own data for our analysis and we do so by using two different methods. We will conduct an eye tracking study and a questionnaire. The eye tracking study is intended to provide a variety of quantitative data on the participants, while the questionnaire provides background information, which can be used to analyze the eye tracking data for specific groups.

In the following we describe how we have chosen to collect our data. First we describe the participants for our test, then we describe our test plan, eye tracking and questionnaire follows next and finally we describe the ethics in relation to conducting our tests.

3.1.1 | Participants

60 participants would provide a clear and statistically sound picture of the tested groups, when conducting eye tracking (Pernice and Nielsen 2016). Having around 40 participants does give an indication of what tendencies might exist, and according to Pernice and Nielsen, an eye tracking study, should at least include 39 participants (Pernice and Nielsen 2016). This number of participants gives us an indication of whether there are any differences between the groups or not and can provide an idea of the need for further research.

Because of the equipment, we will use to test the participants, we have to exclude certain potential participants. After having contacted The Eye Tribe about any problems related to the participants age, we got at reply with the following recommendations: Participants cannot wear glasses with polarized glass or shifting focus and participants with "saggy eyelids" with partly obscured pupils can be difficult to track, will not be part of this test.

To avoid having users who are currently not using borger.dk and not having the test session turn into an instruction in the use of borger.dk, we will not include participants who has little or no experience in using the Internet. This ensures we only test current users. It is also related to our exclusion of the physical aspect of using the Internet and the user's cognitive abilities. We do not focus on issues related to accessibility, as that is a subject with extensive research already done.

3.1.2 | Test plan

We intend to develop a test plan, which shows what and how we will be conducting our tests to gather the data we need.

Test plans are useful as a type of blueprint for the tests. It can be used to communicate what the test is investigating and how, and it helps to ensure that everyone knows what their roles are (Rubin, Chisnell, and Spool 2008).

James A. Whittaker describes one of the strengths of a test plan to be that it forces the planners to consider problems and coming up with solutions (Whittaker 2012). He considers the test plan as a burden, when it is completed, as he feels it is too time consuming having to constantly update it (Whittaker 2012). We agree that it can be time consuming, especially for experienced testers. For us less experienced it can help us, as described earlier, so we intend to keep revising it.

We will base our test plan on the template provided in the second edition of 'Handbook of Usability Testing' (Rubin, Chisnell, and Spool 2008). This template forces us to consider various aspects of the planning process and provides a framework for us to use. We intend to customize it to our needs, as parts of it might not be relevant or we need to add more to it. This template is available for download following the link provided on page 91(Rubin, Chisnell, and Spool 2008).

The template is similar to other templates we have found, such as one by Derk-Jan De Grood in his book 'TestGoal – Result-Driven Testing' (Grood 2010) and one by Rex

Black in 'Managing the Testing Process' (Black 2009). Both of these examples varies mostly in the terminology used and not on what is included.

The purpose of the study described by this test plan, is to investigate whether or not age as a category alters or affects peoples' ability to use a website. In the test the participants will be tested via eye tracking to understand how they perceive or view a website. These participants will be separated by age so that benchmark can serve as a main point of comparison. Testing whether elderly is affected by age related factors in their performance on websites is important as modern society dictates that the elderly has to be able to do so, in order to receive the benefits that constitutes their rights as citizens.

The order of our test plan is identical to the order used by Rubin, Chisnell and Spool, found on page 67(Rubin, Chisnell, and Spool 2008). There is a reason for doing so; if one follows this order in their process, they will find that they most likely will not have to go back and make major changes. If the study objectives have been set first the rest can follow. Arguably it would not make sense to make the tasks as the very first thing, as one would not have determined the exact goals or the participants for the test. E.g. you would not make an elaborate set of tasks for an eye tracking study, only to find that the participants you want to test are blind. With that being said, Rubin, Chisnell and Spool explains that the probability for your test to go exactly like you planned is small. They argue that one can accommodate unwanted or unforeseen changes by developing the test plan in stages(Rubin, Chisnell, and Spool 2008). More methods to reduce issues with the tests are described in chapter 3.1.3.3 | Pilot test on page 57.

In the following, the test plan for this thesis will be accounted for, naturally the first section will cover the purpose and goals for the test plan.

3.1.2.1 | Overall objectives for the study

The goal of the test is to create data for an analysis of possible differences between young and elderly people. The data collected will be used to determine whether age related factors affects a person's ability to use websites.

The goal of this study is to:

- Measure any differences in the way young and elderly people navigate and look at websites, while performing tasks

3.1.2.2 | Research questions

- Do young and elderly people focus on different parts of a website?
- Do they navigate through a website structure differently?
- Are there differences in their time spend on completing tasks?
- Do they spend varying time looking at specific parts of the website?
- Are one group more prone to using a trial-and-error method than the other?
- Are there any tendencies in where each group starts on the website?

3.1.2.3 | Recruiting participants

Test participants will be determined first and foremost by their age. The age groups are 18-24 and 53+ years old to ensure that they are adults who might have used borger.dk and the elderly group are of a high enough age, to have experienced the advent of digitalization. We will also attempt to include participants with different demographic backgrounds. The following table is a representation of our desired participants, how many we hope to include and their diversity. In section 3.3.2 | Test reflection we will discuss how many participants we were actually able to get.

Characteristic	*Desired number of participants*
Participant type	
Old	28
Young	12
Total number of participants	40

Prior knowledge	
Low	13
Medium	14
High	13
Internet usage	
Low	5
Medium	20
High	15
Gender	
Male	20
Female	20

Table 1 – Desired participants

We will exclude people who have:

- Polarized glasses
- Glasses with shifting focus
- Very "saggy eyelids" with hardly visible pupils
- Little to no experience using the Internet
- Limited physical abilities; unable to operate a mouse or being blind

3.1.2.4 | Tasks

The tasks will be centered around the participants locating specific pieces of information on borger.dk. The tasks will be designed so the participants are forced to navigate through different parts of the website, including areas not normally relevant to them. This includes tasks relating to areas relevant to the young group, but not normally relevant to the elderly group and vice versa. Furthermore, we will include tasks that might be relevant to both groups and tasks which might not be relevant to any of the groups.

We have chosen to use the website borger.dk as it is the main hub for all Danish e-government communication between the state and the citizens, with only 10.6% of the adult population exempt from using it. This means that it is an important portal for all citizens, old and young. As 86.9% of those exempt are 55 years or older it indicates that this could be a web site that could help us investigate whether there are differences in the age groups.

Looking at the information architecture of borger.dk it provides possibilities to test the different systems mentioned in 2.1.2 | Information Architecture Systems on page 24. The organization system of the web site is extremely complex with a very broad and deep organization structure and organization schemes covering many different groups of content. This allows us to test how the two groups looks at and navigates this type of structure that the digitalization results in. The labeling of such a large organization system are also complex, where it can be difficult not to create ambiguous labels, which can confuse and irritate users. Borger.dk also demands a well thought out design of the navigation system to allow the users to navigate vertically in the organization structure, but also provide lateral options when necessary, while avoiding making a cluttered and overly complicated amount of contextual links.

The publically available part of the web site is huge on its own, but more navigational options and organization schemes are made available when the citizen are logged in. We have chosen to focus on the public part to avoid having to display any personal and sensitive data. This and other ethical considerations can be found in section 3.3.3 | Ethical considerations on page 75.

The test will consist of three tasks. One mostly relevant for the young participants and one mostly relevant for the elderly. One tasks will be neutral and relevant for both groups and two tasks will be centered on niche areas of borger.dk, which is not relevant for most people. Each task will have multiple paths through the website structure, so the use of shortcuts present on the website, can be compared between the groups. All the tasks will start on the front page of the website and have the user navigate down through the hierarchy.

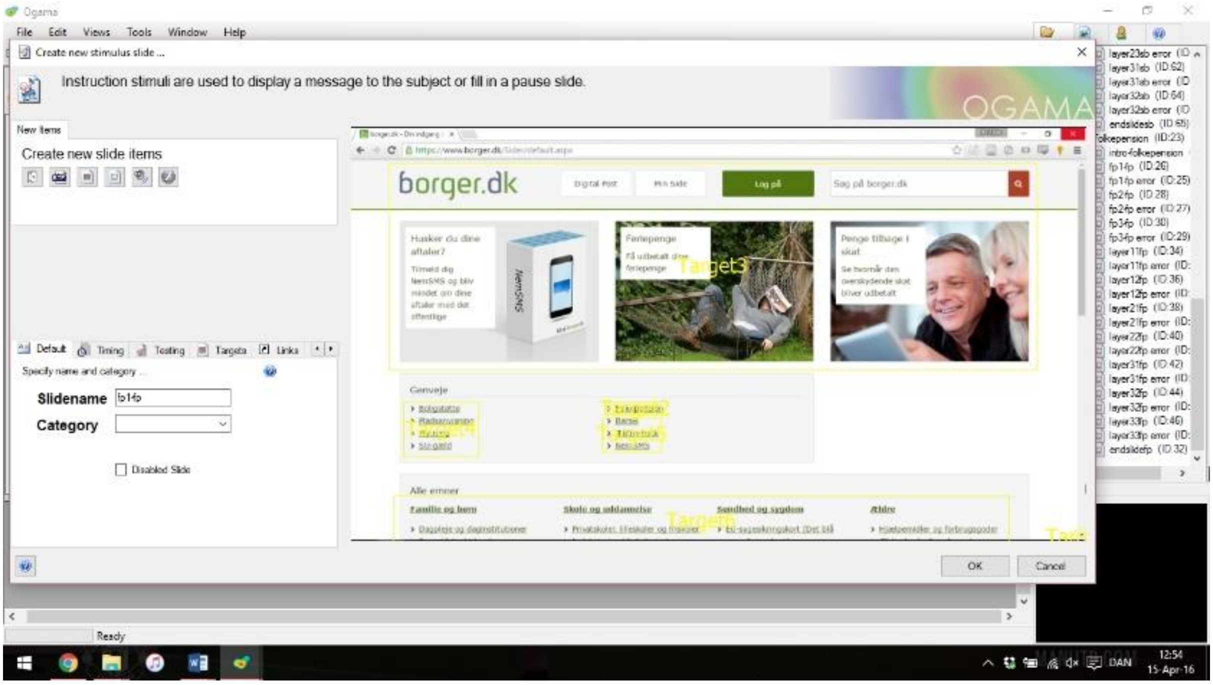
The search function of the website will not be mirrored in our test, as we intend to test the users' navigation through the navigational hierarchy of the website. This means that we will not be testing the search systems of information architecture mentioned in 2.1.2.4 | Search systems on page 29. Even though search systems are an integral part of information architecture, we are ignoring it as it can potentially reduce the need for looking for the answers for our participants.

3.1.2.5 | Task list

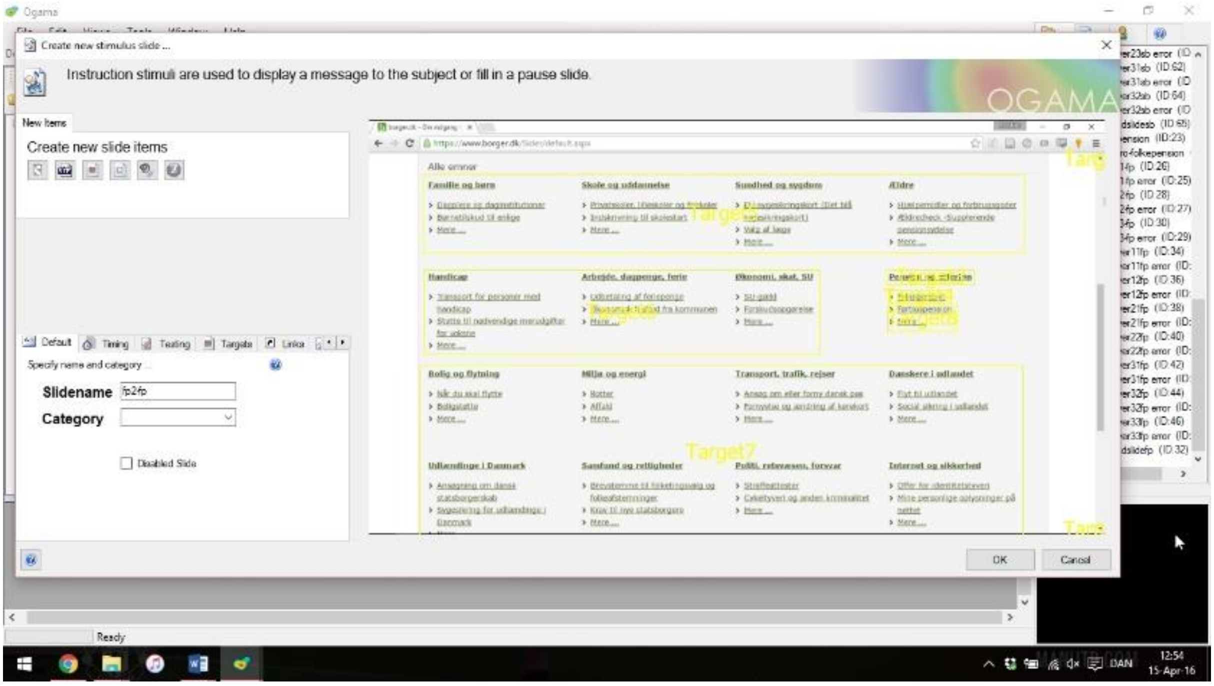
Our study will follow certain guidelines, to ensure the users are tested in similar fashions. Each user will fill out a questionnaire, to provide us with information on their demographic background and prior knowledge. They will be asked to sit down in front of a computer and the test moderator will help the participant calibrate the equipment. The test moderator will explain to the participants what the test is about. When the test begins instruction slides will inform participants how they start the test and when they have completed each task. The goal is to have minimal involvement from the test moderators to minimize the risk of the test results being contaminated but the test moderator shall always be present and nearby to help participants if need be, whether its equipment malfunction or help to proceed with the tasks.

The purpose of this particular OGAMA experiment is to simulate the website borger.dk. To accommodate that we have created an experiment that has the illusion of the site to a certain extent. For every website page the experiment simulates, multiple slides have been added with links between each other. To imitate scrolling the participants can use the directional arrows to move up and down the page. In all slides the browser window is also present in order to support the illusion of using a website. To do this we have taken over 30 screenshots of different areas of the website. These slides and the instruction slides will be in Danish as all our participants will be Danish and we are simulating a Danish site.

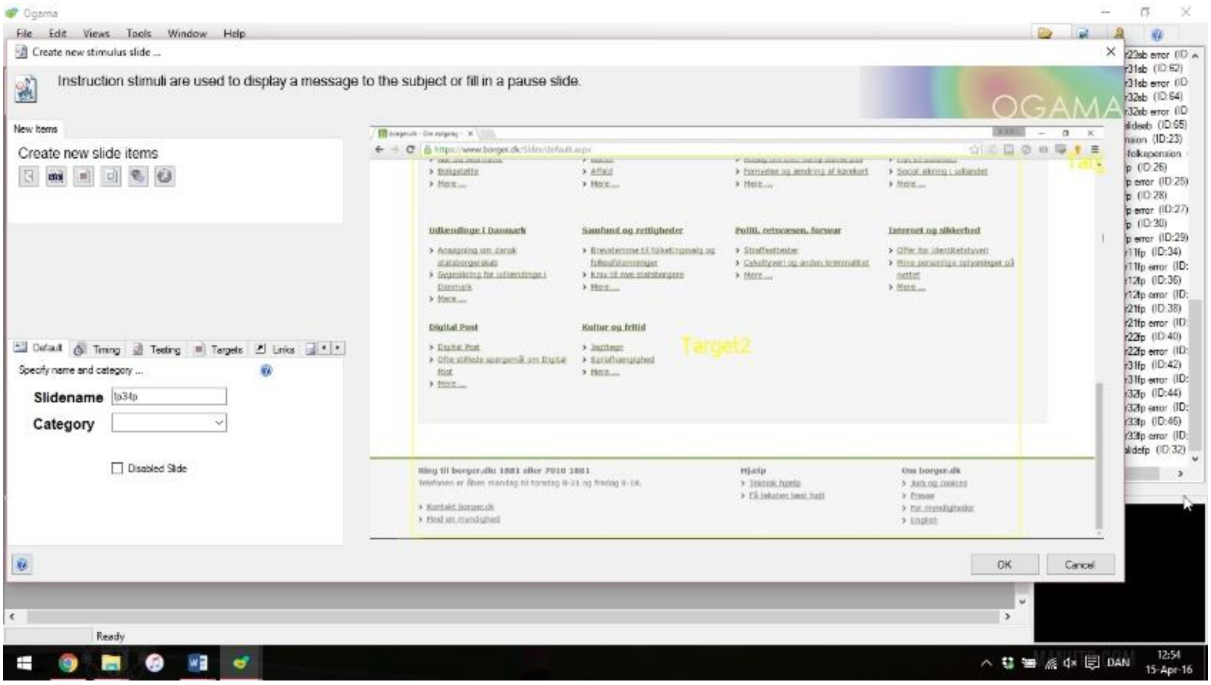
For our test we have created three tasks for our participants to complete. To complete them, they will have to visit different parts of the site. The information they will be looking is located vastly different places visually, as a method of making sure that participants cannot complete the tasks, without looking around on the screen. This is to avoid participants only looking in one location because they were successful twice in doing so. Instructions on how to complete a task is located in the beginning of that task and will explain to the participants what kind of information they a looking for. The tasks will have the same order for all participants and will not be randomized. This is due to a software limitation because OGAMA is as simple as it is. If one were to change the order of slides in OGAMA they would effectively have another experiement on their hands. This makes it very hard to compare results and with the amount of data we are trying to create and analyze it would take a very long time. We fully understand how this affects our results; there is a chance that participants will complete the later tasks faster than the first as they have now understood how the site works. However, given as we are not testing the usability of the site and all tasks are ordered in a way that requires participants too find the information a new place everytime, this is what we have chosen to do.



Picture 1 - Front page 1



Picture 2 – Front page 2

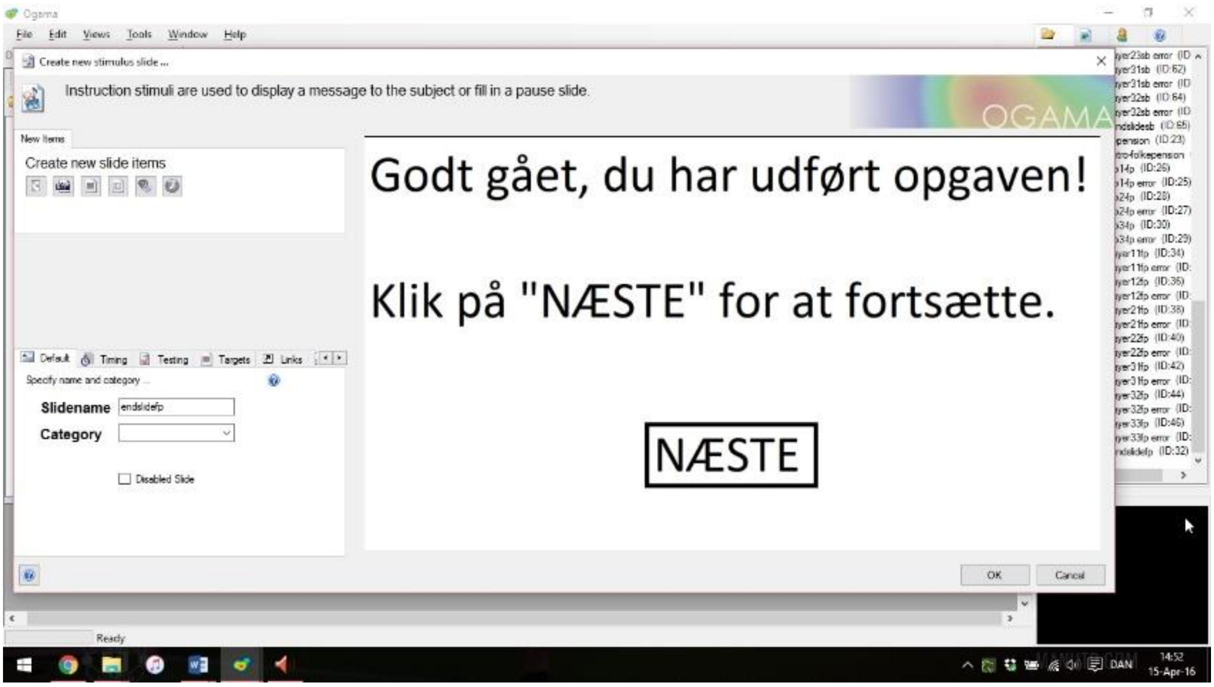


Picture 3 – Front page 3

After reading the task specific instruction slide we start them off at the "front page" of our illusory site. The yellow round targets in the right corners are links between the slides. The rectangle shaped targets fall under two categories, progression targets or error targets we call them. A progression target will lead the participant to the next slide in the experiment and thus allow them to progress towards completing their task. These target rectangles are fit to match the size of the option on the page so participants have to click precisely. The error targets are much bigger as they encompass many more options. These targets will link the participant to an error slide. The purpose of an error slide is to inform the participant that they have selected the wrong option and that the information they are seeking cannot be found there. It also informs them that they can press a "BACK" button to return to the previous slide. This is to help the user accomplish the tasks, as we are not testing their ability to complete the tasks, but how they attempt to do it.

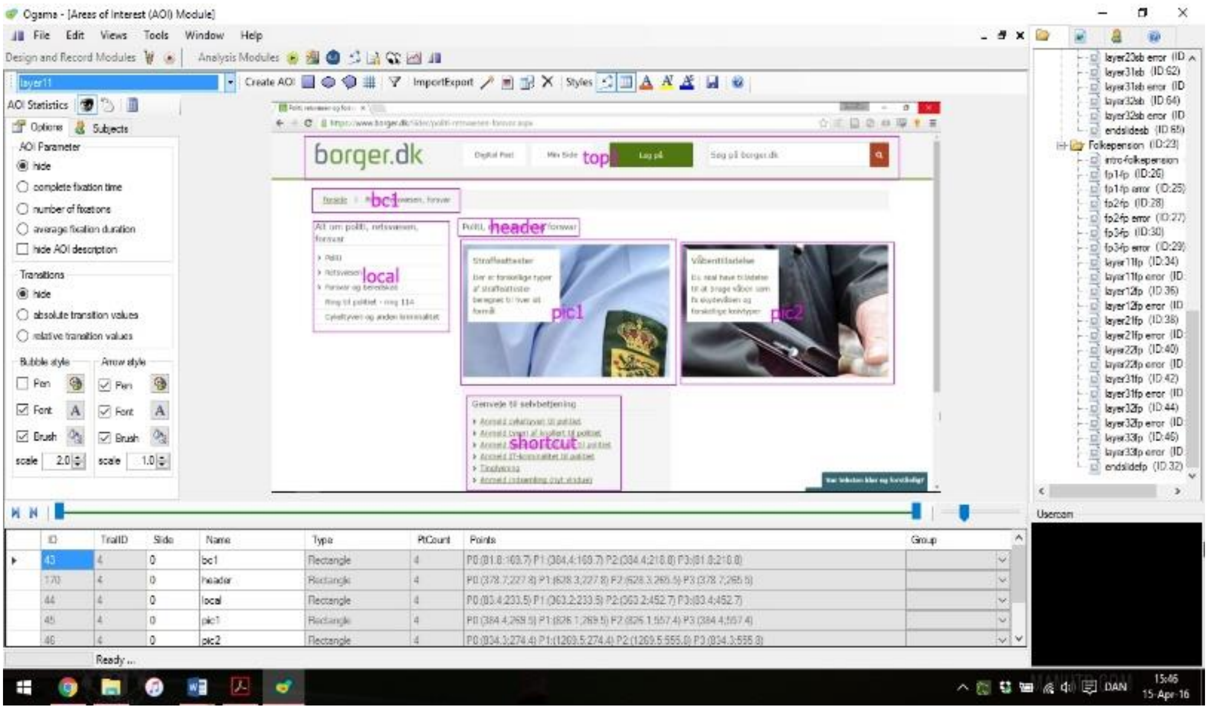
Picture 4 – Error slide

By pressing the target, they get send back to the page they were previously on. As OGAMA is a fairly simple piece of software it lacks a 'previous slide' function. For this reason, we have included a companion error slide for every other slide in the experiment. The purpose of all these slides is that they link back to their companion slide so participants always come back to the slide they were previously on.



Picture 5 – End slide

When a participant has completed a task they will be presented with an end slide which will inform them that they have completed that task and give them a button to press to go to the next task. The last slide in the test is similar to this though it thanks participants for participating and gives them a button to press to end the test.



Picture 6 – AOI

In our experiment we have ensured that it is possible to collect all of the data that we want to as seen in section 3.1.2.10 | Measures on page 51. Our way of ensuring that is to add Areas of Interest(AOI) to all slides. The term AOI were presented in chapter 2.2 | Eye tracking on page 30. These smal areas have been added as squares on the page which each cover things on the page that could be of interest.

The screenshot shown above is a representation of these areas, covering important objects such as the local navigation, pictures and shortcuts. For the picture shown above, the AOIs have been selected on the base of what participants would probably look at to find information such as the local navigation or the shortcuts. We have also added AOIs to the header and the breadcrumbs as it is a possibility that participants might want to reconfirm their current position in the site structure.

Pre-test arrangements

- Prepare software(OGAMA)
- Statements of consent are filled out
- Questionnaires are filled out immediately prior to the test by each participant

Before conducting the actual test, it is advisable to have certain things prepared. According to Pernice and Nielsen it is best to have your eye tracking software ready to go immediately as soon as the participant arrives. Furthermore, they argue that by giving participants too much information before the test they might get nervous or confused. (Pernice and Nielsen 2016) Having the software ready when participants arrive and making sure that they do not see the whole setup makes for a more professional presentation and also correlates with the perspective of making sure participants are more comfortable in an unfamiliar situation.

Session introduction

- Introduce the users to the test environment
- Present the test, its purpose, the equipment, calibration and tasks to the participant
- Calibrate the equipment: The seat, screen and The Eye Tribe equipment

Pernice and Nielsen argues that participants should know just enough about the eye tracking test before starting but not too much. By overdoing it you run the risk of making participants uncomfortable. (Pernice and Nielsen 2016) By giving users the right amount of information you minimize the risk of giving them information overload which can lead to confusion or the participant not focusing on the tasks.

3.1.2.6 | Test environment

Our tests will be conducted as laboratory tests, as described in section 3.1.3.2 | Test setting and testing on page 55. This is done because it is necessary to ensure that the computer used has OGAMA and The Eye Tribe software installed, has the required processing capability to use the software and hardware. The use of a laptop is to give us a mobile laboratory needed, as we have to travel to and conduct the tests at varying locations. At the different locations we have to find places to set up our equipment in a way, so the possible distractions to the participant is as low as possible and with as little direct sunlight as possible. It also has to be a controlled environment with as little distracting elements, such as people walking by, as possible. We will attempt to position the participants so they face a wall so they will not have pedestrians walking directly in to their view. In regards to our position as test moderators, Pernice and Nielsen recommends having the test moderator sitting next to and slightly behind the test participant, to ensure close enough proximity to aid with any problems, but just out of their field of vision, so as to not be a distraction (Pernice and Nielsen 2016). We intend to follow these recommendations.

Picture 7 - Test setup

The equipment used are laptops with OGAMA installed on them and The Eye Tribe tracker, which we have described in more detail in 3.1.3.1 | Hardware and software on page 54. The tracker has to be placed underneath the screen, which means we have to place it on the keyboard of the laptop. This is of course not the best solution, as it can be distracting. We work around this problem, by given the participants time while sitting in front of the test equipment, while being introduced to it, before they have to perform the tasks. This helps reduce how distracting it is. The tasks do not require the use of a keyboard, but only a mouse. We have chosen to use a wireless mouse and not the touchpad. This is done as most people know how to use a regular mouse and as touchpads have varying sensitivity, which can confuse the user, as it might be very different from their own laptop. To improve the visibility of the screen, we are raising the back of the laptop, so the screen is higher.

This is done so the eye tracker, positioned on the keyboard, does not obscure the view of the bottom of the screen.

The picture above shows how we intend to setup our test equipment. As described above, we have raised the screen, positioned the eye tracker on the keyboard and provided a wireless mouse.

3.1.2.7 | Location and setup

The conduction of this test will take place in very different locations. Due to limitations, we have to come to our participants and test where they are available. That will cause some difficulties as we relinquish some control of the test environment. However, we are planning to conduct the test in as similar fashion we can regardless of location i.e. having consistency in the setup of the computers. These specifics can be found in section 3.1.2.6 | Test environment.

Our equipment for conducting the test includes two laptops, two eye trackers and two wireless mice. Both laptops share the same 13.3 inch display size, the eye trackers are identical and the mice share the same size and are both without extra buttons. All questionnaires and Statement of Consents will be in print form and we include pens for the participants.

3.1.2.8 | Project schedule

The deadline for the project is May 31 which forces us to have certain deadlines that we must uphold. We plan to have our testing finished at least three weeks before deadline so we have an appropriate timeframe to sort through and analyze the data.

3.1.2.9 | Moderator role

The moderator's main role is to greet participants, help them to their seat and help them get setup with the equipment. Afterwards the moderator will explain what the test is about and how it is going to be conducted through a manuscript. Trivedi and Khanum argues that social context matters when doing testing and that having social interaction with other people while testing, can have a positive impact on results, rather than a test with minimal social interaction (Trivedi and Khanum 2012). Because this is not a usability test the role of the test moderator will be that of a friendly guide, always on hand should participants need it but never intruding if the participants are doing fine on their own. This social interaction will also be part of the presentation of the task as the moderator will explain participants what the test will be about before beginning to enhance the social interaction. Because each task will be presented on a slide during the test every participant gets the exact same instructions. Removing the moderator from this process and replacing them with written instructions, maximizes the level of consistency in the task presentations.

3.1.2.10 | Measures

Before each eye tracking test, we will have the test participants fill out a questionnaire. This questionnaire will include questions regarding their gender, age, prior knowledge, demographic information and previous experience with borger.dk. More information on the questionnaire can be found in 3.1.4 | Questionnaire on page 58.

We will get the following measures from our eye tracking tests:

- Attention maps
- Fixations
- Errors
- Clicks to complete
- Time to complete
- Scan paths
- Click paths
- Areas of Interest

These will be used to measure, if any, tendencies for each demographic group and if a level of prior knowledge exists.

The measurements are used to answer the test plan research questions. Below we describe what measurements are used for the individual test plan research questions.

'Do young and elderly people focus on different parts of a website?'

This is answered by the attention maps, which shows what has been focused on. Fixations and fixation times in AOI will elaborate on that.

'Do they navigate through a website structure differently?'

Scan paths show how the participants look through the structure and click paths will show how the navigate through it.

'Are there differences in their time spend on completing tasks?'

This will be investigated by the participants' time to completion for each task and can be elaborated on by the time used in each AOI, to understand where the potential differences occur.

'Do they spend varying time looking at specific parts of the website?'

This is closely related to the question above and can be answered by looking at the time spent in the different AOIs and time spent on each slide.

'Are one group more prone to using a trial-and-error method than the other?'

By looking at the amount of errors by the different groups, we can get a broad understanding of whether any of them are using a trial-and-error method. To supplement it we can give a heuristic evaluation by looking through the test as it is recorded in real time.

'Are there any tendencies in where each group starts on the website?'

This can be done by looking at the first couple of fixations. The first fixation will most likely be at a random spot, as the task starts with a black screen. This means we will look at the second fixation to see, where the participants start their information search.

Measurements will be conducted on all the tasks completed by the participants. Every tasks ensure that the participants will visit different parts on the site. The tasks have been created with the goal in mind that participants can use different methods of completing the task through the main navigation, shortcuts etc. With these things in place, the results will show where participants look at the site.

3.1.2.11 | Analysis

Our analysis will comprise of the above-mentioned data to identify differences between the groups, tendencies in certain groups, similarities in the groups, and relate the findings to existing research. This can be found in 4 | Analysis on page 76.

Now that the test plan has been accounted for, the following sections will go further into detail on our data collection in relation to eye tracking, our questionnaire and ethics.

3.1.3 | Eye tracking

We are collecting most of our data by eye tracking, as described in 2.2 | Eye tracking on page 30. The specific measurements used to analyze and visualize the collected data we will be described in more details in 3.2.1 | Eye tracking on page 65, where we discuss how we analyze the collected data.

The specifics of our eye tracking study can be seen late in this chapter. In 3.1.3.1 | Hardware and software on page 54 we describe the tools we use to collect our eye tracking data in regards to the software and hardware. In 3.1.3.2 | Test setting on page 55 we discuss how we intend to conduct the experiments and how we set up our test equipment. Finally, in 3.1.3.3 | Pilot test on page 57 we describe how we conduct the test in accordance with our chosen hardware and software and test setting, to get feedback and make necessary changes and improvements to our test method.

We acknowledge that the hardware and software we use, might give results that can be different if we used another combination of hardware and software, but it is a necessary evil to collect data. It will not, however, effect how the data can be used to provide indications for any differences between groups as we use the same software and hardware for all the tests. Using OGAMA and The Eye Tribe eye tracker gives us the ability to collect the necessary data to compare the way different people look at a website, where and how they look for information, which will in turn help us answer our problem statement and see if there are any differences in how elderly and young people look for information.

3.1.3.1 | Hardware and software

OGAMA stands for Open Gaze and Mouse Analyzer and the program was written solely by Dr. Adrian Voßkühler of the Free University of Berlin("OGAMA" 2016). It supports several commercial eye tracking units including The Eye Tribe which will be described below. With OGAMA you can add slides to an experiment and have users go through those slides one by one in a specific order. In these slides you can add target areas for the participants to click. It allows for a test of finding a specific thing and clicking on it to progress. The software, as the name indicates, excels in analyzing mouse movements and where people look. By employing AOI's on selected slides the software can create data on fixation times and transitions.

OGAMA is a very basic program to work with and construct tests in. Everything needs to be done manually; creating slides, adding target areas, creating completion conditions, creating links and so on. For larger experiments with more than 15 slides there is no way to edit them all at once, each must be manually edited by the creator. This is a significant limitation of the program as it is rather time consuming to work with. If one were, for instance, to merge data from two or more experiments, it would be a hefty task depending on the size of those experiments off course, as everything from the different experiments must be manually paired. This limits how the software can be used, as it more or less removes the option the conduct simultaneous tests on more than one computer. As OGAMA is freeware and a community driven software, there is a risk that it might contain a high number of bugs. We have experienced a few crashes when working with it previously. To reduce the risk of corrupting the data base, we intend to make backup copies of it, for each experiment session. This means we can always go back to the latest working version and avoid having to start over.

We elected to use OGAMA in this project for several reasons; it is freeware with a large community, it works well with our tracker form The Eye Tribe and we have had some training in using it, as part of our Master's program. It provides us with the means to gather the data, described in chapter 2.2 | Eye tracking.

The Eye Tribe is a Danish company based in Copenhagen who sells eye tracking hardware. Their equipment consists of a small tracker with a stand and a connection cord. The Eye Tribe's technology *"relies on infrared illumination and uses advanced mathematical models determine the point of gaze."*("The Eye Tribe" 2016). This means that the eye tracker emits infrared rays, which are reflected by the test participants eyes back towards the tracker, so it can calculate where the participant was looking. This is as described earlier in 2.2 | Eye tracking on page 30. The Eye Tribe eye tracker is a fairly cheap piece of hardware at only $199, which fits the vision of The Eye Tribe of providing eye tracking hardware for the mass market consumer ("The Eye Tribe" 2016).

There are a number reasons why we chose to use the eye tracker from The Eye Tribe. It is very mobile on account of its size and the ease of operating it. It provides us with the ability to go out and conduct the tests where we can find the test participants. It is a non-intrusive piece of hardware, as it does not require the participants to wear any equipment, as the tracker are positioned in front of them, without any need to physically interact with it. It is also available for use from the university, which makes it easier for us to acquire.

3.1.3.2 | Test setting and testing

We are focusing on how the participants complete the set tasks and using the collected data to investigate whether differences exist between the different groups. This is done in a fashion where the test introduction and task presentations are all presented in the exact same way, to ensure as high a level of consistency as possible.

The tasks will be designed to have the participants navigate through different parts of the structure and in areas they might be familiar with and areas with information not usually relevant to them.

The test will be conducted as a laboratory test, which comes with both strengths and weaknesses compared to the opposite in field tests.

The controlled setting of a laboratory test adds to its validity, as outside influences are reduced (Hertzum 1999; Trivedi and Khanum 2012) as it provides a peaceful space (Kaikkonen et al. 2005). The controlled setting ensures the participants experience the test procedures and tasks the same (Trivedi and Khanum 2012). It also allows for the researchers to focus on specific subjects they wish to investigate (Trivedi and Khanum 2012).

Laboratory tests, however, puts the user in an unnatural situation, conducting tasks they might find equally unnatural (Hertzum 1999; Trivedi and Khanum 2012). Trivedi and Khanum describes the difference between field and laboratory settings as field settings sacrifices control for realism, while laboratory settings sacrifices the realism for more control (Trivedi and Khanum 2012). It can also be seen as a problem when the environment is too 'clean' and without anything to interrupt the user, as this is not how they would normally be using a website (Kaikkonen et al. 2005), which would make the results be descriptive of an unnatural situation that might not exist outside the laboratory.

Laboratory tests focusses on how something is done and not what can be done in the system (Hertzum 1999). This fits the goal of our study, as we intend to investigate how the groups look at a website and not what the website can do. Furthermore, field tests are more time-consuming (Trivedi and Khanum 2012; Kaikkonen et al. 2005), which does not fit into our schedule.

According to research done by Kaikkonen et al. the number of problems found are similar in laboratory and field tests and the severity of the problems did not differ (Kaikkonen et al. 2005). This adds validity to our chosen approach.

3.1.3.3 | Pilot test

To evaluate and validate our questionnaire and eye tracking test, we will conduct a couple of pilot tests. This is done to ensure that the actual study can be conducted with as few problems as possible. The results will be described in 3.3.2 | Test reflection. This is done in accordance with the results of the study 'The importance of pilot studies' by Van Teijlingen and Hundley, which states that by doing so increases the validity and ensures the best use of research results (Van Teijlingen and Hundley 2002). This is, according to their paper, something that is rarely done, as some see it as devaluing their work and some publishers sees the description of problems encountered as a weakness (Van Teijlingen and Hundley 2002). We include it to follow the recommendations and to explain the possible changes to our tests.

Pilot tests can be used to ensure that everything needed is in order. This includes ensuring all materials are printed and that the test product are ready, checking whether the tasks are understandable and helps getting an idea of how long time each test takes (Schade 2015). It helps the test moderator getting a feel of problem areas and how to use the equipment properly, in a low pressure environment (Van Teijlingen and Hundley 2002). For questionnaires it can help with spotting errors, the range of multiple choice questions and choosing the best possible order of questions (Van Teijlingen and Hundley 2002). Furthermore, it is useful in an organization to convince the managers that the study is worthwhile and thereby securing funding (Van Teijlingen and Hundley 2002). It can also help in spotting problems with the chosen test method before the actual test, so a more appropriate method can be chosen (Schade 2015).

There are, however, some reservations regarding the use of pilot tests. As pilot tests are mostly done on small groups, it does not carry a lot of statistical certainty and are thus not a guarantee for the success of the study (Van Teijlingen and Hundley 2002). Including the results of the pilot test can have a negative influence on the test results as well, as they might be conducted in one way and then the actual study are slightly different, to avoid problems encountered in the pilot test (Van Teijlingen and Hundley 2002). It is important to note that qualitative studies are often done in a dynamic fashion, where each subsequent test is improved on the basis of the previous test. This means that the problem stated above is mostly relevant for quantitative studies (Van

Teijlingen and Hundley 2002). Amy Schade from the Nielsen Norman Group are more open to the inclusion of the results of pilot tests, but does state that it should not be included at all costs (Schade 2015). In some organizations the costs of the pilot study might result in the actual study being forced to continue, even though the pilot test showed it had flaws, as it will mean taking a loss on the study. It can also have the opposite result, where the funding body stops the main study, as they feel the pilot test has giving adequate results (Van Teijlingen and Hundley 2002).

We will follow the suggested guidelines for conducting a pilot test set forward in the article by Amy Schade (Schade 2015) and the paper by Van Teijlingen and Hundley (Van Teijlingen and Hundley 2002) described above.

The questionnaire will be handed out at the same time during the test, as we will be doing it in the main study (Van Teijlingen and Hundley 2002). The questionnaire will be handed out before each eye tracking test, but after the statement of consent. We will ask for feedback on the questions and how easy to understand they are and make changes as needed.

The pilot tests will be timed, to get a more precise idea of how long each will take (Schade 2015; Van Teijlingen and Hundley 2002).

To ensure the usefulness of the pilot tests, they will be conducted on two participants, one from each age group, to attain information on both groups responses (Schade 2015; Van Teijlingen and Hundley 2002).

The pilot tests will be conducted at least one day prior to the first tests in the main study are planned to allow time for necessary changes (Schade 2015).

The pilot test results will not be included in the main study, to avoid contamination of the research data (Schade 2015; Van Teijlingen and Hundley 2002).

3.1.4 | Questionnaire

The complete questionnaire can be found in Appendix 1.

Our questionnaire has been designed so we can collect the data we are interested in, which are demographics, described in 2.1.1.2 | Demographics on page 13 and prior knowledge described in 2.1.1.3 | Prior Knowledge on page 14. This information will help analyze and compare the data collected through eye tracking, between different groupings of participants. We can compare the participants on their demographics and their level of prior knowledge, to test whether there are any differences between the young people and the elderly.

As will described later in

3.2.2 | Questionnaire on page 67, we intend to calculate the median of the participants' prior knowledge, so we can compare them based on their level. The level of prior knowledge is based on a range between one and six, with one being the lowest. This means that the questions relating to prior knowledge, has to be designed, so we get values between one and six. This has been done for most questions, however we do have a few yes/no questions asking the participants' if they have ever done certain tasks online. These questions have been bundled up into a group of six questions. When analyzing the questionnaires, we can count the number of yes' and no's and have a value between one and six that can be used on calculating the levels of prior knowledge.

Our questionnaire will be used to provide extra data, which cannot be collected through eye tracking; demographical information and their prior knowledge. As our thesis deal with the concept of age and whether or not it makes a difference we have to uncover what age our participants are. This has been done in many studies in the past in relation to using the internet. Zukowski and Brown found that age was an important factor in how users felt about privacy with e-shopping and they also argue that young people are faster to embrace the internet (Zukowski and Brown 2007). In 2001 Thompson discovered that younger people were more likely to be proficient at downloading and messaging on the internet, compared to older respondents (Thompson S.H. Teo 2001). Kalmus, Realo and Siibak found that younger people were more likely to use social media and entertainment on the internet than older users and that younger users were more driven by free will to do so. (Kalmus, Realo, and Siibak 2011). Dwivedi and Williams was able to confirm their hypothesis, that older users were less likely to adopt e-government in the UK, to be true (Dwivedi and Williams

2008). All of these studies show that age makes a difference in the way that people interact with technology. This means that it is crucial for us to get that information so we can use it in our analysis.

In our questionnaire we will also ask participants to state their gender. For our part this is another way of differentiating between individuals. However, there are several studies that have specifically investigated genders importance. Zukowski and Brown found that gender had little to no influence on how users felt about privacy when shopping online (Zukowski and Brown 2007). Thompson discovered that males more so than females use the internet for downloading and purchasing activities, while females tend to engage in message activities more than males do (Thompson S.H. Teo 2001). Dwivedi and Williams found the difference between males and females to be insignificant when relating to e-government adoption (Dwivedi and Williams 2008). We will include gender in the questionnaire as it allows us to see if there is a difference between genders across the two age groups.

In our questionnaire we ask participants to state their educational level as it might have an effect on results. The following studies have investigated this particular parameter. Zukowski and Brown was able to confirm their hypothesis; that people with higher levels of education were less concerned about information privacy than people with lower levels of education (Zukowski and Brown 2007). In his testing Thompson found that levels of education did not scale positively with the usage of computers and argued that the accessibility of the technology is the reason (Thompson S.H. Teo 2001). Kalmus et al. discovered that higher levels of education scaled positively with internet usage related to work and information seeking (Kalmus, Realo, and Siibak 2011). Dwivedi and Williams was able to confirm their hypothesis on education, that higher levels of education scaled positively with e-government adoption rate (Dwivedi and Williams 2008). These studies have shown how education makes a difference when using the internet, how it is used and what reservations the users might have. For this reason, we include it in our questionnaire.

Much like demographics we are also interested in investigating participants' prior knowledge. Using the theory from 2.1.1.3 | Prior Knowledge on page 14 we can develop the questions needed in our questionnaire to measure that parameter.

As prior knowledge is a combination of tacit and explicit knowledge, it can be difficult to express and therefore difficult to measure. Measuring of prior knowledge are done in various ways, but not all are equally useful and some gives varying results (Dochy, Segers, and Buehl 1999).

Dochy, Segers and Buehl argue that having the participants assess their own level of prior knowledge or having the experimenter judge the participants level, is a flawed method, as it is based on subjective thoughts. The same goes for free recall tests and interviews, as their outcome are affected by the participants verbal abilities (Dochy, Segers, and Buehl 1999).

Other methods are viewed as being more reliable. These methods includes multiple choice test, open questions tests, close tests, completion tests, recognition tests and matching tests (Dochy, Segers, and Buehl 1999). These tests are what we will base our questionnaire questions on to measure prior knowledge. Specifically, we will use multiple choice tests, as it has been used in related research, which we have discussed earlier. Beier and Ackerman used a questionnaire, where the users was asked multiple choice questions to measure different values, such as demographics, cognitive abilities and prior knowledge. This was done by setting up questions for the participant to rate their level of agreement on a 6-point scale, where 1 was strongly disagree and 6 was strongly agree (Beier and Ackerman 2005). Crabb and Hanson did their measurement of prior knowledge in a similar fashion. They used a 5-point scale for measuring the participants confidence in using the Internet, which also ranged from strongly disagree to strongly agree, and a 6-point scale to measure their Internet usage, which ranged from never to everyday (Crabb and Hanson 2014). These methods and ranges from the inspiration for our structuring of questions in the questionnaire. We do, however, intend to include a few questions, where we ask the users to assess themselves. This is necessary, as we wish to know how much the participants have used computers in their work and studies. This is hard to do if they have six different amounts to choose from, as it most likely have changed during the elderly participants' careers. Instead we ask them how much they themselves feel they have used computers and then have a critical approach to the collected data.

In addition to wanting to know participants experience with computers in their careers we also intend to investigate their regular internet usage. To measure this, we will use a range closely related to Crabb and Hanson's: 1.Everyday 2.Several times a week 3.Several times a month 4.Every few months 5.Less often 6.Never (Crabb and Hanson 2014). The labels can be changed to fit different types of questions, that are not related to time, as the ones above. This allows us to calculate a median of the participants Internet usage, which can be used to give an indication of their level of prior knowledge, on scale from one to six. What constitutes a high or low level of prior knowledge will be based on the median of all participants. With that value, it is possible to place each participant as either above, at or below the median.

In order for us to get all this information on participants' prior knowledge we propose to ask the following questions.

The first grouping is related to their usage of different functions on the internet. Participants choices will range between rarer than twice a month and more often than 4 times a week. The participant will be asked how often they do any of the following online activities;

- Use online banking
- Contact the commune/state
- Use a search engine
- Read news
- Watch films, series or documentary
- Digital post
- Social media
- Email
- E-shopping
- Take part in online discussions

The next grouping is related to participants daily and weekly usage of the internet in general.

- Internet usage – Times a day; ranging between rarer than 1-2 times a day to more often than 8-10 times a day

- Internet usage – Hours a week; ranging between less than 3-9 hours to more than 30-39 hours

The following grouping is related to participants' experience with computers in either previous jobs or studies. They will be asked to grade the following questions on a scale of 1-6, in which 1 is a little and 6 is a lot.

- Use computers in previous jobs
- Use computers in studies

This is to understand how much the participants have used computers and get an estimation of how they perceive their own prior knowledge.

Finally, we intend to investigate participants experience with borger.dk. We include these questions for several reasons; if certain participants are vastly different to the rest it could be because they have either used the site many times before or not at all. Having this information allows us to get an understanding of the users' prior knowledge of the specific site. however, we do not intend to ask users specifically about the site borger.dk; instead we will ask them if they have ever done any of the following;

- Used the internet to change their home address
- Used the internet to renew their passport
- Used the internet to renew or change their driving license
- Used the internet to order a new medical card
- Used the internet to choose a new doctor
- Used the internet to apply for public financial support

We ask these questions to investigate if people have used borger.dk before without asking them specifically. We do this because we want to make sure that participants do not remember incorrectly. All of the questions we ask are purposefully related to actions citizens must do on websites that represents the Danish state. In addition, we also intend to ask participants if they have visited the following sites;

- Borger.dk
- SU.dk

- Sundhed.dk
- Skat.dk
- E-boks.dk

We ask all of these questions to investigate how much experience users have with borger.dk and other public Danish sites. We can use this data to explain possible anomalies and to provide more detail to our participants as a whole.

3.1.5 | Ethics

Ethics are an important part of doing data collection as it is the collectors responsibility to act in an ethical manner. Geoffrey Mills wrote that in action research there are ethical guidelines data collectors can follow (Mills 2010). Informed consent is important to make sure that participants understand what they are getting involved with and to reduce the likelihood that they will be exploited by the researcher. Freedom from harm is the act of making sure that participants will not harmed in any way, be it physically or mentally. It also includes handling their individual data with care so it will not be exposed to others who could abuse it. Mills argues that anonymity is a good method of making sure that the data has no personal ties to avoid problems like the one mentioned above (Mills 2010). If said anonymity is not an option due to the projects nature, confidentiality is a method of ensuring that participants can still be protected from unwanted attention or ridicule. If the researcher promises that personal information and test results from participants will only be seen by said researcher an amount of trust is required from both parties (Mills 2010).

In order to maintain an ethical standard for this project we have decided to create a written Statement of Consent that all participants must sign if they want to participate. This document informs the would be participants of our intentions with the test and the results, so we can ensure that they understand who we are, where we come from and why we are doing the tests. Furthermore, we explain these details to them before the test in our face to face introductions.
In the statement, it also says what the equipment we use to conduct the test and how it is not harmful for the participant. We also explain to the participants face to face that the infrared lights on the equipment is not harmful, to prevent participants from

becoming nervous. We do this face to face so we can show it directly on the equipment while the participant is sitting in front of it.
The Statement of Consent informs participants what their role in the test is; that they are anonymous and the data we collect will be assigned a number rather than their name. We do not require any personal information about participants other than gender and age. The statement also says how participants' data will be handled, that we will categorize them in two groups. We explain this again in the test introduction face to face as well.

Almost everything that our participants read in the statement of consent is repeated by us in the face to face introduction before each test. We do this to ensure that participants have understood all of these things, in case the wording in the document was not clear enough or they simply did not remember some of them. When we repeat it we minimize the risk of having participants who do not understand what our intentions and purpose are and what the test is.

3.2 | Data Analysis

Having collected the data, we need ways to analyze it. To do so we use a wide selection of measures, based on the eye tracking data and the data collected in our questionnaires. In the following we will describe what these measures are and how we intend to use them.

3.2.1 | Eye tracking

We have briefly discussed some eye tracking measurements, which can be used to locate usability problems and make conclusions on what the users' experience during their search for information. This was done in chapter 2.2 | Eye tracking on page 30. In the following we will describe which measures we intend to use, to answer our problem statement and research questions.

The data collected by The Eye Tribe and OGAMA, can be displayed in OGAMA in a variety of ways. We intend to use most of these.

Attention maps will be used to show what areas of the website the different groups focus the most on. It can show if either of the groups focus most on one area or scans the different parts of the website.

The fixations can indicate which areas of the website draws the most attention. As fixations are used to develop the attention maps, they will be used with attention maps to analyze what the groups looks at.

Scan paths show us what the participants looks at and in what order. This will help us in finding tendencies in the different groups and comparing to see if any differences exists. We will specifically use the Levenshtein algorithm in OGAMA, to compare the similarities between the different groups.

Our AOI's shows how much time the participants spend looking at specific parts of the website and their transition between each AOI. This measurement is a supplement to the attention map, but a tailored supplement, as it is made to fit our needs. The transition values between AOI's supplements the scan paths, as it shows us the routes the participants eyes take to locate information. AOI's are created by us manually, prior to the testing taking place. This is described in more detail in 3.1.2.4 | Tasks on page 40, where we discuss which areas have been marked.

We are also going to look at the time-to-complete for each task and the three tasks combined. This shows us if one group is faster than the other and can give indications to whether one group is faster at learning to use the website, by comparing the differences between each group for each task and see whether the differences increase or decrease for each task.

A few measurements that is not included in OGAMA, but can be done manually by us will also be used in the analysis. We look at and compare the click paths of the users, to see if there are any differences in how the participants click through the website. To build on click paths we are also going to look at a combination of errors and clicks-to-complete. Errors can indicate if one group experience more difficulties navigating and understanding the website. However, it can also mean that the participant just clicks on the first link that seem related to the subject and uses a trial and error method of finding information.

All of these will be used to measure, if any, tendencies for each demographic group and level of prior knowledge. It will furthermore be used to compare the two age groups.

Measurements will be collected on all the tasks completed by the participants. Every tasks ensures that the participants will visit different areas of Interest on the site. The tasks have been created with the goal in mind that participants can use different methods of reaching their goal through the main navigation, shortcuts etc. With these things in place, the results will show where participants look at the site.

To analyze and interpret the data collected and presented in OGAMA, we will use the lists provided by Ehmke and Wilson in their 2007 paper 'Identifying Web Usability Problems from Eye-Tracking Data' (Ehmke and Wilson 2007), who have paired usability problems to certain eye tracking measures. This will help us reduce the amount of times, where we have to guess at why certain tendencies appear, as described in 2.2 | Eye tracking on page 30.

All of these measurements will be analyzed alongside the information collected through our questionnaire described in the next chapter.

3.2.2 | Questionnaire

To provide information we can use to divide the participants into groups, we use our questionnaire.

The demographic information can be used to divide the participants into groups based on their age, sex and educational background. The rest of the questionnaire focusses on their prior knowledge. The prior knowledge will be measured by finding the median of the values from one to six, the participants use to describe their use of the Internet, to find what the level of prior knowledge is for the participants. This is done to compare different groups of participants, with differentiating levels of prior knowledge to each other, but also to see the differences within each group. The level of prior knowledge might be why differences between groups exists, so it is important to be able to compare this as well.

In the following and final chapter of the methodology we account for our reflections in doing the tests and using the method.

3.3 | Method Reflection

In the following we provide our reflections on our chosen methods. We will describe how it worked as intended and what we could have done differently or better, when looking at it with the clarity of hindsight.

3.3.1 | Pilot tests

Our pilot test showed us a few issues with our questionnaire and eye tracking test, which we have used to make changes to both.

In the questionnaire we found that we had made an error, which confused the participants, as it asked them to circle the correct answer, but the reply method was intended to put a cross in checkboxes. We made changes to improve the clarity in multiple choice questions, where it could be difficult to see if they put their cross in the correct row. This was done by adding lines between each question that went on through to the end of the multiple choices.

In regards to eye tracking, we found numerous issues with the equipment setup. We had thought of positioning the eye tracker, so it was just above the screen. However, this was impossible, as it is designed to be positioned below the screen. This presented a new problem because, as we are using laptops, the eye tracking hardware had to be positioned on the keyboard. This resulted in us having to place a piece of cardboard on the keyboard to stop the legs of the hardware from pressing the buttons. It was also obscuring some of the screen. We solved this by positioning the box the eye tracking software came in, underneath the computer which raised the screen and allowed the participants to see the screen properly.

The lighting proved to be a problem in one of the pilot tests, as the sun was setting, it shone in on the participant and made the tracking problematic. We solved this by ensuring that no direct sunlight was present.

We had apprehensions regarding the use of glasses, as we were worried that they could affect the eye tracking hardware's ability to track the participants' eyes, but we found that the one person using glasses, was calibrated and tracked perfectly.

We were confirmed in our belief that our task descriptions were clear and understandable. We found that it was incredibly important to clearly describe the limitations in the software, as the test participants representing the elderly group, was confused by the inability to use the mouse wheel to scroll. We solved this problem by taking a little more time to explain and show this limitation.

Having concluded the pilot tests we can move on to the actual tests that make up the study. Before presenting the data and analyzing it, we will discuss strengths and weaknesses of our test method, to comply with the recommendation presented in 3.1.3.3 | Pilot test on page 57.

3.3.2 | Test reflection

Having conducted our tests, we have had an opportunity to evaluate on our experiences and what went well, but also on issues with the testing. In the following we will describe what we experienced during our tests, how the actual tests compared to the pilot tests and what could have been done differently.

Our pilot tests provided us with a lot of valuable information. From the error with the instructions for one question, to the clarity of the multiple choice questions, we were able to make improvements to our data collection.

We found it difficult to recruit participants, even from sources we would have thought to be interested in helping us. We contacted the local branch of Ældresagen, where the phone tender asked us to send an email explaining, what we were asking for and what we were doing. We were told that we would get an answer three weeks later, when they have had a meeting and discussed. Unfortunately, we never got a reply. In the meantime, we contacted more Ældresagen branches, but none was interested in helping us. Instead we turned to Aalborg Centralbibliotek. We got a reply were we were told we could not just recruit participants from the library, but we were welcome at their cafés on Tuesdays, Wednesdays and Thursdays where they help citizens with computers, tablets and so on. Attending three of these resulted in 5 subjects tested. The rest we got from personal relationships, were we particularly had to use our parents, their network and other family. This resulted in us having to recruit participants from large parts of Denmark and travel to get them done. We experienced that being two students without resources or professional connections made it hard to get anywhere. In the end we tested 26 elderly participants and five young.

Additionally, we experienced some elderly participants being hesitant, when asked to participate. A close friend to one of our parents simply did not want to participate and seemed genuinely afraid. We of course had to respect this, but also found that a lot of the fear, in lack of a better word, could be mitigated by our thorough explanation of the test, our intention with it and the equipment. It was especially effective explain that we did not test them personally and that we did not test their ability to solve the tasks. We found that our manuscript was well designed and helped us explain the relevant information to the participants.

Recruiting the young participants were easier, as we were allowed to come set up our equipment at local high schools. Unfortunately, we could only set up at spots, were there would be a lot of distractions, which could influence the results. We were also only allowed to recruit the students during breaks, which meant we could only do a couple a day at a school. We did not go this route, instead we used students at the university which we found through a personal connection.

We believe it could have been easier recruiting participants, if we had started recruiting earlier and had the resources to compensate them for their time and professional connections to make us of. We are, however, happy with the number of participants in the study and how diverse they are. In chapter 3.1.2.3 | Recruiting participants we had a table showing how many participants we would have like to include in the test, in the table below is an account of how many we actually got.

Characteristic	*Desired number of participants*
Participant type	
Old	26
Young	(5)
Total number of participants	31
Prior knowledge	
Low	13(1)
Medium	12(1)
High	1(3)
Internet usage	
Low	7(0)
Medium	12(0)
High	7(5)
Gender	
Male	17(2)
Female	9(3)

Table 2 – Recruited participants

Our questionnaire proved to be well designed, as we did not experience problems with it, after we made the changes based on the pilot tests. The only issues were with the final question, as it is in relation to the use of computers during the participants' studies. Some elderly was unsure, what they should do with that one, as computers were not around, when they were students.

In line with the theory described in 3.1.4 | Questionnaire on page 58, some participants had difficulties evaluating themselves and spend a fair amount of time trying to figure out where on the spectrum, they should position themselves. One participant seemed like she was attempting to answer the questions in the way she thought was the correct way, meaning in a manner that were like everyone else. This seems like a direct consequence of the questions themselves as they all deal with how often you would check your digital mail and the like. This can be seen as an important action that responsible people would do often. We got the impression that she wanted to appear responsible and "correct" which can have influenced her answers.

Despite the limitations of having the participants evaluating themselves, the questionnaires were quickly filled out and we got a lot of information, usable for analyzing the data collected by eye tracking.

The tasks were easy to understand for the majority of the participants and none got completely lost. A few forgot what they were supposed to do, because they clicked through the instruction slide too fast. We had a couple who were not sure when they were done with the tasks, even though we explained that a slide would be displayed, saying the task were completed, when they had finished the task.

Many of the participants used the same click paths which might reflect how some paths were easier or more obvious than others. We investigated the routes they took to see if participants would spend time examining options in depth or choose the first option that looked promising.

The equipment we used proved to have both positive and negative aspects. The first problem was found during the first pilot test and has to do with the placement of the eye tracker. We found that it had to be positioned underneath the screen, which proved that a laptop is not optimal tool to use with the eye tracker from The Eye Tribe. It had to be positioned on the keyboard, which meant it would press the buttons and block parts of the screen. We found a method of setting it up, so it could track the eye movements, without blocking the screen. A better solution would have been to bring a monitor, keyboard and mouse to the tests and have the participant sitting at the second screen, with the keyboard and mouse. We, as the test moderators, could then sit next to them and start the test from the laptop. The eye tracker could be positioned directly under the monitor, as The Eye Tribe recommends and would have been easier to hide to make it feel less intrusive.

A positive aspect of the equipment was that it was both easy the explain the calibration process, easy to do and easy for the participants to understand.

The problem put forward by The Eye Tribe regarding saggy eyelids, did not prove to be a big problem. Most of our participants were between 55 and 65 years, a few older and up to 88 years old. Of these we only experienced problems caused by saggy eyelids in one participant, but that also proved so problematic that we could not do the test.

We found that even though the pilot test, described in 3.1.3.3 | Pilot test on page 57 showed glasses was not a big problem, it actually proved to be, during our tests. Most of our conducted tests that we cannot use, due to bad or lacking data was caused by glasses. We found that the problematic types of glasses were not only glasses with shifting focus, but regular glasses were also causing problems. This was a major problem, as most of the elderly participants were using glasses.

The mobility of The Eye Tribe eye tracker makes up for the limitations of it and we have been able to conduct more than 30 tests, with most of them providing good and usable data.

Head movement also proved problematic, as some elderly participants had a tendency to move in closer to the screen, when trying to read something or when they were confused by the website. It is possible that this is a direct consequence of using a 13.3 inch display; a standard size laptop screen. This issue of participants moving closer to the screen so see something better could perhaps have been resolved by using a bigger screen. However, that would have an effect on the test in an undesirable way. Many of our participants were avid users of laptops and iPads and are thus forced to use small screens in their everyday life. We would argue that using a bigger screen would not give the true results because it is unlikely that participants use one themselves. This is backed up by data from Ældresagen which states that 72% of the age group 65+ primarily uses laptops and that iPads are increasing in popularity("Går På Nettet via Computeren" 2016; Lauterbach 2015).

The software has been passable and is very good as freeware by freeware standards. The problems we experienced were varied and, at times, very strange and unexpected. We had two occurrences of the software crashing. One happened during a calibration and only required us to restart the program. The second one was a bigger problem, as it happened during a test. Luckily it was a test where the calibration was not very good and might not have been usable anyway even if OGAMA had not crashed.

For one test participant we experienced a very strange error. When he completed a task he clicked the button to proceed to the next task, but was sent to an error slide and from that he was sent back to the previous task. We have not been able to reproduce this and there are no links from the slide to that error slide, so we do not understand what happened. The test participant took it lightly and continued without any problems but the results have obviously been altered by this error, namely the time to complete as the participant had to go through the previous task again. This will, of course, be considered in our analysis.

The fact that the software uses slides disrupts the feeling of being on a website, as you do not get any mouse-over feedback, when having the mouse moving over a link and not being able to search or use the mouse wheel to scroll.

We found it to be a hindrance that we as test moderators were not able to follow the tracking of the participant's eyes in real time, as we could not ensure that the participants eyes were tracked. This could have improved the data collection, if we could tell the participants to sit still for a second or two, to allow the tracker to reacquire the tracking of the participants' eyes, when it had lost it. However, this would have required a dual monitor setup which is not supported by OGAMA, but we find it to be a significant feature that could immensely improve the software.

We had several experiences, both good and bad, with our data analysis. Using OGAMA to analyze data is generally easy and straightforward, but we have experienced a few problems. When using the playback of each test it sometimes would not change slide or complete the entire test session. There is also no feature to see the users time-to-complete, which means we have to do it manually. Looking at the fixations in our experiments, we found that some was empty and it could indicate that the collected data are at times skewed. This is a hindrance, but as long as it is taken into consideration, the data are still usable. The fixations, scan paths and more, can still show tendencies.

Despite these issues we found that OGAMA did what was needed and have been able to extract a lot of data from it.

The combination of our questionnaire and eye tracking have proved to be a useful and well-rounded solution for our data collection. The questionnaire helped us collect the data needed to establish different groups, which were necessary to be able to compare and analyze on our collected data. The eye tracking provided a lot of data, which can be analyzed in a wide variety of ways and provide valuable insight into how the groups scan for information in an information architecture. Our chosen approach to the test conduction helped the different users complete them without many difficulties. We found that collecting data on demographics and prior knowledge to establish different groups and then collecting data via eye tracking, was a useful and informative method of collecting data on users scanning behavior.

3.3.3 | Ethical considerations

In our test we have made some ethical decisions on the basis of our usage of borger.dk. Using this website could create some ethical problems which we have tried to avoid. Normally when people use borger.dk they log in to their personal account. We wanted to avoid that so we made sure that all tasks in test could be completed without ever logging in. That way, we also ensured that participants did not have to use one of our own personal accounts which could also cause some of them to be uncomfortable.

Even though our usage of a simulation of a website was more of a technical limitation that anything else there are also some possible ethical problems that have been avoided. Borger.dk is a website where citizens, or would be citizens, of Denmark go to find important information and perform important actions. Whether users are finding information on how to change their home address or renew their passport, this is the website to do it. Furthermore, to do any of these things they need to login with their personal information.

For these reasons we would argue that borger.dk is a site associated with a degree of seriousness. If test participants were to visit the live website borger.dk this serious attitude could have affected their behavior and might even made them uncomfortable, as their relationship with the site was usually a personal one and not something to be monitored by strangers; much in the same vain as home banking. In our test introductions we explain participants that there are several limitations to the site; they cannot use the mouse wheel or use the search function. When participants are doing the tests they get messages when they have completed a task or when they have selected the wrong option. You could argue that these factors help convince the participants that they are in a simulation; a fake site. If any of our participants were uncomfortable with using the site this assurance of the site and their actions not being real and not having consequences might alleviate that. All of the above goes well with the idea of laboratory tests which is described in 3.1.3.2 | Test setting on page 55 because we create a controlled test environment.

4 | Analysis

Our analysis takes its point of departure in the measures described in both 2 | Theory and 3.2 | Data Analysis. We have divided it into two parts.

1. Our results presentation and discussion, which consists of two sub-sections:
    - The questionnaire and the data collected through it
    - The eye tracking study data and how it can be analyzed using the data from the questionnaire
2. A summary of our findings and how the relate to our research questions

4.1 | Results presentation

In this section we will present data collected in this study. For each section is this chapter, we will present the data or the elderly group first, then compare it to data for the young group. There will be several instances where the elderly group will be split up in order to analyze on internal differences, these instances will be clearly labeled as such. The primary focus of this chapter is to describe the elderly, as it is the group with the largest sample size. The young participants' results will be included to make comparisons. The data we present is split up in two sections; questionnaire and eye tracking studies.

4.1.1 | Questionnaire

Our questionnaire provides us with useful data on the participants, which can be used to add more detail to our eye tracking study and explain some of the tendencies in the results. In the following we will describe the data collected through our questionnaire. It is split in two sections: Demographics and Prior Knowledge.

4.1.1.1 | Demographics

Our questionnaire provided us with information on the age, gender and educational level of the participants. Below we will go through these three statistics.

Our elderly participants are a broad sample of different ages, as the youngest were 53 and the oldest 88 years old. The median for our participants are 63 years old with a standard deviation of 8.74 years. Eight of our participants were between 50 and 59 years old, 15 were between 60 and 69 years old and the remaining four are older than 70 years old. Three of the 70+ participants were in their eighties. For the young participants the median are 23 years old with a standard deviation of only 1.6 years. The age of the participants spans from 20 to 24 years old.

		Age		
		50-59	60-69	70+
Gender	Men	4	4	0
	Women	4	11	3

Table 3 - Elderly participants age and gender

Of our 26 elderly participants nine were men and 17 were women. Optimally we would have had a fifty-fifty distribution, but this was not feasible with the problems recruiting participants. Four of each gender were between 50 and 59 years old, four men and 10 women between 60 and 69 years and three participants over 70 are women while one is a man. The age median for women are 63, just as the median for all participants, while the men are slightly younger, with an age median of 62. The young group consists of three males between 21 and 23 years old, while the two females are the youngest and oldest of the group at 20 and 24 years old.

		Education level			
		Municipal primary and secondary school	Craftsmanship school	Upper secondary school	Higher education
Gender	Men	2	6	0	1
	Women	8	5	1	3

Table 4 - Gender and educational level of the elder group

The educational level of the participants is 10 at just municipal primary and lower secondary school (two men, eight women). 11 finished craftsmanship school (six men, five women). Only one, a woman, at upper secondary school. Four participants have a higher education (one man, three women). The age median for each are: 62.5 at municipal primary and lower secondary school, 63 at craftsmanship school, 83 at upper secondary school and 59.5 with a higher education. With only one at upper secondary school we cannot make any conclusion based on it. The median for municipal primary and lower secondary school and craftsmanship school are close to or similar to the median for all the elderly participants. The participants with a higher education are younger at 59.5 years old. The educational level of the young group is flat, as all five are studying for their bachelor degrees.

4.1.1.2 | Prior Knowledge

In this chapter we will refer to data which can be found in Appendix 2 with questionnaire results. The most important parts of that data will be described in a table down below. When it comes to the level of prior knowledge for the participants, we calculated the median, from the scores in the questionnaire relating to prior knowledge. The scores were from one to six, with one being the lowest and six the highest.

		Gender	Age	Prior knowledge median
Elderly group participants	A2	Female	64	4
	A3	Female	64	3
	A4	Female	62	4
	A5	Female	68	3.5
	A6	Female	57	1
	A7	Female	64	3
	A8	Female	63	2
	A9	Female	88	2
	A10	Female	55	3
	A11	Male	59	1
	A12	Female	65	6
	A13	Male	67	2
	A14	Female	53	2
	A15	Male	55	4
	A16	Male	55	2
	A18	Female	61	3.5
	A19	Male	68	1
	A20	Female	83	3.5
	A21	Female	83	2.5
	A22	Male	57	2
	A23	Male	62	1
	A24	Female	58	3
	A25	Female	60	2
	A26	Male	73	3
	A27	Male	65	4
	A28	Female	63	1
Median			63	2.75
Standard deviation			8.89	1.23

Table 5 - Elderly participants' prior knowledge

The participants have a median prior knowledge score of 2.75 and a standard deviation of 1.23. This show us that the participants have a fairly high prior knowledge, considering they should have a low one, if the notion of native and immigrants is to be believed. The standard deviation is fairly low as well, indicating a homogenous group. Comparing the six oldest and six youngest participants, we find that the oldest have a median prior knowledge score of 2.75. The exact same as for the entire group. The youngest six have a score of two, considerably less than the oldest, which goes against the notion of immigrants and natives. When dividing them into age groups, the numbers are: 50-59 = 2, 60-69 = 3 and 70+ = 2.75. These numbers show us that the difference between the three age groups are limited, but the youngest are the lowest, with a score lower than the median for the entire group. This might be because only two of the nine 50 to 59 years old, have a college degree and the rest have attended craftsmanship school (five) and municipal primary and lower secondary school (two). Looking at the self-evaluation score for use of computers in work life, we found that the 50-59 year olds and 60-69 year olds both scored high, with a median of 5 and 6 respectively. The 70+ year olds scored three. This goes against what was just discussed on why the difference in prior knowledge exist and we were not able find anything in our data that can explain it. It can be caused by our sample size and the participants recruited. We might have gotten different results with a larger and more demographically diverse group of participants.

Comparing the two genders shows that the women have a higher prior knowledge score than the male. The scores are three and two respectively, with similar standard deviations of 1.21 and 1.20 respectively.

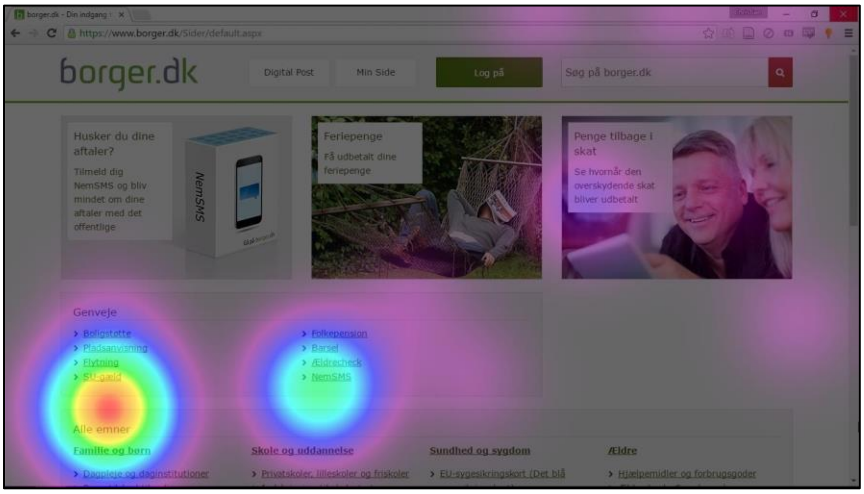
Of our 26 participants only five did not have any prior experience with borger.dk. The prior knowledge for these five participants are two, considerably less than the median for the entire group of 2.75, which shows that they are more inexperienced navigating this type of website. The age for these participant's ranges from 53 to 88, which includes both our youngest and oldest participant in the elderly group. Only one of the young participants did not have any prior experience with borger.dk. This participant was also the youngest, which might explain why, as she has not have had much reason to use it up to this point.

Among the young group the prior knowledge median score is 5 with a standard deviation of 1.3. The highest score is shared by three of the five participants, while the lowest is at two. The prior knowledge score for the young participants are almost double that of the elderly. Scores for how much they use the Internet are sixes for all in both daily internet usage and time spent each week online. The elderly has prior knowledge median scores of three and two respectively, which is far less than the young participants. The lowest scorer among the young participants said that she did not use the Internet for activities, such as what borger.dk facilitates, but are constantly active on social media and browsing the Internet. This show that the prior knowledge measurement in our study is more focused on domain expertise, but it does, however, include technical expertise, as we measure the participants use of computers in general. The young participants have a prior knowledge median score of one for work use of a computer, with a standard deviation of 2.49. This shows that there is a big difference in the five participants, where three have marked the score as one, one marked it as five and one as six. The lack of experience, compared to the elderly with a score of five with a large standard deviation of 2.07, is offset by their experience and prior knowledge gained from their studies, where they have scored six all of them.

4.1.2 | Eye Tracking Study

In the following we present our findings in our eye tracking study and use the data described in the previous two chapters on demographics and prior knowledge to compare the different groups. The analysis will be conducted by including the measures mentioned in 2.2 | Eye tracking.

Analyzing our eye tracking data have shown us some issues with the data. When choosing some participants, it is clear that their data is skewed. By this we mean that we have good data, but it is slightly offset from where the participant was actually looking. See Picture 8 for an example, where it is clear that the participant, was looking at the shortcuts, but the data shows the participant was looking a little below it. This affects the results, so we will



Picture 8 - Skewed data

attempt to clean the data and exclude the participants that provided no data on certain slides, as they might have taken another click path or their eyes were simply not tracked. Participants with less than four recorded fixations are excluded from the scan path analysis as well, to ensure the scan paths have lengths long enough for a proper comparison. If the data from a participant is clearly skewed on a slide, we will exclude that as well, as the skewed data might indicate that the participant was looking at an AOI he/she was not and fixations wrongly placed might influence the data negatively.

4.1.2.1 | Click paths

In this section we will present our results of which click paths our participants had taken through the tasks. Firstly, we will present each available path to complete a given task. Then we will present the amount of users who went through each path. In this section click paths will be referred to as CPs. Furthermore, we will account for our participants clicks to complete.

In Table 6 below is the result of our elderly participants' uses of click paths. In the following these will be elaborated on in detail.

Click paths	Task 1	Task 2	Task 3
Occurences			
A	26	0	12
B	0	1	1
C	0	0	0
D		25	1
E			12
F			0
Percentage			
A	100,00%	0,00%	46,15%
B	0,00%	3,85%	3,85%
C	0,00%	0,00%	0,00%
D		96,15%	3,85%
E			46,15%
F			0,00%

Table 6 - Click Paths – Elderly participants

Task 1 – Cykeltyveri

In this task users were asked to find information of how to report a bicycle theft. There were three paths to the final destination.

Click path A - This is the fastest way of completing the task alongside CP B. The participants select a sub header called "Cykeltyveri og anden kriminalitet", which is the exact thing they are looking for information about; "cykeltyveri" or bicycle theft. Selecting that option brings them to a landing page where they have the option of reporting the theft by clicking a big green button called "START". The following page will have a box that is identical but this time it will read "VIDERE", in English in would similar to a button labelled Next. This is the final act for this click path. The clicks to complete amount to three.

Out of 26 participants, 26 chose this click path. The reason for this we will discuss after presenting the other CPs.

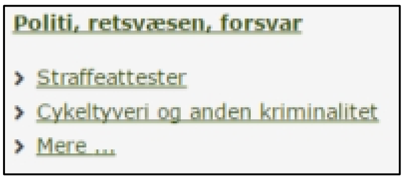
Click path B - Alongside A this is the fastest way om completing the task. While A would have participants click the sub header of a menu, this click path regards the main header of that menu; "Politi, retsvæsen of forsvar" which means police, legal system and defense. In the following page participants would have to find the right option in a list of shortcuts. The right option reads "Anmeld cykeltyveri til politiet"; Report bicycle theft to the police. The last page is identical to the last page in click path A. As with CP A the clicks to complete amounts to three.

Out of 26 participants, nobody used this path.

Click path C - This is the longest and most complicated way to get to the desired information. It follows the first step in CP B; selecting the header. However, for users to take this path they would have to ignore the shortcuts on the bottom of the page and use the local navigation to select: "Cykeltyveri og anden kriminalitet" which is the exact same page that the sub header from CP A directly linked to. The rest of the path is identical to click path A. Consequently, this click path requires four clicks to complete.

Out of 26 participants, nobody used this CP.

Seeing as there is such a clear tendency in these results we will discuss briefly what this could mean. The nature of the task requires users to find the appropriate box with the header and the three sub headers. From that we can argue that users must have seen both of the two entry points, since they are located directly above and below one another. See Picture 9 to see the positions. Looking at fixations for our participants, we can see that they indeed did look at both the main header and the sub headers. This is shown in Picture 10. This ties in well with the wording found on the instruction slide to this task. The instruction tells the participants specifically to find information on how to report a bicycle theft; "cykeltyveri". The sub header that 100% of participants selected was named "cykeltyveri og anden kriminalitet". From this data you could argue that participants read both options and chose the one most closely related to the instruction. For this task the average clicks to complete amongst participants was three, naturally, because every participant used the same click path.



Picture 9 - Clean menu

Picture 10 - Menu fixations

Task 2 – Statsborgerskab

In this task users were required to find information on how to apply for Danish citizenship. There were four paths leading to completion. Much like in Task 1 the CPs have two entry points in the form of a header and sub header.

Click path A - This CP starts with users selecting the header "Udlændinge i Danmark. On the following page they will select an option from the local navigation to make it drop down, then select an option from the now expanded menu. This leads to a page very similar to the one found in Task 1 with the green start button and the progress button after that. This amounts to four clicks in total.

This is the joint longest click path in this task and none of our participants went this route.

Click path B - This path also begins with selecting the header. However, on the following page participants can click on a big picture that says "Dansk statsborgerskab" or Danish citizenship. Then much in the vain of CP A, participants must select the green buttons. This CP requires four clicks to complete.

Two participants of 26 went this route.

Click path C - Like the previous two this CP begins with users selecting the header. However, this time user select a shortcut found in a list of shortcuts which will end the task. This is by far the shortest click path in Task 2 as it only requires two clicks to complete.

None of our participants used this CP.

Click path D - Unlike previous paths this one has a different entry point; a sub header just like in Task 1. This sub header functions exactly like it did in that task too where it is found below the header and it reads: "Ansøgning om dansk statsborgerskab"; applying for Danish citizenship. Selecting this lead the participant to a page with a green box which read start and then another green box which read progress. Very similar to CP A in Task 1. Conversely, all participants used that CP in Task 1 and in this task 25 out of 26 used CP D. The clicks to complete for CP D amounts to three.

Given the similarities between the first two tasks there are a few things we can discuss. There is a possibility that what was discussed in Task 1 applies here in the same way; that users read both the header and the underlying options to then identify the specific word they were looking for. Picture 11 supports this notion, as it is clear from the fixations that the participants read both the header and the sub header. Picture 12 shows the header and sub headers without the fixations. However, there is also the possibility of them having success in the previous task, had an effect on this one. We know from the click path results in Task 1 that every participant used the same method previously so there is a high possibility that some were affected,

Picture 11 - Menu Fixations

which can have caused them to use the same method again. Additionally, having these two tasks solved in such a similar fashion might be an indication of the participants search behavior being very similar. We will investigate our participants search behavior later in the analysis, in section 4.1.2.5 | Scan paths.

PIcture 12 - Clean menu

Task 3 – Folkepension

In this final task participants were asked to find information on how they could apply for state pension. This task was designed with the aim of having more click paths then previous tasks and thus have six CPs. Due to having so many and how similar they are, they will be presented in groups.

Click path A & B - Both of these CPs have their first click at a shortcut presented on the first slide. This is different to both previous tasks as they have required participants to go the another slide to find the first click target. This shortcut leads to a page which has a list of popular choices. The first one listed is a click target which concludes CP A with two clicks. If participants instead are inclined to use the local navigation they will be presented with a page very similar to them at this point, a green button on the right that reads "START" which is the click target; this is CP B and requires three clicks. Out of 26 participants, 12 used CP A and 1 participant used CP B.

Click path C & D - These click paths starts at a header much like in previous tasks; the header called "Pension og efterløn" meaning pension and early retirement benefits. Clicking here will lead to a page closely related to the one in CP A & B but they are not the same. On this page there is a list of popular choices, the first listed option is a click target and this is CP C which brings the total clicks to complete to two. The two click paths diverse if users instead make us of the local navigation and selects the option "Folkepension og tillæg" which is the same page the shortcut from previous CPs lead to. Technically this means that the CPs overlap and there would be more than six click paths. However, seeing as only one person used one of these options we will not go into detail with all available solutions. The one participant who used either of these CPs used D which required four clicks.

Click path E & F - Both these click paths starts with a sub header found beneath the header in C and D. This sub header reads "Folkepension". This link is identical to the shortcut found on path A and B and leads to the same page. Therefore, the rest of the paths E and F are identical to A and B respectively which also means that they share the same amount of clicks to complete. Out of 26 participants 12 used path E and nobody used path F.

To comment, the same tendency as discussed in Task 1; where users would select the option with related word, seems to prevalent here as well and this indicates the participants employed lexical matching. 12 participants made use of the shortcut which said the related word; folkepension. 12 participants, when given the choice of the header or sub header, selected the sub header with the obvious word attached. When comparing these two groups of 12, we found that the participants that chose CP A had a prior knowledge median score of 2.0 which is lower than the median for all participants of 2.5. The participants that took CP E has a score of 3.0 which is above the general median. It is difficult to make any conclusions based on this, but it might be that the CP E group's high prior knowledge helped them learn to use borger.dk and thus knew where to go, to find the correct link. Of the 12 participants who took CP A three had no prior experience with borger.dk, while it is only one of 12 in CP E. This might be a reason, as the less experienced Internet users and borger.dk users did not know how the navigation work on borger.dk. When looking at the two groups age, we found that the CP E group had an age median of 64.5, while the CP A group had one of 60. The difference is not big and is within the standard deviation for the age median for all participants and it does not seem that age was a factor.

To sum up our 26 elderly participants use of click paths, in Task 1 and 2 the differences were almost nonexistent as everyone did the same in Task 1 and 25 did the same in Task 2. However, in Task 3 we experienced much more diversity in the click paths as shown in Table 7 where the distribution of all CPs can be seen. The CPs for Task 1 and 2 clearly shows that the elderly participants follow the same paths, which are the fastest as well with three clicks in both. Task 3 divided the elderly into two groups where the difference was whether they had found the shortcut or not. Interestingly, the group that used the shortcuts where also the group that had the lowest prior knowledge score and the most participants with no prior experience with borger.dk. Age did not seem to be a factor, even though a small difference in age median of 4.5 years existed. Noticeably, both of the popular click paths in Task 3 were also the shortest with two clicks in each.

Click paths	Task 1	Task 2	Task 3
Occurences			
A	26	0	12
B	0	1	1
C	0	0	0
D		25	1
E			12
F			0
Percentage			
A	100,00%	0,00%	46,15%
B	0,00%	3,85%	3,85%
C	0,00%	0,00%	0,00%
D		96,15%	3,85%
E			46,15%
F			0,00%

Table 7 - Click paths for the elderly

The results collected on click paths show that our participants were homogenous in Task 1 and 2 which are quite similar in their structure. However, the separation between users in Task 3 was evenly divided between two paths. Interestingly, as explained earlier these CPs are extremely similar after the initial click and even have the same amounts of clicks required to complete. In summary, most of our participants have used the same amount of clicks for their task completion even though in some cases they did not follow the same paths.

If we compare this data to the young group, we find similar results as four out of five used the same paths for Task 1 and 2 as the elders did. One participants used a different CP for Task 1 making them the only one out of 31 participants. In regards to Task 3, the young group are very similar as well as participants are divided into two paths; coincidentally the same paths used by the elders.

These results on click paths and clicks to complete does not correspond with established literature on the subject that states that elderly use more clicks to complete tasks than young people (Kurniawan, Zaphiris, and Ellis 2002; Mead et al. 1997). This literature was described in section 2.1.1.4 | Search behavior in elderly users.

Given these results on both the elderly and the young group, one could suspect that this would mean that they have all spend the same amount of time completing the tasks. In the following section we will investigate our participants' completion times.

Option	Task 1	Task 2	Task 3
Occurrences			
A	4	0	2
B	1	1	0
C	0	0	0
D		4	0
E			2
F			1
Percentage			
A	80.00%	0.00%	40.00%
B	20.00%	20.00%	0.00%
C	0.00%	0.00%	0.00%
D		80.00%	0.00%
E			40.00%
F			20.00%

Table 8 – Click paths for the young participants

4.1.2.2 | Time to complete

In this section we will present the data we have collected on participants' completion time for the tasks in our test. These results will be condensed numbers as the time taken to read instructions and going to the next task has been left out.

Test participant	Task 1	Task 2	Task 3	Total time spent
	Time spent			
A2	0:01:03	0:00:31	0:00:16	0:01:50
A3	0:00:52	0:00:24	0:00:50	0:02:06
A4	0:01:22	0:00:35	0:00:28	0:02:25
A5	0:00:53	0:00:23	0:00:07	0:01:23
A6	0:00:36	0:00:30	0:00:44	0:01:50
A7	0:01:36	0:00:30	0:00:51	0:02:57
A8	0:00:49	0:00:16	0:00:06	0:01:11
A9	0:02:14	0:01:17	0:01:24	0:04:55
A10	0:01:09	0:00:36	0:00:52	0:02:37
A11	0:01:13	0:00:50	0:00:14	0:02:17
A12	0:01:09	0:00:38	0:00:18	0:02:05
A13	0:01:52	0:01:10	0:00:19	0:03:21
A14	0:00:24	0:00:22	0:00:08	0:00:54
A15	0:00:33	0:00:32	0:00:06	0:01:11
A16	0:00:17	0:00:13	0:00:11	0:00:41
A18	0:00:12	0:00:12	0:00:45	0:01:09
A19	0:03:33	0:05:02	0:01:00	0:09:35
A20	0:02:15	0:00:45	0:01:05	0:04:05
A21	0:01:54	0:00:24	0:00:25	0:02:43
A22	0:00:49	0:00:25	0:00:18	0:01:32
A23	0:01:24	0:00:43	0:00:35	0:02:42
A24	0:00:43	0:00:24	0:00:06	0:01:13
A25	0:02:08	0:01:02	0:00:37	0:03:47
A26	0:00:42	0:00:38	0:00:15	0:01:35
A27	0:00:31	0:00:38	0:00:08	0:01:17
A28	0:01:00	0:00:51	0:00:24	0:02:15
Average	0:01:12	0:00:46	0:00:29	0:02:27
Standard deviation	0:00:46	0:00:55	0:00:22	0:01:47

Table 9 – Elderly time to complete

This table presents the completion times for our participants. The total time spent column, however, is not indicative of how long time the full test took to conduct.

Instruction slides and end slides have been excluded from these to only illustrate the times that participants were actively completing tasks; not reading instructions.

Our results show that our elderly participants are different from each other. While our fastest participant finished the tasks in 41 seconds, the slowest spend 9 minutes and 35 seconds. This slow time is a big deviation from the rest, as the second slowest used 4 minutes and 55 seconds. In order to explain the variations, we will group the participants time in minutes.

Our two largest groups have spent between two and four minutes completing the tasks. These two groups combined make up 19 of 26 participants. Contrasted to the fastest and slowest it is closer to the fastest participant's time but the amount of people faster than these two groups amount to two while the amount that were slower amounts to five. It shows that the deviation from the two large groups are frequently slower. When accounting for gender as a demographic factor in completion times, results show us that male participants were slower than female with an average time spent on all tasks of 2:41 compared to 2:19 minutes. However, one male participant proved to be very different from the others as his total completion time was calculated to be 9:35 minutes. If we were to remove this outlier from the rest, the male average would decrease to 1:49 minutes. Noticeably, this would make them faster than the female group. If we were to remove the slowest of the female participants, with a total time of 4:55, the average time would decrease to 2:09 minutes which would still be slower than the modified male average.

Time	Participants
1-59 sec	2
1:00-1:59 min	10
2:00-2:59 min	9
3:00-3:59 min	2
4:00-4:59 min	2
5:00-9:59 min	1

Table 10 - Time groupings

In our tests, the females have proven to be faster at completing the tasks. However, we suspect that a fairly large anomaly has caused this to be the case. In the following we will discuss individual task completion times.

The table in Appendix 3 represents our results on participants' individual task completion times as well as how the concepts of prior knowledge and age have affected these. In this table the amount of participants for each time interval will be shown, alongside the participants' age median and prior knowledge median score.

As can be seen in the table there are some clear differences in the amount of time spent completing the tasks. Most notably, many participants take less time for each completed task. Task 1 and 2 are very similar in their structure, the participant will select an option from a menu, then click twice and the task is done. In theory their completion time should be similar, however, 23 participants out of 26 spend far more time on the first task than they did their second. This could be an indication of learning the site and the structure which can happen when testing without a random order of tasks. The transition from Task 2 to 3 is somewhat similar as 21 participants were faster as well. This supports the notion of participants gaining domain expertise and learning the structure of the website, as described in chapter 2.1.1.3 | Prior Knowledge. However, one must note that the third task has a "shortcut" involved which is located on the first slide in contrast to the first two tasks which had their options on the second slide. This could indicate several things. That some participants had observed these shortcuts a third time despite not having success in doing so the first two times or it could mean that participants had remembered this option.

It can be noted that the fastest time was 12 seconds while the slowest was 3 minutes and 33 seconds for Task 1. This is a significant difference between two users. The reason for highlighting these figures is because it is the first task that participants has a chance to scan the site for the information they need. Compared to the following tasks, this is a part of the test that can give an indication of how fast or slow participants would find the desired information, on a site they might or might not have experience using. In this task there is also a minimal risk of having participants learning the site as there would be in Task 3.

In regards to the notion of prior knowledge and age having an effect on participants' completion times the effect seems to be very miniscule. The prior knowledge median scores for the elderly participants were evenly distributed amongst the first three completion time groups in the table in Appendix 3 ranging from two to three. Only in the slowest time were the median lower at one. Even though the difference between a median score of one and three is significant when six is the maximum, there are not enough instances of it happening for it to be named a tendency. From these results it would seem that prior knowledge did not affect our elderly participants' competition times in a significant way. When examining the age differences of our elderly participants we see a small pattern. In the fastest time groups, the age median scores are generally lower than in the slower time groups. This is indicative of age affecting our participants' ability to solve tasks faster.

In the previous section we investigated click paths and how our participants went through our tasks. As described in that section, a vast majority of participants used the same, fast click paths. This corresponds well with the similar completion times found in this section. However, there are cases of participants with the same click paths with vastly different completion times. The most extreme case is two participants with the times 1:11 and 9:35. Both of these participants were required to make eight clicks to complete their tasks but had severely different completion times. In section 2.1.1.4 | Search behavior in elderly users on page 18 we described the concept of verification strategies; the idea of users revisiting previously seen menus to verify that they had reviewed all options before selecting one. To investigate this concept, we isolated our participants who had used the most frequent and fastest click completions. 24 out of 26 participants used a click path with eight clicks. The time to complete median score for these participants was 1:58. Of the 24 participants, only 17 had valid data that we could analyze upon. Of the 17 remaining 12 was found to be faster than the median and five were found to be slower. Of the 12 faster participants, two were found to have used a verification strategy while three out of five had done so in the slower group. The difference between having two participants in one group and three in the other is miniscule, however, with large amount of data being unusable, there is no way of knowing what that data would have shown. Additionally, it is interesting that the invalid data for this particular investigation was all from the slower group. One could speculate that poor eyesight could have affected users' completion time negatively since it was the eye tracking data that proved invalid. To summarize on clicks to complete having an effect on completion time, it would seem that it makes a very small difference. From the limited data we have we can conclude that five of 24 used verification strategies. Since two were from the fast group and three from the slow group this could indicate that it did not make a difference. However, all of the missing data was from the slow group so if there were to be more, they would all be in that slow group.

When comparing all these results to the young group there are both similarities and differences.

Test participant	Task 1	Task 2	Task 3	Total time spent
	Time spent			
B1	0:00:41	0:00:28	0:00:09	0:01:18
B2	0:00:39	0:00:22	0:02:10	0:03:11
B3	0:00:18	0:00:18	0:00:05	0:00:41
B4	0:00:25	0:00:18	0:00:07	0:00:50
B5	0:00:31	0:01:44	0:00:07	0:02:22
Average	0:00:31	0:00:38	0:00:32	0:01:40

Table 12 - Time to complete young participants

By going over the data of all three tasks at once it becomes obvious that there two major deviating participants; B2 and B5. Since both of these participants are also the only ones for the young group with any errors in their tasks, they will be discussed in the next section of the analysis. As a whole, the results are very homogenous just like with the elderly. Noticeably, the young group has faster results for each task they complete and the variation in completion time gets smaller and smaller for each task. As a whole the young group is faster at completing Task 1 and 2 but not 3. This corresponds well with existing research which also states that elder users are slower than their younger counterparts(Artis and Kleiner 2006; Kurniawan, Zaphiris, and Ellis 2002; Biterman, Lerner, and Bitterman 2007; Mead et al. 1997; Chadwick-Dias, McNulty, and Tullis 2003; Czaja et al. 2001; Etcheverry, Terrier, and Marquié 2012). This existing research was described in section 2.1.1.4 | Search behavior in elderly users.

In summary the completion time for our tasks has showed that participants are fairly similar with a few exceptions. It was found that age, prior knowledge and click paths had a small effect on participants' completion times. In the following section we will present data on how many errors our participants made during their tests.

4.1.2.3 | Errors

In our test we made a simulation of a website; borger.dk. In this simulation we were interested in giving participants a way to go back to a previous slide just like you would go back to the previous page in a browser. However, a limitation in our chosen software made that impossible and we had to go another route. Instead we added error slides. The error slides were added to serve three purposes; to give users the knowledge that they had selected the wrong option, to give them a way of going back to the previous slide and give ourselves a method of logging how many times users made the wrong choices. As explained in section 3.1.2.5 | Task list on page 42, that meant that for every slide in the OGAMA experiment we had to add an error slide.

In the following we will account for the amount of errors that the elderly group made during their tests so we can use this data to compare with the young group but also to supplement possible tendencies. In the table, the 10 participants who made errors are accounted for.

The results show that the amount of errors are somewhat small. Only three participants made errors in several tasks. There were two participants who made a significantly higher amount of errors, three and six respectively. These two are both characterized by the errors falling within the same tasks; e.g. the participants with six errors had five of those in Task 2. Both participants followed the fastest and most popular click paths available, just like 22 other participants. All participants who made errors had a prior knowledge median score of 2.5, which is slightly below that of all the participants (2.75), which does not indicate that prior knowledge influenced how many mistakes they made. The same goes for age, as their mean age are 64 years old, only one year older than the mean age for all participants (63).

Test participant	Total errors	Task 1	Task 2	Task 3
	Errors			
A2	1	1		
A3	2	1		1
A4	1		1	
A7	1	1		
A9	1		1	
A13	3	3		
A19	6	1	5	
A20	1	1		
A23	1			1
A28	2		1	1
Total	19	8	8	3

Table 13 - Errors elderly participants

In the previous section we stated that we would follow up on results found in the completion time for the young participants.

Test participant	Task 1	Task 2	Task 3	Total time spent	Total errors	Task 1	Task 2	Task 3
	Time spent				Errors			
B1	0:00:41	0:00:28	0:00:09	0:01:18	0	0	0	0
B2	0:00:39	0:00:22	0:02:10	0:03:11	5	0	0	5
B3	0:00:18	0:00:18	0:00:05	0:00:41	0	0	0	0
B4	0:00:25	0:00:18	0:00:07	0:00:50	0	0	0	0
B5	0:00:31	0:01:44	0:00:07	0:02:22	7	1	6	0
Average	0:00:31	0:00:38	0:00:32	0:01:40				

Table 14 - Time to complete and errors young participants

As can be seen by this table there is a direct connection between the completion time and the errors committed by each specific participant. For that reason, we are going to investigate why these errors were made. For participant B2 five errors were made in Task 3. In Task 3 participants were charged with finding information on state retirement, in this case "folkepension". This participant made four of the five errors on the same page. The following screengrabs are where the participant clicked for each error. The first picture will highlight the target groups that the participants are supposed to click.

Picture 13 - Error 1

Picture 14 - Error 2

Picture 15 - Error 3

Picture 16 - Error 4

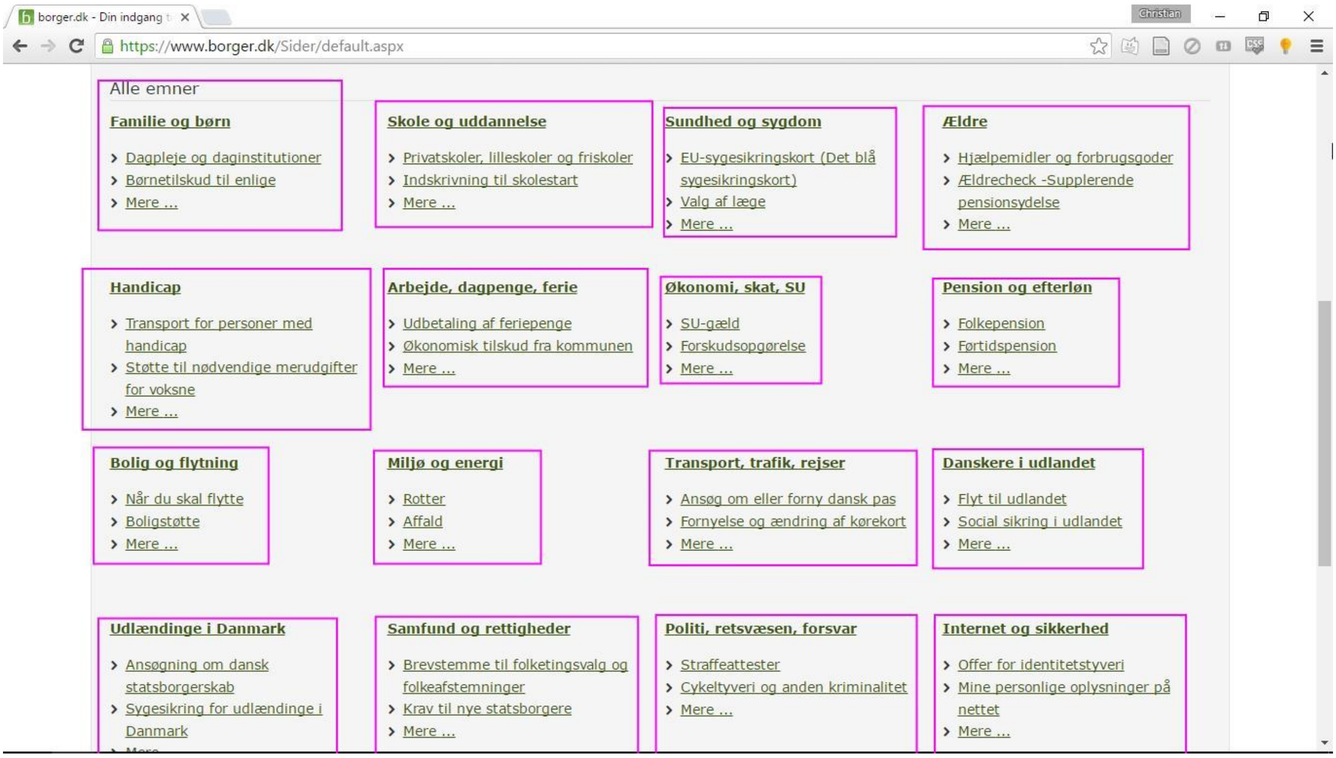
These screengrabs show that this participant clicked four times in the same menu, which was also a wrong menu. It would seem that this person had decided that even when told that it was the wrong choice he still believed the answer to be found in this menu. Noticeably, he clicked the same sub header twice even though he were already informed of this being the wrong option. The key to this behavior would seem to lie in the name of the sub header clicked twice; "pensionsydelse". This is of course closely related to the correct option. It is interesting how the menu below, which contains the "magic word" three times, gets ignored while the participant keeps on getting error messages. Unfortunately, our data on this participant's fixations is not of a high enough quality to support of dismiss any claims. Based on the data collected on errors there is the possibility that this participant has utilized lexical matching in his information search. Furthermore, it would seem that he did not use a verification strategy as the data clearly shows how the same wrong menus is being selected without the other menus being scanned. In regards to the other young participant with many errors, all of these six errors was all found on two front page related pages. In this case the participants started clicking on seemingly random menus with relatively short interval, indicating a sort of trial and error approach.

If we compare this error behavior to our elderly group, where we had two participants with three and six errors, there is no correlation to the first participant of the young group. The elderly participants did not click the same menu more than once when they were told it was the wrong option. However, the one elderly participant with six errors did click on different menus with widely different content suggesting a trial and error approach.

In this section we have found that age does not affect how many errors the participants made. This data corresponds well with existing research, described in 2.1.1.4 | Search behavior in elderly users, as several studies have found that age do not make a difference when making errors in searching for information(Artis and Kleiner 2006; Kurniawan, Zaphiris, and Ellis 2002; Biterman, Lerner, and Bitterman 2007).

4.1.2.4 | Areas of interest

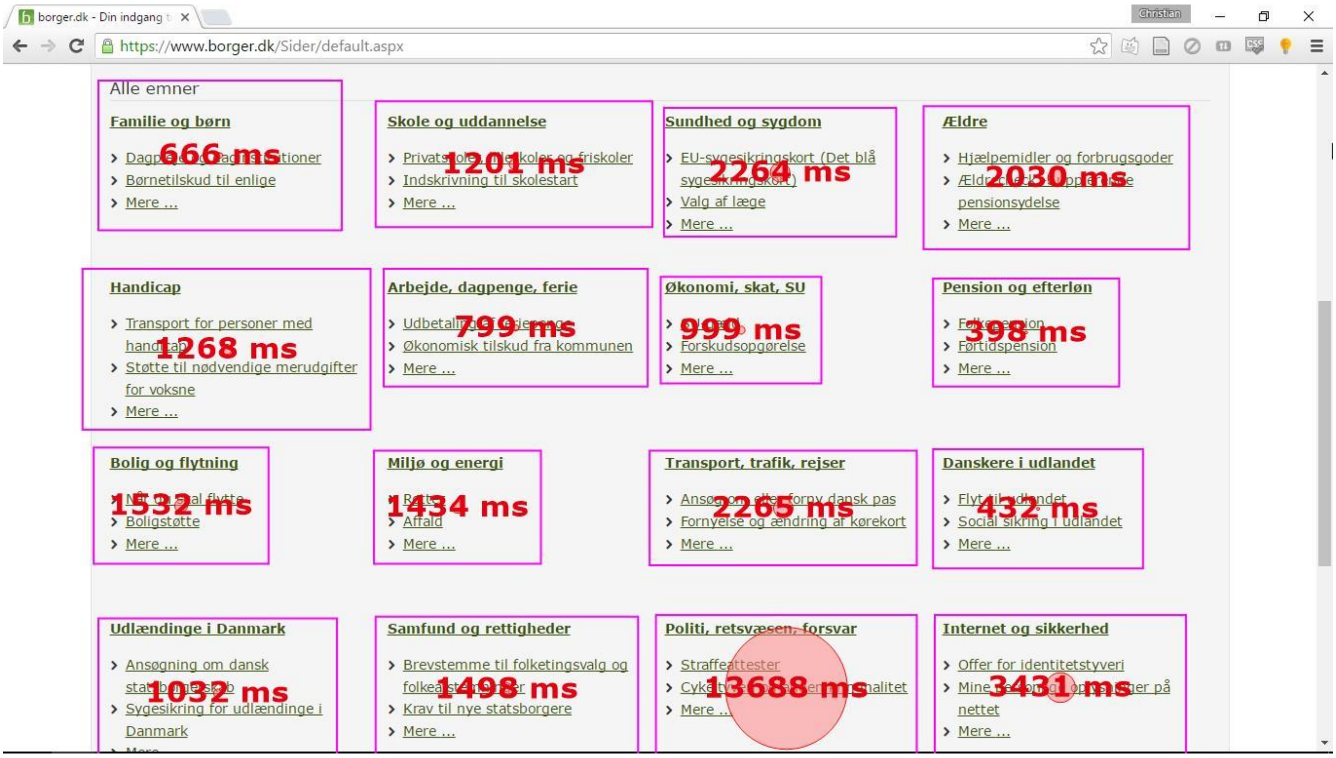
In this study we have created a OGAMA experiment in order to conduct eye tracking research. AOIs have been placed in manually chosen places on the slides. This was described in both section 2.2 | Eye tracking and 3.1.2.5 | Task list. In this section we will present data on how much time our participants have spent in each AOI, the amount of fixations and the transition values between these areas. It is important to note that some of our data proved to be too low quality when we examined it further, which means that these results are based on 12 elderly participants, not 26. The data presented will be connected to Task 1 in our test; because it is the first time in the test that users see this menu grid, without the purple boxes of course:



Picture 17 – Clean AOIs

These purple boxes each represent an AOI. Our reasoning for placing them like this is so we can measure and create data in accordance with these menus. Our reason for using this particular slide from the experiment is this; this is the first time in the test that our participants have seen this grid. Therefore, we must assume that they are going to scan it for information. Had we used the identical slide from the other two tasks there is a risk of participants having learned the grid from before.

We will start by presenting how much time each of our elderly participants have spent looking at these menus. The first screenshot is the total amount of time spent by all elderly participants:

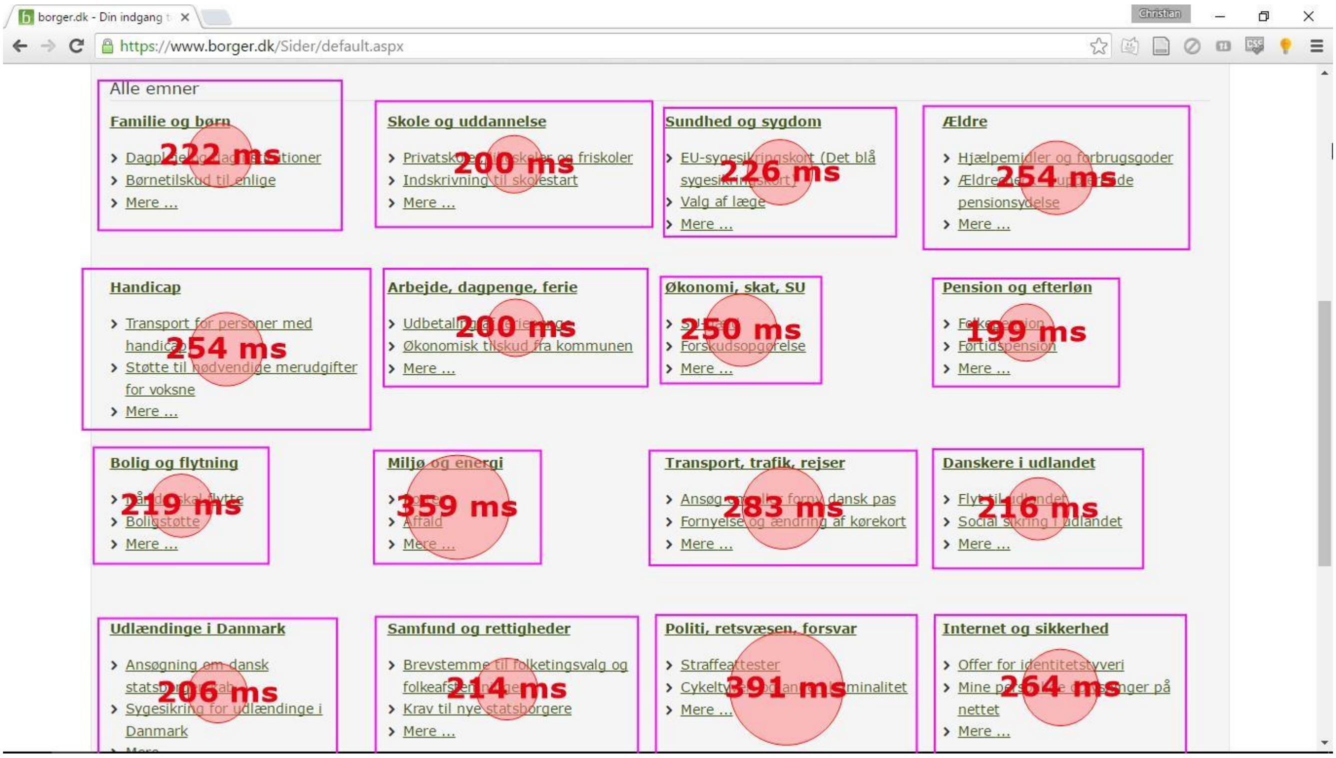


Picture 18 - Total fixation time

As can be seen by the screenshot, some menus have been looked at more than others. Obviously the menu in the bottom right is the target which the participants are trying to find. By looking at the screenshot it becomes clear that all menus have been seen and looked at. One might suspect that the smaller menus would have lower values because they fade in comparison to the big but that is not the case as can be seen clearly by the small menus to the left having equal or higher values than most others. From this data one might suspect a higher number of fixations located in the more popular menus.

Picture 19 – Total fixations elderly

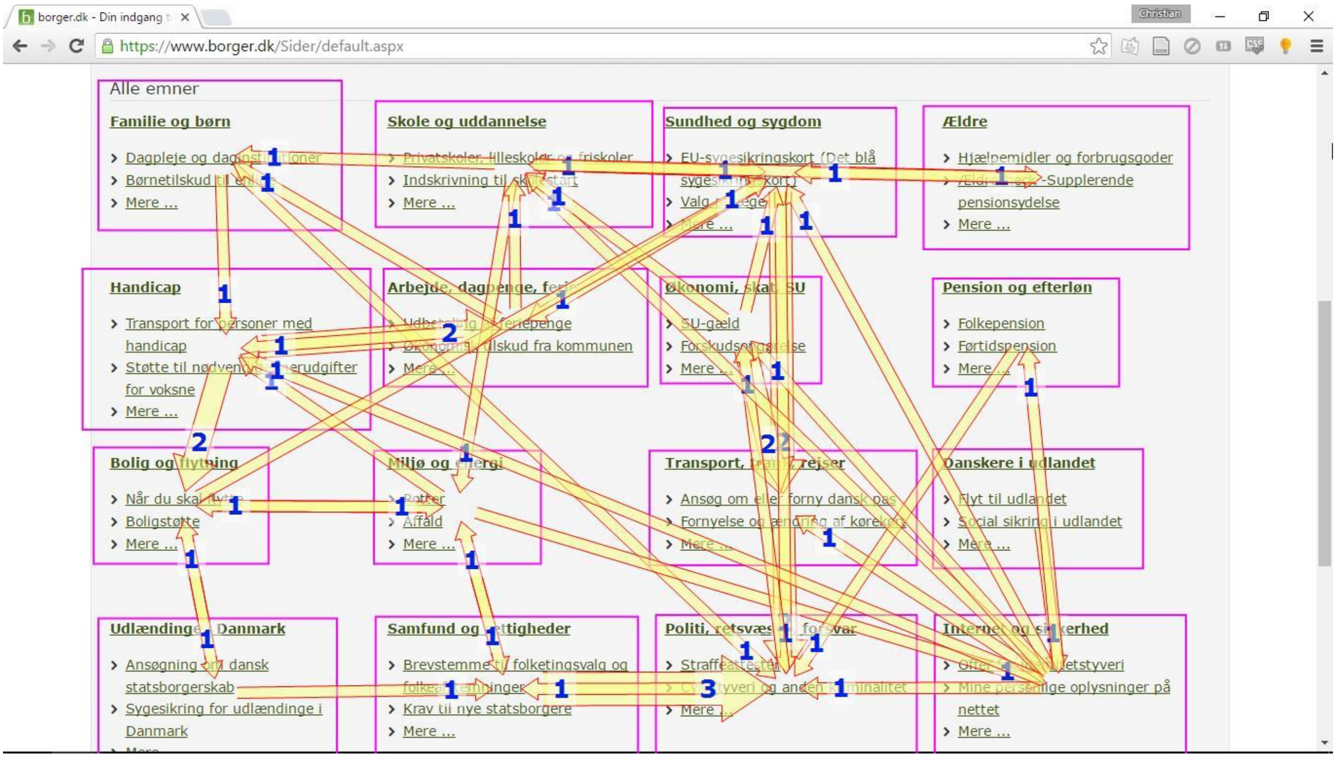
The number of fixations suggests that participants did spend some time reading in all menus. It suggests that some menus were glossed over fairly quickly while others caught the eye. Obviously, the target area had most fixations as one would think that most participants spend time reading the content in depth. To support this claim we present the average fixation duration for each menu.



Picture 20 – Average fixation times

By looking at the average fixation duration it becomes clear that the fixations recorded in each are on average of a somewhat large value. Interestingly, given that the main target menu had 34 fixations, it is telling that the average time per fixation is still the highest. In our theory, in chapter 2.1.1.4 | Search behavior in elderly users, we described how Al Maqbali et al. states that the minimum time for a fixation to be classified as a fixation being between 100-200 milliseconds(Al Maqbali et al. 2013). Seeing as the average fixation on the target area was almost double as high as the minimum requirement for a fixation to exist, we would argue that this data supports the claim that some users have been reading the content of the menu. This claim was previously discussed in section 4.1.2.1 | Click paths.

Since the previous screenshots have indicated that every menu had been inspected by our participants, the following will present the transition values between the menus, to account for how participants have moved between them.

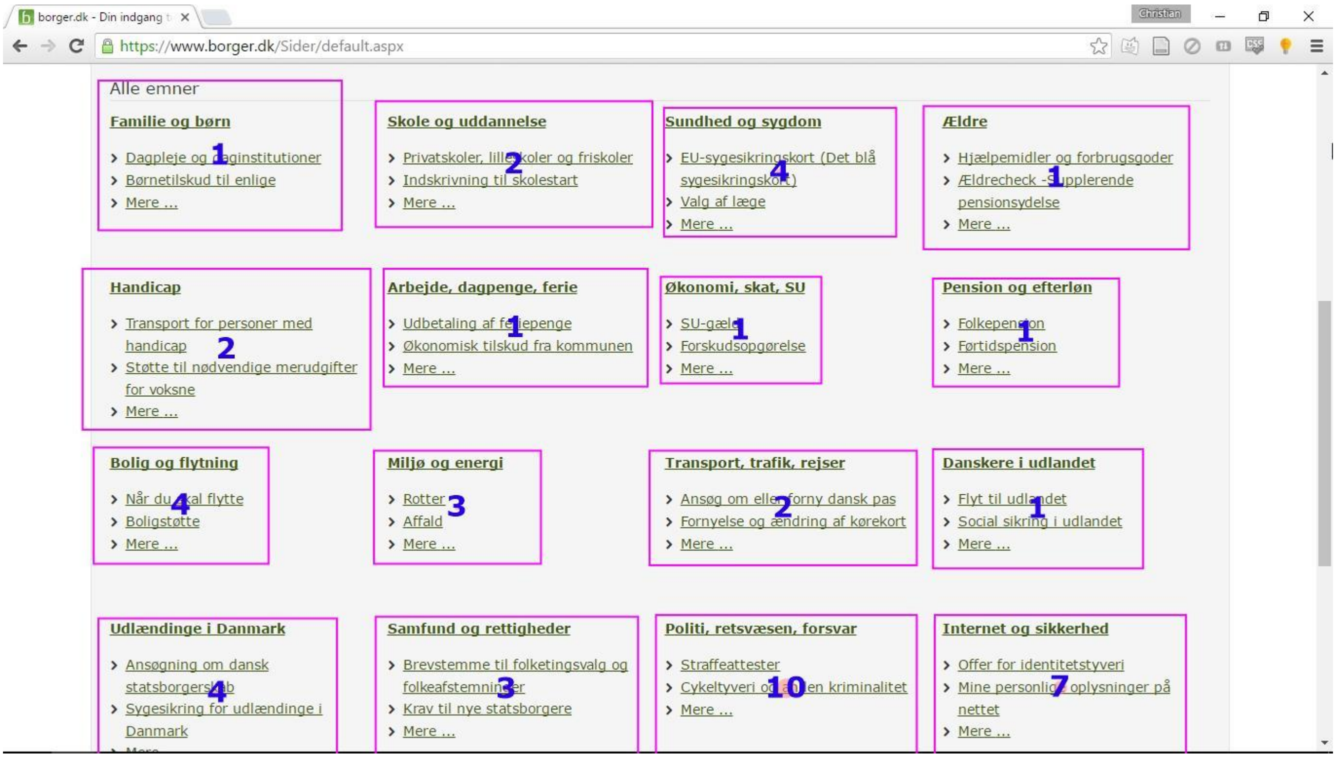


Picture 21 – Total transitions elderly

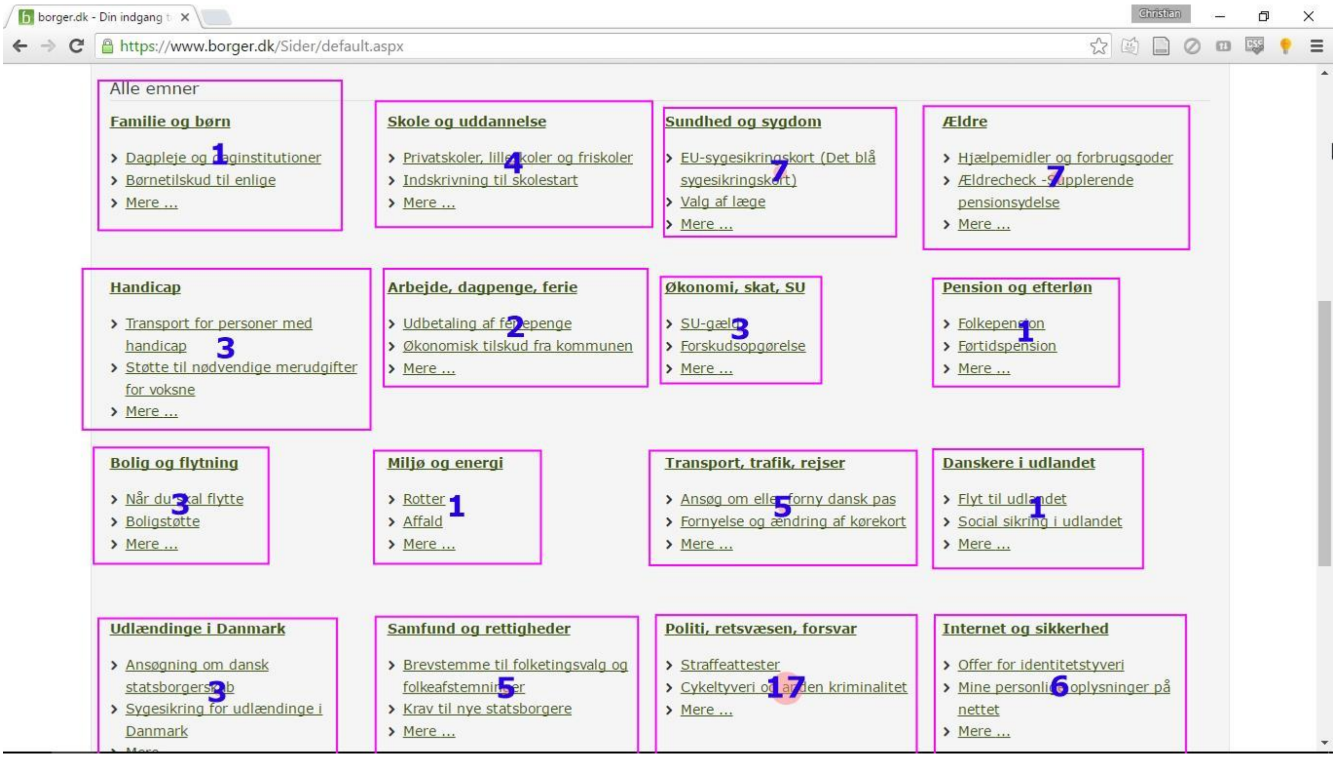
The arrows indicate how participants have transitioned through the menus. It is interesting to note that participants have had very different transitions through these menus as there are not many repeats. The few instances of two identical transitions cannot suffice in showing a pattern. However, it is interesting that three instances exist of participants reaching their goal through the same menu. Hidden in the numbers, in the target menu, is another value of two; which means that there are also instances of users coming from two menus above directly to the target menu. All in all, these transitions show how diverse the group of participants are in their scanning.

As a side note, it is interesting how the bottom right AOI has so much traffic. When participants enter this slide in the experiment, they have just pressed a click target in the bottom right corner. This means that there is high chance of them focusing in that corner when the new page emerges on the screen. By examining the transitions between menus we can see just how many paths that leads away from that area, suggesting that it is a starting point for some. Additionally, the transitions to this area are very few while the transitions going away from it are many.

As discussed earlier in section 4.1.1.2 | Prior Knowledge, data indicated that the age group 50-59 had less prior knowledge than the age group of 60-69. In the following we will investigate whether or not that has had an effect on their behavior in the areas of interest. The first two screenshots will present the fixations for both groups, starting with the group of 50-59 year olds which consists of eight participants.



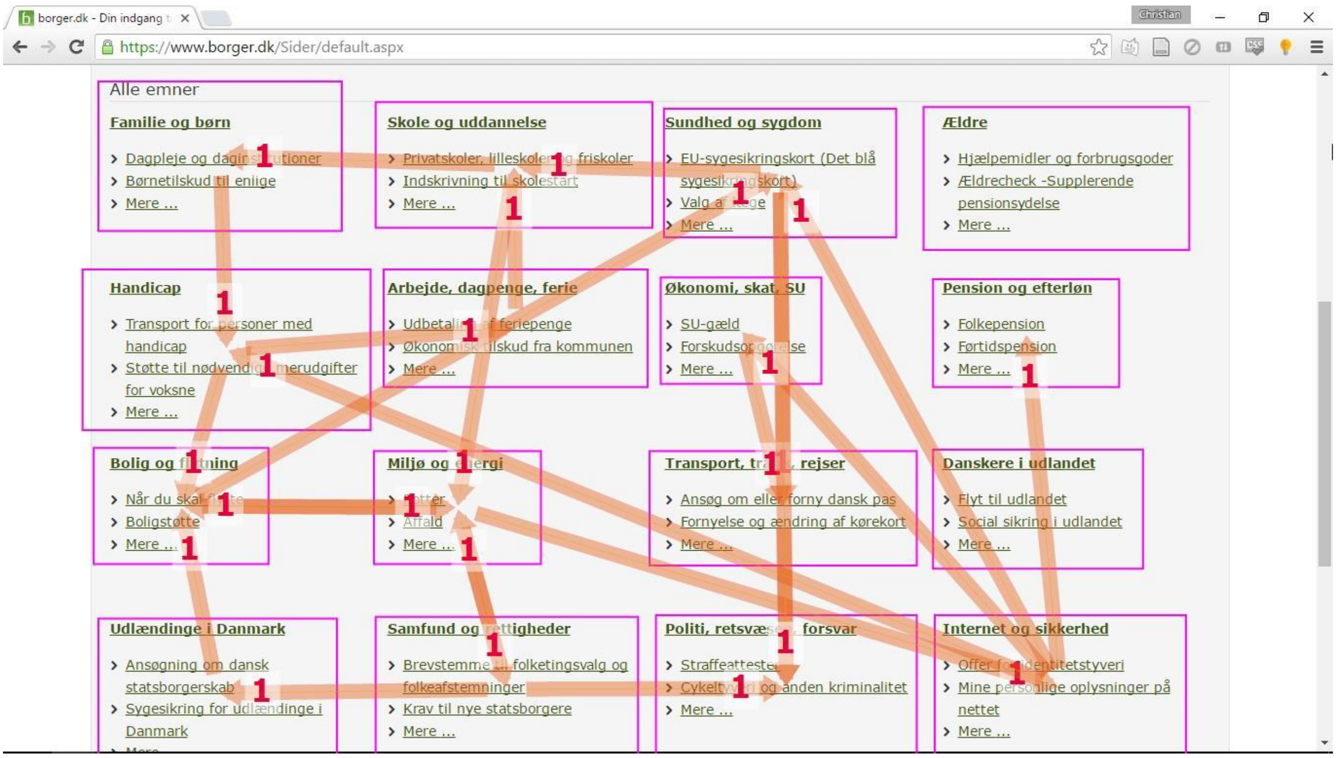
Picture 22 – Fixations 50-59 year olds; 8 participants



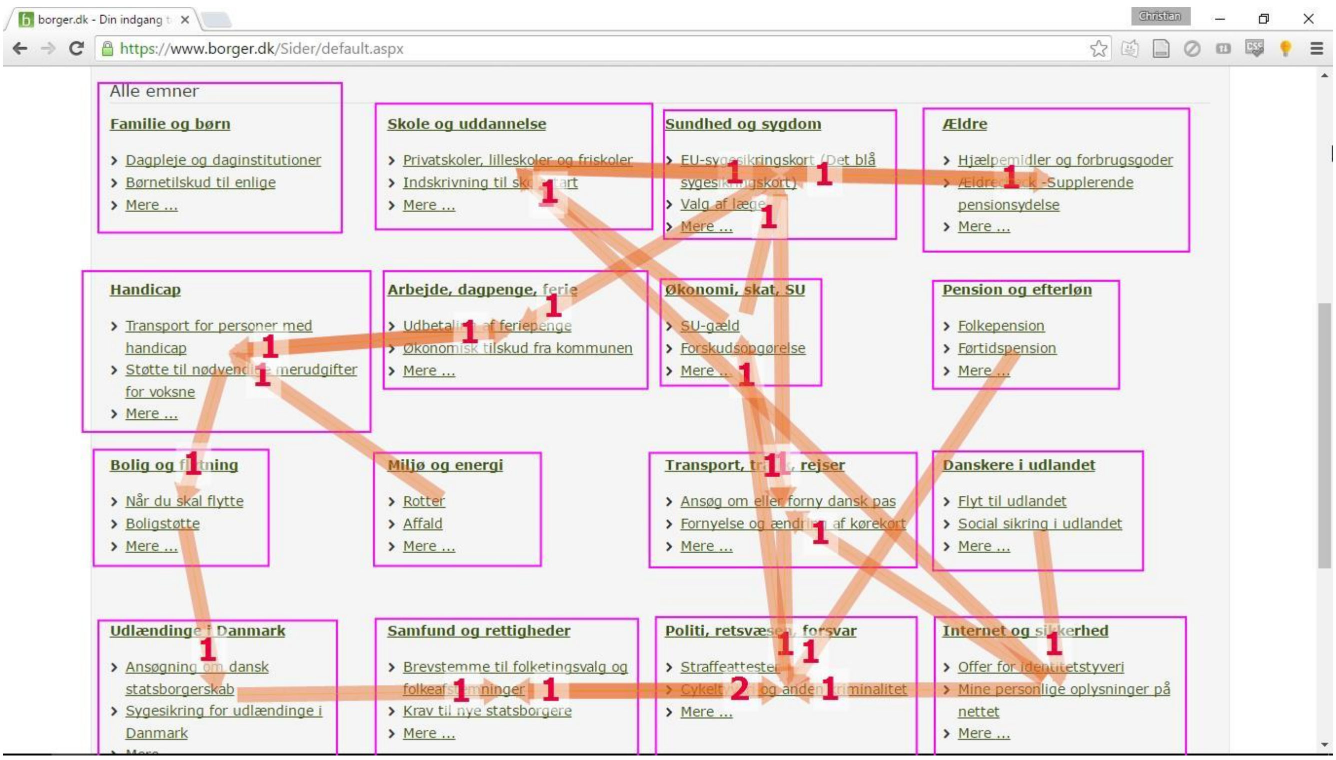
Picture 23 – 60-69 year olds; 14 participants

The amount of fixations themselves are not what is important as the younger and older group have different sizes and thus different amounts of fixations. However, the placement of these fixations are what to notice here. Despite the differences in group sizes and the number of participants, the data shows us that each menu has at least one fixation for both groups and in some cases the amount of fixations are identical. The most noticeable difference lies in top right corner as it shows a large difference between the two groups. There is the possibility of the older group, with more prior knowledge, found this particular menu interesting. It should be mentioned that the menu is related to elderly and thus relevant to both the 50-59 and 60-69 year olds. The prior knowledge possessed by the one group might have caused them to view this menu, it could have been a menu they are used to finding on their own time.

To examine if prior knowledge has affected these two groups in their transitions between menus the following screenshots will show just that; as before the first is of the 50-59 group.



Picture 24 – 50-59 year olds transitions

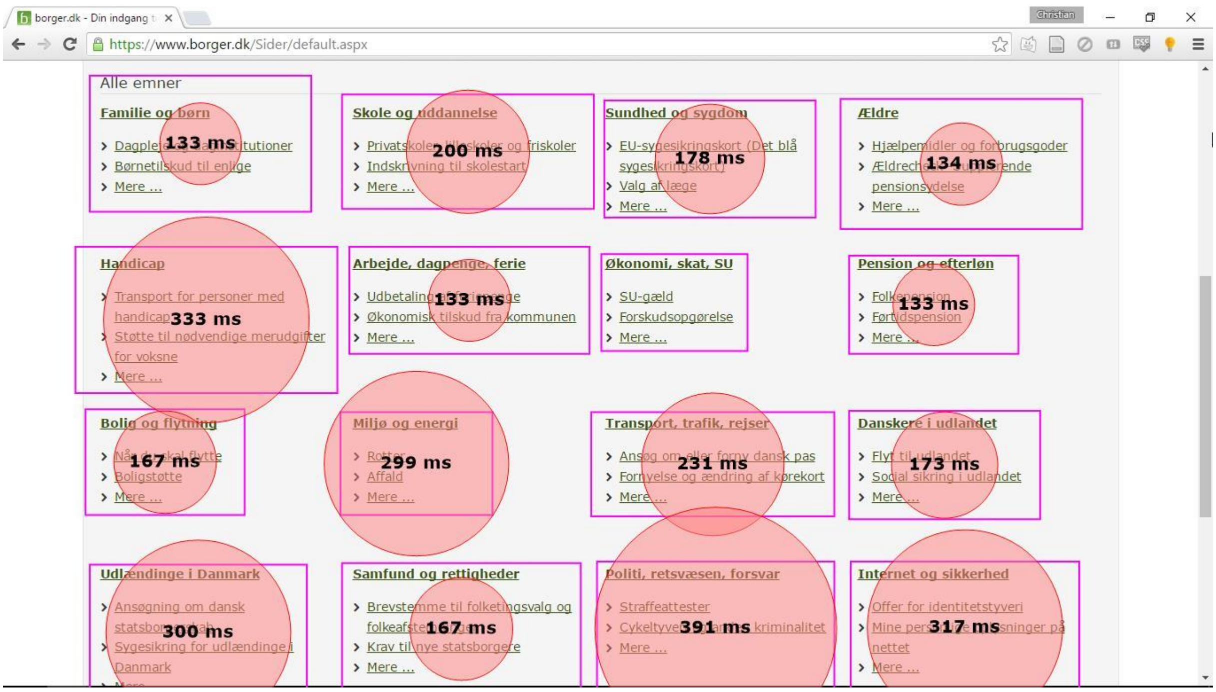


2875

Picture 25 – 60-69 year olds transitions

Comparing the two groups transitions there is something of a pattern to notice. The group with the least prior knowledge seems to have scanned the menus more thoroughly displayed by the transitions on the left, whereas the group with more prior knowledge has more transitions going directly to the target menu. It is also interesting that the first group has less participants, but about the same amount of transitions which supports the claim of them scanning more menus before finding the answer. In the second screenshot the only transition with more than one entry is found as well.

In the following we will compare the results for the elder participants (53+) with the young group of participants (20-24). We will display the average fixation times and transitions for the younger group.



Picture 26 – Average fixations young participants

Compared to the average fixation time for the elderly group, it is noticeable how the AOIs with lower values are much lower for the young participants, indicating that they have scanned and found menus to be irrelevant faster than the old group. Furthermore, the AOIs with the higher values are also different from the elderly group as the average fixation of the five most viewed is much more similar. This could indicate a pattern of the young users choosing some menus to focus on while not paying much attention to others. Interestingly, the target AOI for this slide in Task 1 has the exact same average fixation time for all participants in our test, both young and old.

4.1.2.5 | Scan paths

In the following we will go through age and prior knowledge to see if any differences exist within the elderly group, which can be attributed to either of the demographic subjects. To do this we will look at the Levenshtein value between the groups to compare their similarities, as described in chapter 2.2 | Eye tracking on page 30. We will illustrate our findings with visualizations from OGAMA.

A limitation with OGAMA is the limited options to compare groups, as you have to manually write a group name by each participant, to be able to compare them. We have ended up with an individual experiment database for age and prior knowledge. This way we can switch between them, without having to change the groups each time.

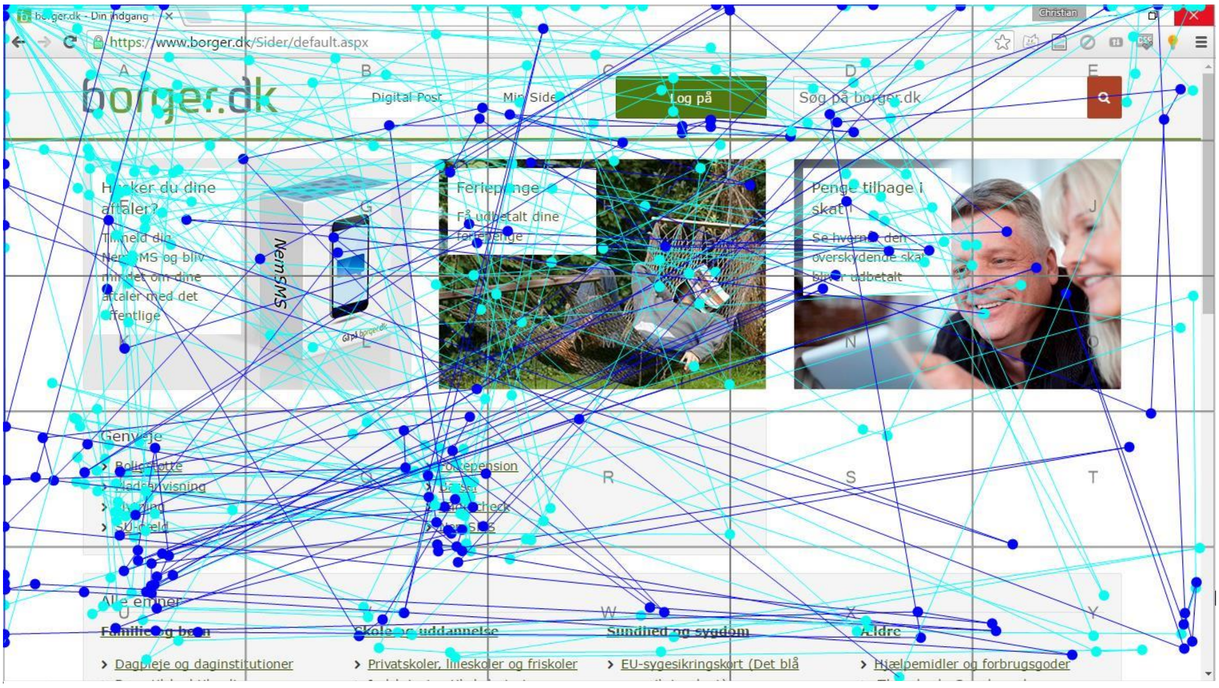
We have chosen six slides on which we have good data from more than 10 different participants. These slides can be seen in Appendix 4 and will be referred to by the number from one to six, they have been assigned in the appendix.

To compare the similarities in the scan paths, we use Levenshtein’s value based on fixations in the AOIs and in a 5x5 grid.

	Above	Below
Above	13.66191	13.87416
Below	13.87416	14.14688

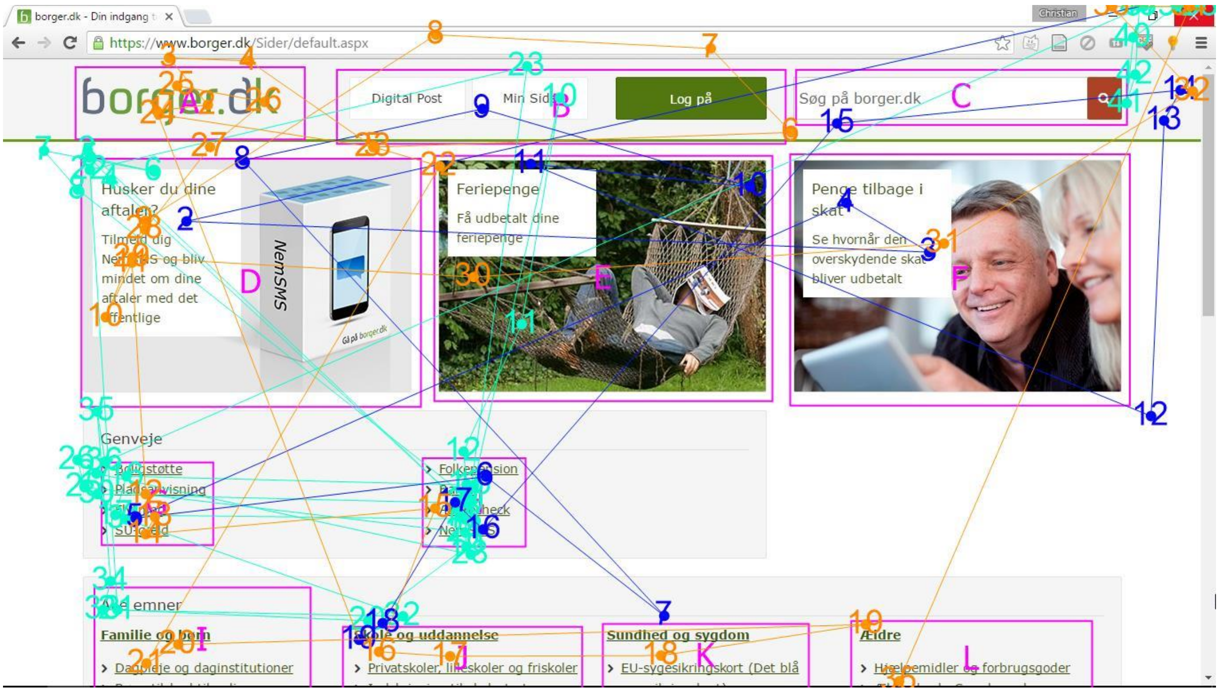
Table 15 – Slide 1 scan paths

By looking at the age of the participants and the scan paths taken, we see that the scan path sequences are very similar, when comparing differences between the age groups and internally in each group. This can be seen in the example from slide one in Table 15. This table show the similarity between the two groups and internally in each, in percentages. The calculation is based on the participants' sequences of fixations in their scan path and compared on 5x5 grid.



Picture 28 – Scan paths 5x5 grid

In Picture 28 we can see the scan paths of the two groups and the 5x5 grid. The light blue scan paths are for the participants above the age median and the dark blue is for those below. This looks confusing, so to get a clearer picture of the similarities or differences between participants we can look at Picture 29 for participants A6, A24 and A27.

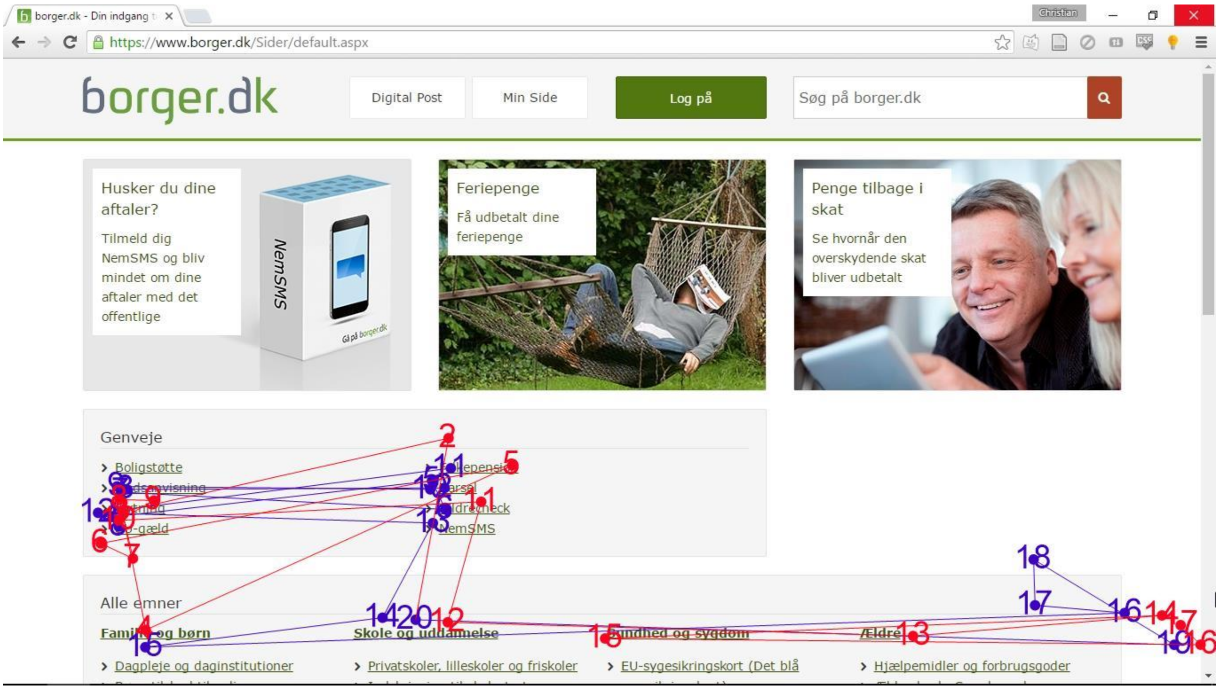


Picture 29 – Scan paths three participants AOI

The similarity between these three participants are as show in Table 16. The similarity between them ranges from 17% between A27 and A24 to 27% between A27 and A6. This means that A6 and A27 are quite similar, at least compared to A27 and A24. We can compare this to how the scan paths of A14 and A28 are similar on slide 4. Their similarity is 55%, which must be regarded as high. The similarity between their fixations are even higher at 75%. It is clear from Picture 30 how similar they really are.

	A24	A27	A6
A24	100%	17%	23%
A27	17%	100%	27%
A6	23%	27%	100%

Table 16 – Similarities three participants



Picture 30 – Scan paths two users

The similarities on the six slides for the participants with useful data, can be seen in the table below. It displays the similarity in the participants' fixations and scan paths in percentages, when using a 5x5 grid.

	Slides					
Similarity	1	2	3	4	5	6
Fixations	44.1871	34.54413	35.43433	37.20134	30.07408	33.43538
Scan paths	12.79582	14.98595	16.96399	17.39849	10.33645	14.60695

Table 17 – Similarities 5x5 grid

The median for fixations is 34.99% with a standard deviation of 4.74 percentage points, which shows us that about a third of all fixations are similar for the participants. This shows us the participants are looking at the same areas of the website, but when comparing it to the scan path similarity we see another pattern. The scan paths have a similarity median of 14.80% with a standard deviation of 2.64 percentage points and is considerably less than the fixations. With a scan path similarity of only 14.80%, it is clear that the participants have vastly different ways of looking for information on the website. To get a better understanding of the differences we can look at the same statistics, but using the AOI's to compare fixations and the scan paths. These can be seen in Table 18 below.

	Slides					
Similarity	1	2	3	4	5	6
Fixations	52.23777	42.91219	42.69231	40.11728	48.88888	50.43956
Scan paths	31.27207	21.51058	25.80955	26.7003	25.64021	25.6316

Table 18 – Similarities AOI

The fixation similarity when comparing using AOI are much higher opposed to a comparison based on the 5x5 grid. The median for fixations are 45.90% with a standard deviation of 4.93 percentage points. The standard deviation is slightly higher, but only 0.19 percentage point. The fixation similarity is 10.91 percentage point higher. This is because the AOIs cover all the text blocks, images, navigation, links and so on, which ensures all the participants looks at are used for calculating the similarities. Using the 5x5 or any other grid size, might have the participants looking at the same navigation option, but parts of it are in one grid, which means the order might differ even though they are actually looking at the same thing. It is clear the participants fixate on similar parts of the website. Looking at the scan paths we see that the median scan path similarity is up 10.93 percentage points to 25.73% with a standard deviation of 3.12 percentage points. This shows that the participants search behavior are similar in where they look, but their scan paths differentiate quite a lot.

To understand this better, we will look at the similarity within the older age group, to see if there are any differences in the age group, based on age and then we will compare the results to our young participants. In the table below we have compared the similarity between those above the age median and those below. The table shows the similarity, when using the AOIs.

	Slides					
Similarity	1	2	3	4	5	6
Fixations	51.81184	40.46329	41.6667	42.62153	40.0463	53.10416
Scan paths	31.97686	21.19544	24.4609	23.33982	14.30912	23.52544

Table 19 – Similarities for elderly above and below age median AOI

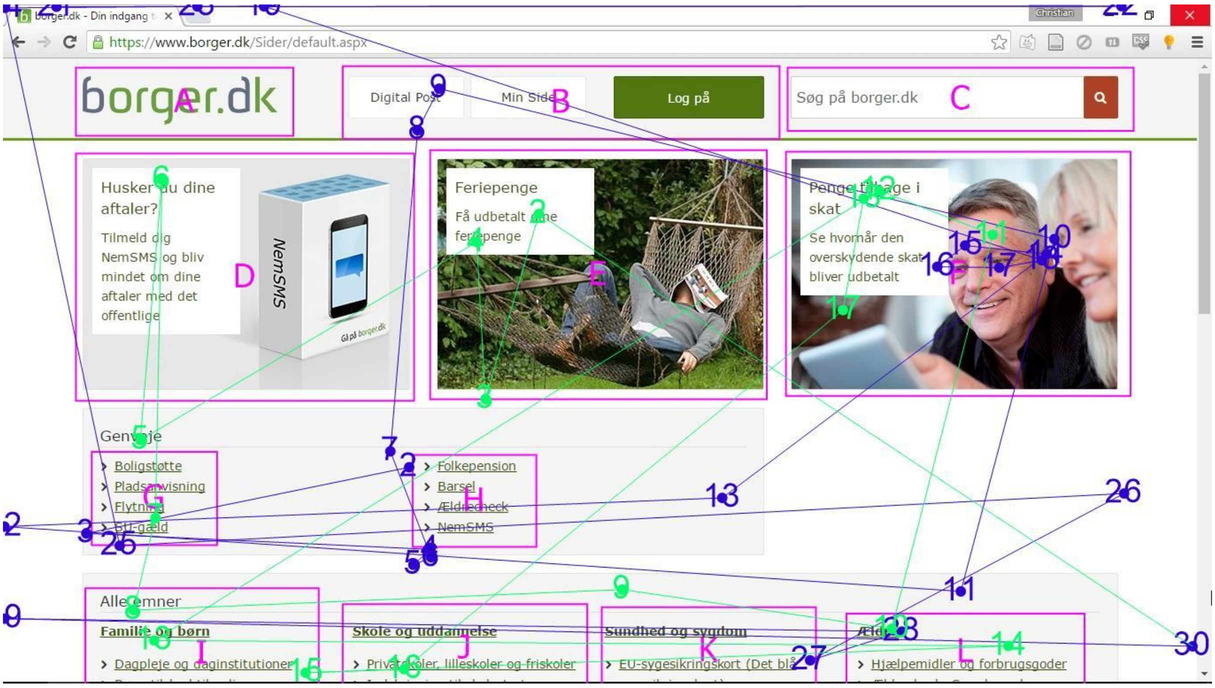
Comparing the differences between the similarity for the entire age group and the similarity of the age group split in two, show that there is little difference between the two. The similarity median for fixations only drops by 3.76 percentage points and the same number for scan paths are only 2.29 percentage points lower, at 42.14% and 23.43% respectively. The standard deviations are increased by 0.96 and 2.57, indicating that it is more varied. The differences are so minor that we do not believe them to be significant enough to indicate a difference caused by age. Now we compare the entire elderly age group to the young age group, to see if any differences that can support the notion of immigrants and natives exist between them.

	Slides					
Similarity	1	2	3	4	5	6
Fixations	42.73217	34.04	44.28571	41.35556	36.25	32.08334
Scan paths	21.42884	18.26151	29.68434	26.92386	21.30952	18.23873

Table 20 – Similarities Elderly and young AOI

The median for fixations when comparing the two groups are at 38.80% with a standard deviation of 4.93 percentage points. This is 7.1 percentage points lower than the similarity within the elderly age group. This proves that there are differences between the two groups in our study, but it is a minor difference, which we do not see as significant enough to indicate the notion of immigrants and natives are true. For the scan paths we also find a small difference. The median for scan path similarity between the two groups are 21.37%, 4.35 percentage points lower than the similarity for the elderly age group. This, again, does not indicate a major difference between the two age groups as whole, but rather that the individual participants look for information in different ways. Picture 30 shows two very similar scan paths for two elderly participants. These two participants have looked at the front page of Task 2 in in more or less the same manner, with a similarity of 55%.

Comparing those two scan paths to the two in Picture 31, which are vastly different, with a similarity of only 11%, we can see completely different search behaviors. The first example is very focused on the shortcuts and the few visible navigation options near the fold and ending om looking at the button to move down the page. These users seem to understand how the website works and know where to look for the information. In the second we see two participants searching all over the front page for the information, before moving down the site.



Picture 31 – Two different scan paths

The differences cannot, from our data, be explained by age differences. Below we will discuss whether prior knowledge might influence the search behavior.

From the table below we can see that the median for similarity between the top and lower elderly participants scan paths, based on prior knowledge, is 25.35%. When compared to the total similarity for the entire age group of 25.73%, we can see that the difference is only 0.38 percentage point. This difference is close to non-existent and indicates that prior knowledge, internally in the elderly age group, does not influence their search behavior.

	Slides						**Median**	**Standard deviation**
Similarity	1	2	3	4	5	6		
Fixations	47.81941	42.49612	43.70834	37.48264	53.17461	50.7963	45.76388	5.79342
Scan paths	31.43503	25.40303	24.40774	25.6146	23.26531	25.30182	25.35243	2.845738

Table 21 – Similarities for elderly above and below prior knowledge median AOI

In the following we compare the participants above the prior knowledge median with the young age group. Afterwards we do the same, this time for the elderly participants below the prior knowledge median. The first comparison is between the elderly above the prior knowledge median.

	Slides						Median	Standard deviation
Similarity	1	2	3	4	5	6		
Fixations	41.56746	31.79654	26.25	36.33333	33.33333	40.33333	34.83333	5.708438
Scan paths	20.96627	18.52778	18.91667	23.88258	15.15873	24.4	19.94147	3.509284

Table 22 – Similarities for elderly over prior knowledge median and the young age group

In this table we see that the young and elderly with a prior knowledge above the median have a similarity of 19.94%, which is 1.43 percentage point less than the similarity for the young and old based on gender at 21.37%. The difference is again minor and does not indicate any differences based on prior knowledge. Next we look at the similarity based on the elderly below the prior knowledge median and the young.

	Slides						Median	Standard deviation
Similarity	1	2	3	4	5	6		
Fixations	43,6057	34,60087	26,7	42,61111	27,14286	43,24075	38,60599	8,003448
Scan paths	21,77576	18,19494	17,36818	27,68419	17,85714	26,97954	19,98535	4,679491

Table 23 – Similarities for elderly below prior knowledge median and the young age group

With a similarity median of 19.99% it is almost the same as the similarity for the elderly above the prior knowledge median and does not show any tendency for prior knowledge affecting the participants search behavior.

We have found by comparing scan paths that the differences based on age are insignificant and does not point to the notion of natives and immigrants being true. It does not show any tendency towards the elderly and young view a website differently or at least not in a significant manner. The same is true when using prior knowledge for comparing the groups. We have not found any differences internally in the elderly age group either and any differences seems to be based on individual differences that cannot be attributed to age or prior knowledge.

4.2 | Summary

In the following we will summarize the key findings we have collected in this study and link these findings to the research questions we have based our study on.

By conducting the questionnaire, we collected data on participants' demographics and their prior knowledge. We were able to calculate median scores to help with understanding the elderly group and identify characteristics of the group as a whole. We were then able to compare this data to the young participants. We found that the elderly participants had a lower prior knowledge than the young participants.

By conducting the eye tracking study, we identified that age did not affect which click paths a user would take, making the two groups have the same number of clicks to complete. However, though we did find that though click paths were similar their completion times were not, as the young group of participants were clearly faster. When examining the errors committed, the differences were few in the amount of errors per participant on average. For the participants who made errors, they were often numerous and in single tasks.

The age groups behavior in our areas of interest was similar as their average fixation times were very similar. Investigating their scan paths proved that age, prior knowledge and gender did not have an effect on how these participants would navigate the slides in the test.

5 | Conclusion

In existing research, we have found that age cannot be used as a grouping mechanism on its own, but can be used as a framework for the groups, which then includes more aspects. We have found that the users' prior knowledge and the demographics can be important factors. These factors have thus been included in our test and in our comparisons.

It is possible to test how users' navigate through a websites information architecture by conducting eye tracking tests using appropriate hardware and software and following a prepared test plan. However, there are certain limitations in regards to eye tracking on elderly people that make it more difficult compared to testing young people. These include the larger portion of elderly who uses glasses and, to some extent, physical aspects such as saggy eyelids. Despite these limitations, we were able to collect data from more than 30 test participants and with most of it usable.

Data showed that the differences between age groups were few and the similarities many. The data collected in this study determined that both groups use the same click paths and thus the same clicks to complete. The amount of errors made by the two groups were similar. The similarity of scan paths was similar for both the groups as well. The differences between the young and elderly were not larger than the differences within the larger elderly group. Only one of our measures showed a difference between the two groups and that was the time to complete. This is in line with previous studies. Our study indicates that the differences in time to complete is caused by the older users spending more time on verification strategies, to ensure they pick the correct option.

We have not found any evidence, which supports the notion of digital natives and digital immigrants, except the differences in time to complete. The elderly age group have a lower prior knowledge, but manages to complete all the tasks and without making more errors than the young group. The elderly group spend more time completing the task, but it was not more difficult for them, when compared to the young age group. Age and prior knowledge as factors did not affect the participant's ability to complete the tasks, how many errors they made or how they looked for specific information in the information architecture of borger.dk.

For these reasons, we have not found the notion of digital natives and digital immigrants to be true in the context of this thesis.

6 | List of Literature

Al Maqbali, Hilal, Falk Scholer, James A. Thom, and Mingfang Wu. 2013. "Using Eye Tracking for Evaluating Web Search Interfaces." In *Proceedings of the 18th Australasian Document Computing Symposium*, 2–9. ACM. http://dl.acm.org/citation.cfm?id=2537734.2537747.

Artis, Sharnnia, and Brian M. Kleiner. 2006. "The Effects of Age and the Design of Web-Based Training on Computer Task Performance." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50 (2): 185–89. doi:10.1177/154193120605000204.

Atwater, Mary M. 1996. "Social Constructivism: Infusion into the Multicultural Science Education Research Agenda." *Journal of Research in Science Teaching* 33 (8): 821–837.

Beier, Margaret E., and Phillip L. Ackerman. 2005. "Age, Ability, and the Role of Prior Knowledge on the Acquisition of New Domain Knowledge: Promising Results in a Real-World Learning Environment." *Psychology and Aging* 20 (2): 341–55. doi:10.1037/0882-7974.20.2.341.

Bennett, Sue, Karl Maton, and Lisa Kervin. 2008. "The 'Digital Natives' Debate: A Critical Review of the Evidence." *British Journal of Educational Technology* 39 (5): 775–786. doi:10.1111/j.1467-8535.2007.00793.x.

Berner, Jessica, Mikael Rennemark, Claes Jogréus, Peter Anderberg, Anders Sköldunger, Maria Wahlberg, Sölve Elmståhl, and Johan Berglund. 2015. "Factors Influencing Internet Usage in Older Adults (65 Years and above) Living in Rural and Urban Sweden." *Health Informatics Journal* 21 (3): 237–49. doi:10.1177/1460458214521226.

Billestrup, Jane, and Jan Stage. 2014. "E-Government and the Digital Agenda for Europe." In *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments*, edited by Aaron Marcus, 71–80. Lecture Notes in Computer Science 8518. Springer International Publishing. http://link.springer.com/chapter/10.1007/978-3-319-07626-3_7.

Biterman, Noemi, Eyal Lerner, and Haim Bitterman. 2007. "E- Health System for Coagulation Function Management by Elderly People." In *Universal Access in Human-Computer Interaction. Applications and Services*, edited by Constantine Stephanidis, 853–59. Lecture Notes in Computer Science 4556. Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-540-73283-9_92.

Black, Rex. 2009. *Managing the Testing Process: Practical Tools and Techniques for Managing Hardware and Software Testing*. 3 edition. Indianapolis, MN: Wiley.

Business Wire. 2010. "Recent ICMI and inContact Survey Reveals Contact Centers Spending More on Self-Service Solutions, but Not Realizing Forecasted Results." *Business Wire*.

Busjahn, Teresa, Carsten Schulte, Bonita Sharif, Andrew Begel, Michael Hansen, Roman Bednarik, Paul Orlov, et al. 2014. "Eye Tracking in Computing Education." In

Proceedings of the Tenth Annual Conference on International Computing Education Research, 3–10. ACM. http://dl.acm.org/citation.cfm?id=2632344.

Case, Donald Owen. 2012. *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior*. Emerald Group Publishing.

Chadwick-Dias, Ann, Michelle McNulty, and Tom Tullis. 2003. "Web Usability and Age: How Design Changes Can Improve Performance." In *Proceedings of the 2003 Conference on Universal Usability*, 30–37. CUU '03. New York, NY, USA: ACM. doi:10.1145/957205.957212.

Crabb, Michael, and Vicki Hanson. 2014. "Age, Technology Usage, and Cognitive Characteristics in Relation to Perceived Disorientation and Reported Website Ease of Use." In , 193–200. ACM SIGACCESS Conference on Computers and Accessibility. ACM. doi:10.1145/2661334.2661356.

Czaja, Sara J., Joseph Sharit, Raymond Ownby, David L. Roth, and Sankaran Nair. 2001. "Examining Age Differences in Performance of a Complex Information Search and Retrieval Task." *Psychology and Aging* 16 (4): 564–79. doi:10.1037/0882-7974.16.4.564.

"Digital Post." 2016. Accessed May 9. http://www.digst.dk/loesninger-og-infrastruktur/digital-post.

Dillon, Andrew. 2002. "Information Architecture inJASIST: Just Where Did We Come From?" *Journal of the American Society for Information Science and Technology* 53 (10): 821–23. doi:10.1002/asi.10090.

Dochy, Filip, Mien Segers, and Michelle M. Buehl. 1999. "The Relation Between Assessment Practices and Outcomes of Studies: The Case of Research on Prior Knowledge." *Review of Educational Research* 69 (2): 145–86. doi:10.3102/00346543069002145.

Duchowski, Andrew T. 2002. "A Breadth-First Survey of Eye-Tracking Applications." *Behavior Research Methods, Instruments, & Computers* 34 (4): 455–70. doi:10.3758/BF03195475.

Dunne, Jennifer E. 2002. "Information Seeking and Use by Battered Women: A 'person-in-Progressive-Situations' Approach." *Library & Information Science Research* 24 (4): 343–355.

Dwivedi, Yogesh K., and Michael D. Williams. 2008. "Demographic Influence on UK Citizens' E-Government Adoption." *Electronic Government, an International Journal* 5 (3): 261–74. doi:10.1504/EG.2008.018874.

Ehmke, Claudia, and Stephanie Wilson. 2007. "Identifying Web Usability Problems from Eye-Tracking Data." In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but Not as We Know It-Volume 1*, 119–128. British Computer Society. http://dl.acm.org/citation.cfm?id=1531311.

"Electronic Government — eGov Conference." 2016. Accessed May 17. http://www.egov-conference.org/glossary/electronic-government.

Etcheverry, Isabelle, Thierry Baccino, Patrice Terrier, Jean-Claude Marquié, and Mustapha Mojahid. 2012a. "Age Differences in Information Finding Tasks: Performance and Visual Exploration Strategy with Different Web Page Layouts." *Computers in Human Behavior* 28 (5): 1670–80. doi:10.1016/j.chb.2012.04.005.

Etcheverry, Isabelle, Patrice Terrier, and Jean-Claude Marquié. 2012. "Assessing Web Interaction with Recollection: Age-Related and Task-Related Differences." *Computers in Human Behavior* 28 (1): 11–22. doi:10.1016/j.chb.2011.08.004.

Evernden, Elaine, and Roger Evernden. 2003. *Information First*. 1 edition. Oxford England; Burlington, MA: Routledge.

"Går På Nettet via Computeren." 2016. Accessed May 20. https://www.aeldresagen.dk/viden-og-raadgivning/hverdagsliv/digital-kontakt-til-det-offentlige/fakta/aeldre-bruger-den-baerbare.

Goldberg, Joseph H., and Jonathan I. Helfman. 2010. "Comparing Information Graphics: A Critical Look at Eye Tracking." In *Proceedings of the 3rd BELIV'10 Workshop: BEyond Time and Errors: Novel evaLuation Methods for Information Visualization*, 71–78. BELIV '10. New York, NY, USA: ACM. doi:10.1145/2110192.2110203.

Goldberg, Joseph H., Mark J. Stimson, Marion Lewenstein, Neil Scott, and Anna M. Wichansky. 2002. "Eye Tracking in Web Search Tasks: Design Implications." In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, 51–58. ACM. http://dl.acm.org/citation.cfm?id=507082.

Grood, Derk-Jan de. 2010. *TestGoal: Result-Driven Testing*. Softcover reprint of hardcover 1st ed. 2008 edition. Leiden, The Netherlands: Springer.

Guan, Zhiwei, Shirley Lee, Elisabeth Cuddihy, and Judith Ramey. 2006. "The Validity of the Stimulated Retrospective Think-Aloud Method As Measured by Eye Tracking." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1253–1262. CHI '06. New York, NY, USA: ACM. doi:10.1145/1124772.1124961.

Hailikari, Telle, and others. 2010. "Assessing University Students' Prior Knowledge." https://helda.helsinki.fi/handle/10138/19841.

Hamer, Judah S. 2003. "Coming-out: Gay Males' Information Seeking." *School Libraries Worldwide* 9 (2): 73.

Hertzum, M. 1999. "User Testing in Industry: A Case Study of Laboratory, Workshop, and Field Tests." *User Interfaces for All. Proceedings* GMD--74: 59–72.

Hurtienne, J., and P. M. Langdon. 2009. "Prior Knowledge in Inclusive Design: The Older, the More Intuitive?" In . Cambridge, UK. http://publications.eng.cam.ac.uk/324162/.

Kaikkonen, Anne, Aki Kekäläinen, Mihael Cankar, Titti Kallio, and Anu Kankainen. 2005. "Usability Testing of Mobile Applications: A Comparison Between Laboratory and Field Testing." *J. Usability Studies* 1 (1): 4–16.

Kalbach, James. 2007. *Designing Web Navigation / James Kalbach*. Elektronisk udgave.-1st ed, Elektronisk udgave. Boston, Mass: Safari Books Online. http://svcs.aub.aau.dk/primo/ebooklauncher/index.php?pkg=safari&url=https://login.zorac.aub.aau.dk/login?url=http://proquest.safaribooksonline.com/?uiCode=aalborg&xmlId=9780596528102.

Kalmus, Veronika, Anu Realo, and Andra Siibak. 2011. "Motives for Internet Use and Their Relationships with Personality Traits and Socio-Demographic Factors." *Trames* 15 (65/60): 385–403.

Kurniawan, S.H., P. Zaphiris, and R.D. Ellis. 2002. "Older and Younger Adults' Time Spent and Errors Made in Expandable and Sequential Hierarchies." *Gerontechnology* 2 (2). doi:10.4017/gt.2002.02.02.005.00.

Kurzhals, Kuno, Brian Fisher, Michael Burch, and Daniel Weiskopf. 2014. "Evaluating Visual Analytics with Eye Tracking." In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, 61–69. ACM. http://dl.acm.org/citation.cfm?id=2669560.

Laberge, Jason C., and Charles T. Scialfa. 2005. "Predictors of Web Navigation Performance in a Life Span Sample of Adults." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 47 (2): 289–302. doi:10.1518/0018720054679470.

Lauterbach, Thomas. 2015. *It-anvendelse i befolkningen*. Vol. 2015. Danmarks Statistik. http://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=20737&sid=itbef2015.

Maloney, Krisellen, and Paul J. Bracke. 2004. "Beyond Information Architecture: A Systems Integration Approach to Web-Site Design." *Information Technology and Libraries* 23 (4): 145–52.

"Månedlig Statistik Om Digital Post." 2016. Accessed May 9. http://www.digst.dk/Loesninger-og-infrastruktur/Digital-Post/Digital-Post-statistik.

Mead, Sherry E., Victoria A. Spaulding, Richard A. Sit, Beth Meyer, and Neff Walker. 1997. "Effects of Age and Training on World Wide Web Navigation Strategies." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 41 (1): 152–56. doi:10.1177/107118139704100135.

"Meet Marc Prensky." 2016. *Powerful Motivational Speaker for Education*. Accessed April 12. http://marcprensky.com/about-marc/.

Merleau-Ponty, Maurice, and Colin Smith. 1996. *Phenomenology of Perception*. Motilal Banarsidass Publishe. https://www.google.com/books?hl=en&lr=&id=7mBOMDRl83QC&oi=fnd&pg=PR7&dq=phenomenology+of+perception&ots=5vlKjMisu9&sig=NvYiY7w3CnxLvJKY73dIsI6QSFE.

Meuter, Matthew, Amy Ostrom, Robert Roundtree, and Mary Bitner. 2000. "Self-Service Technologies: Understanding Customer Satisfaction with Technology-Based Service Encounters." *Journal of Marketing* 64 (3): 50.

Meyer, Beth, Richard Sit, Victoria Spaulding, Sherry Mead, and Neff Walker. 1997. "Age Group Differences in World Wide Web Navigation." In , 295–296. Conference on Human Factors in Computing Systems. ACM. doi:10.1145/1120212.1120401.

Mills, Geoffrey E. 2010. *Action Research: A Guide for the Teacher Researcher*. 4th ed. Boston, Mass: Pearson.

Morville, Peter, and Louis Rosenfeld. 2007. *Information Architecture for the World Wide Web*. http://shop.oreilly.com/product/9780596527341.do.

"OGAMA." 2016. Accessed March 28. http://www.ogama.net/.

Paavola, Sami, Lasse Lipponen, and Kai Hakkarainen. 2004. "Models of Innovative Knowledge Communities and Three Metaphors of Learning." *Review of Educational Research* 74 (4): 557–76. doi:10.3102/00346543074004557.

Pernice, Katie, and Jakob Nielsen. 2016. "How to Conduct Eyetracking Studies." Nielsen Norman Group. Accessed March 28. https://media.nngroup.com/media/reports/free/How_to_Conduct_Eyetracking_Studies.pdf.

"Popper: Critical Rationalism | Internet Encyclopedia of Philosophy." 2016. Accessed May 19. http://www.iep.utm.edu/cr-ratio/.

Prensky, Marc. 2001a. "Digital Natives, Digital Immigrants Part 1." *On the Horizon* 9 (5): 1–6. doi:10.1108/10748120110424816.

Prensky, Marc. 2001b. "Digital Natives, Digital Immigrants Part 2: Do They Really Think Differently?" *On the Horizon* 9 (6): 1–6. doi:10.1108/10748120110424843.

"Prior Knowledge | Learning Sciences." 2016. Accessed May 17. https://learningsciences.utexas.edu/teaching/learning/prior-knowledge.

Resmini, Andrea, and Luca Rosati. 2012. "A Brief History of Information Architecture." *Journal of Information Architecture* 3 (2). http://journalofia.org/volume3/issue2/03-resmini/jofia-0302-03-resmini.pdf.

Rubin, Jeffrey, Dana Chisnell, and Jared Spool. 2008. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. 2 edition. Indianapolis, IN: Wiley.

Russell-Rose, Tony, and Tyler Tate. 2013. *Designing the Search Experience: The Information Architecture of Discovery*. 1 edition. Amsterdam: Morgan Kaufmann.

Schade, Amy. 2015. "Pilot Testing: Getting It Right (Before) the First Time." *Nielsen Norman Group*, May. https://www.nngroup.com/articles/pilot-testing/.

Shapiro, Amy M. 2004. "How Including Prior Knowledge As a Subject Variable May Change Outcomes of Learning Research." *American Educational Research Journal* 41 (1): 159–89. doi:10.3102/00028312041001159.

"The Computer - Stephen Hawking." 2016. Accessed May 17. http://www.hawking.org.uk/the-computer.html.

"The Eye Tribe." 2016. Accessed April 18. https://theeyetribe.com/.

Thompson S.H. Teo. 2001. "Demographic and Motivation Variables Associated with Internet Usage Activities." *Internet Research* 11 (2): 125–37. doi:10.1108/10662240110695089.

Trivedi, Munesh Chandra, and Mohammadi Akheela Khanum. 2012. "Role of Context in Usability Evaluations: A Review." *arXiv Preprint arXiv:1204.2138*. http://arxiv.org/abs/1204.2138.

Van Teijlingen, Edwin, and Vanora Hundley. 2002. "The Importance of Pilot Studies." *Nursing Standard* 16 (40): 33–36.

Whittaker, James A. 2012. "The 10-Minute Test Plan." *IEEE Software* 29 (6): 70–77. doi:10.1109/MS.2012.25.

Williamson, Kirsty. 1997. "The Information Needs and Information-Seeking Behaviour of Older Adults: An Australian Study." In *Proceedings of an International Conference on Information Seeking in Context*, 337–350. ISIC '96. London, UK, UK: Taylor Graham Publishing. http://dl.acm.org/citation.cfm?id=267190.267211.

Zhu, Zhiwei, Kikuo Fujimura, and Qiang Ji. 2002. "Real-Time Eye Detection and Tracking under Various Light Conditions." In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, 139–144. ACM. http://dl.acm.org/citation.cfm?id=507100.

Zukowski, Tomasz, and Irwin Brown. 2007. "Examining the Influence of Demographic Factors on Internet Users' Information Privacy Concerns." In *Proceedings of the 2007 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries*, 197–204. SAICSIT '07. New York, NY, USA: ACM. doi:10.1145/1292491.1292514.

7 | Appendices

Appendix 1

Køn:

Sæt venligst kryds ud for det rigtige svar.

Mand ☐

Kvinde ☐

Alder:

______

Uddannelsesniveau:

Sæt venligst kryds ud for det rigtige svar

Folkeskole ☐

Teknisk uddannelse(håndværkerfag) ☐

Gymnasie uddannelse ☐

Videregående uddannelse(Universitetsuddannelse) ☐

Erfaring med internettet:

Sæt venligst et kryds ud for det rigtige svar

Har du nogensinde brugt internettet til at ændre din hjemmeadresse?	Ja ☐	Nej ☐
Har du nogensinde brugt internettet til at forny dit pas?	Ja ☐	Nej ☐
Har du nogensinde brugt internettet til at forny eller ændre dit kørekort?	Ja ☐	Nej ☐
Har du nogensinde brugt internettet til at bestille nyt sundhedskort (sygesikringsbevis)?	Ja ☐	Nej ☐
Har du nogensinde brugt internettet til at vælge ny læge?	Ja ☐	Nej ☐
Har du nogensinde brugt internettet til at søge om offentlig støtte i form af penge?	Ja ☐	Nej ☐

3325

Sæt venligst kryds ud for de hjemmesider du har benyttet før:

Borger.dk	☐
SU.dk	☐
Sundhed.dk	☐
Skat.dk	☐
e-boks.dk	☐

Vær venlig at markere med kryds hvor ofte du udfører følgende handlinger på internettet:

	Sjældnere	2 gange om måneden	3 gange om måneden	2 gange om ugen	4 gange om ugen	Oftere
Benytter netbank						
Er i kontakt med det offentlige						
Bruger en søgemaskine						
Læser nyheder						
Ser film, serier og dokumentarer						
Tjekker din e-boks						
Benytter sociale medier						
Skriver og læser e-mails						
Handler online						
Deltager i online debatter						

	Sjældnere	1-2 gange	3-4 gange	5-7 gange	8-10 gange	Oftere
Vær venlig at markere med kryds hvor mange gange du benytter internettet om dagen						

	Færre timer	3-9 timer	10-19 timer	20-29 timer	30-39 timer	Flere timer
Vær venlig at markere med kryds hvor mange timer du benytter internettet om ugen						

	1	2	3	4	5	6
På en skala fra 1-6, hvor 1 er lidt og 6 er meget, vurder hvor meget tidligere/nuværende jobs har påkrævet dig at bruge computere						

	1	2	3	4	5	6
På en skala fra 1-6, hvor 1 er lidt og 6 er meget, vurder hvor meget du bruger en computer til dine studier						

3330 # Appendix 2

	Gender	Age	Education	Experience	Borger.dk experience	Online actions										Internet usage	Weekly usage	Work use	Study use	Prior knowledge median
						1	2	3	4	5	6	7	8	9	10					
A2	Female	64	Videregående	1	Yes	4	1	6	5	1	4	6	6	2	1	4	3	6		4
A3	Female	64	Teknisk skole	1	Yes	3	1	5	6	5	3	6	6	1	1	3	2	6		3
A4	Female	62	Folkeskole	1	Yes	5	1	6	5	5	4	4	6	1	1	3	3	5		4
A5	Female	68	Folkeskole	3	Yes	4	1	6	6	4	2	1	6	1	1	5	3	6		3.5
A6	Female	57	Folkeskole	1	Yes	3	1	1	1	1	3	1	4	1	1	3	2	6		1
A7	Female	64	Teknisk skole	1	Yes	3	1	4	4	2	3	5	5	1	1	3	2	3		3
A8	Female	63	Teknisk skole	4	Yes	3	1	5	1	1	2	6	6	1	1	2	2	6		2
A9	Female	88	Folkeskole	0	No	4	2	6	2	1	3	1	5	1	1	2	2	1		2
A10	Female	55	Folkeskole	4	Yes	5	1	1	1	6	2	6	6	3	2	3	3	1		3
A11	Male	59	Teknisk skole	0	No	1	1	6	5	3	1	1	6	1	1	2	1	4		1
A12	Female	65	Videregående	0	Yes	6	1	6	6	6	5	6	6	6	1	6	5	6		6
A13	Male	67	Folkeskole	2	No	1	1	6	2	1	6	2	6	1	1	6	6	6		2
A14	Female	53	Videregående	0	No	2	1	6	6	1	3	1	6	1	1	2	2	6	6	2
A15	Male	55	Videregående	3	Yes	4	4	6	6	1	4	6	6	1	2	6	2	6		4
A16	Male	55	Teknisk skole	2	Yes	4	1	1	5	1	4	1	6	1	1	5	2	2		2
A17	Male	67	Teknisk skole	0	Yes	4	1	5	6	1	5	1	3	1	1	3	1	3		2
A18	Female	61	Folkeskole	1	Yes	3	4	5	6	3	6	6	6	1	1	6	2	2		3.5
A19	Male	68	Teknisk skole	0	Yes	5	1	3	1	1	4	1	1	1	1	2	1	1		1
A20	Female	83	Teknisk skole	0	Yes	4		6		3	4		6	1		4	3	1		3.5
A21	Female	83	Gymnasie	0	Yes	3	1	5	3	1	2	1	5	1	1	3	3	6		2.5
A22	Male	57	Teknisk skole	0	Yes	1	1	6	6		1	1	6	2	1	4	3	3		2
A23	Male	62	Teknisk skole	0	Yes	5	1	4	1	1	4	1	4	1	1	1	1	1		1
A24	Female	58	Teknisk skole	2	No	5	1	6	6	1	3	1	6	2	1	3	3	6		3
A25	Female	60	Folkeskole	2	Yes	2	1	4	4	1	4	1	4	1	1	3	2	4		2
A26	Male	73	Folkeskole	0	Yes	6	3	6	6	1	3	1	6	1	1	4	3	5		3
A27	Male	65	Teknisk skole	0	Yes	4	1	6	6	4	3	6	6	2	1	6	3	6		4
A28	Female	63	Folkeskole	0	Yes	3	1	3	1	1	3	1	1	1	1	1	2	6		1
Median		63		1		4	1	6	5	1	3	1	6	1	1	3	2	5	6	2.5
Standard deviation		8.7		1		1	1	2	2	2	1	2	1	1	0	2	1	2		1.213645

Appendix 3

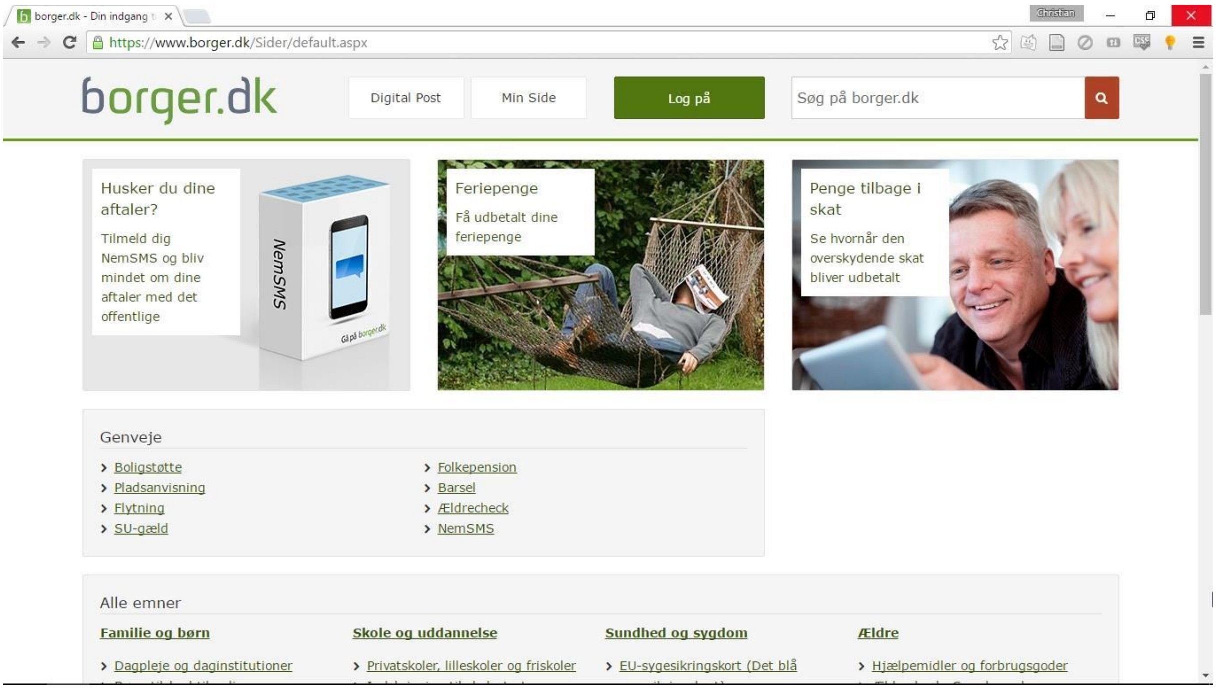
	0-29 sec			30-59 sec			1:00-1:29 min		
	Participants	Age Median	Prior Knowledge Median	Participants	Age Median	Prior Knowledge Median	Participants	Age Median	Prior Knowledge Median
Task 1	3	55	2	9	63	3	7	62	3
Task 2	9	62	2	13	63	3	3	67	2
Task 3	16	63	3	7	61	3	3	83	2

	1:29-1:59 min			2:00-2:29 min			2:30+ min		
	Participants	Age Median	Prior Knowledge Median	Participants	Age Median	Prior Knowledge Median	Participants	Age Median	Prior Knowledge Median
Task 1	3	67	3	3	83	2	1	68	1
Task 2							1	68	1
Task 3									

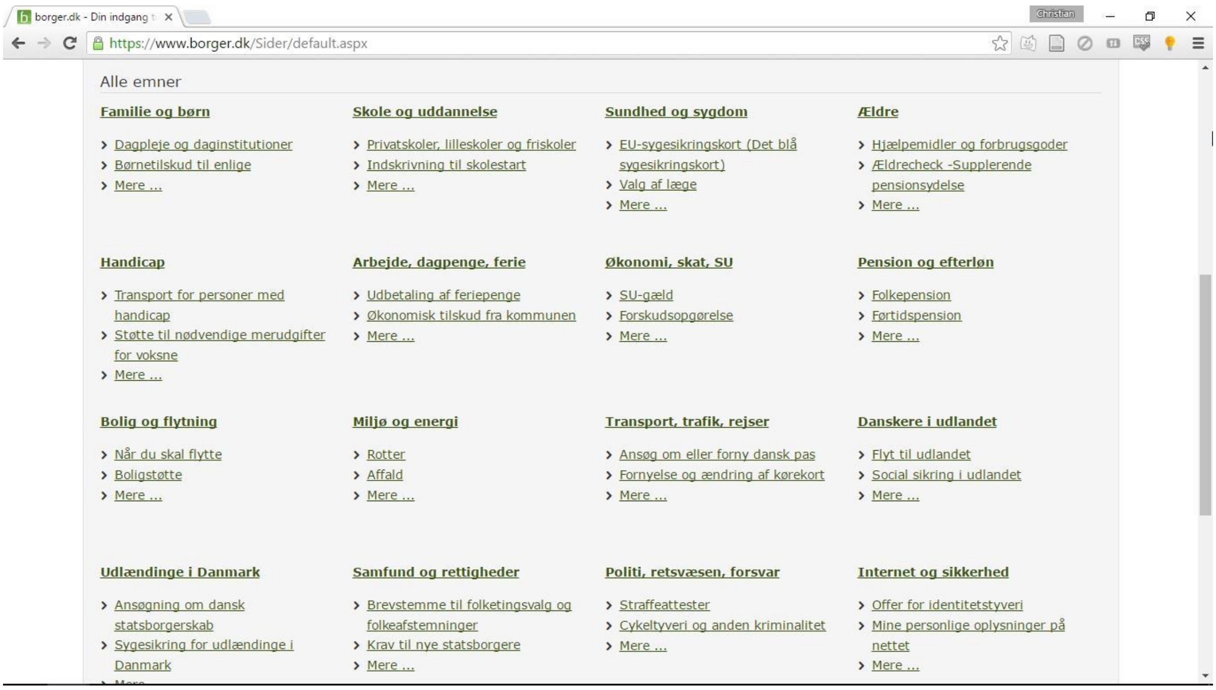
3335

Appendix 4

Slide 1

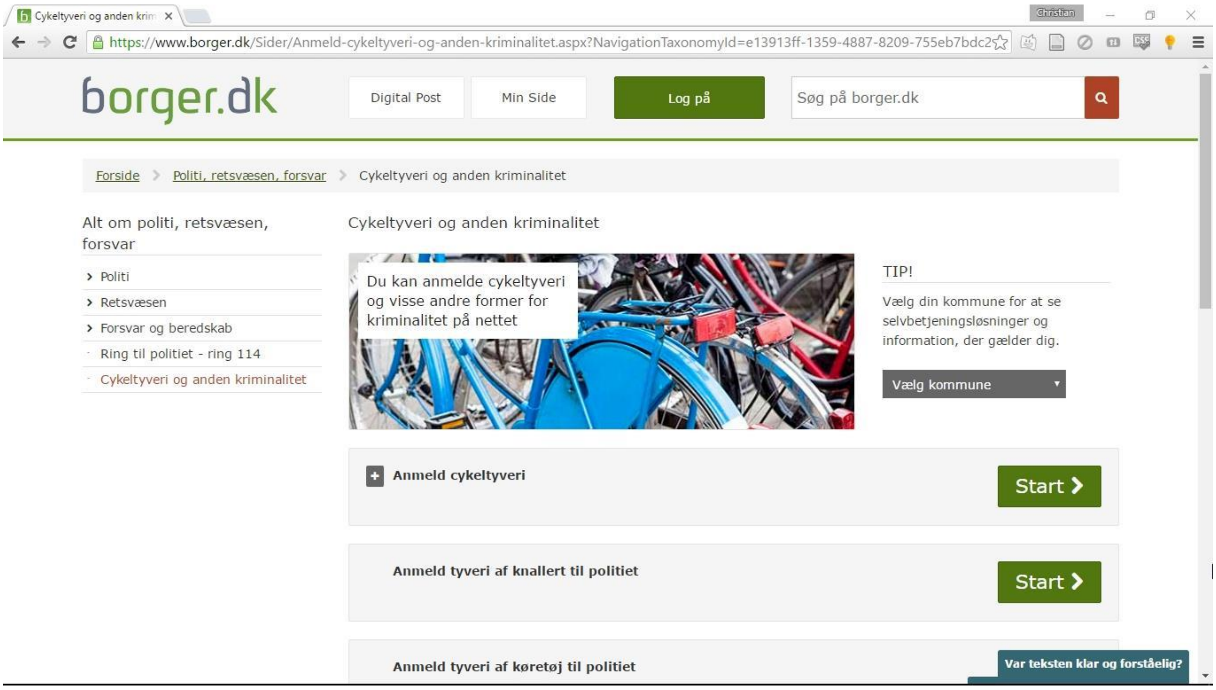


Slide 2

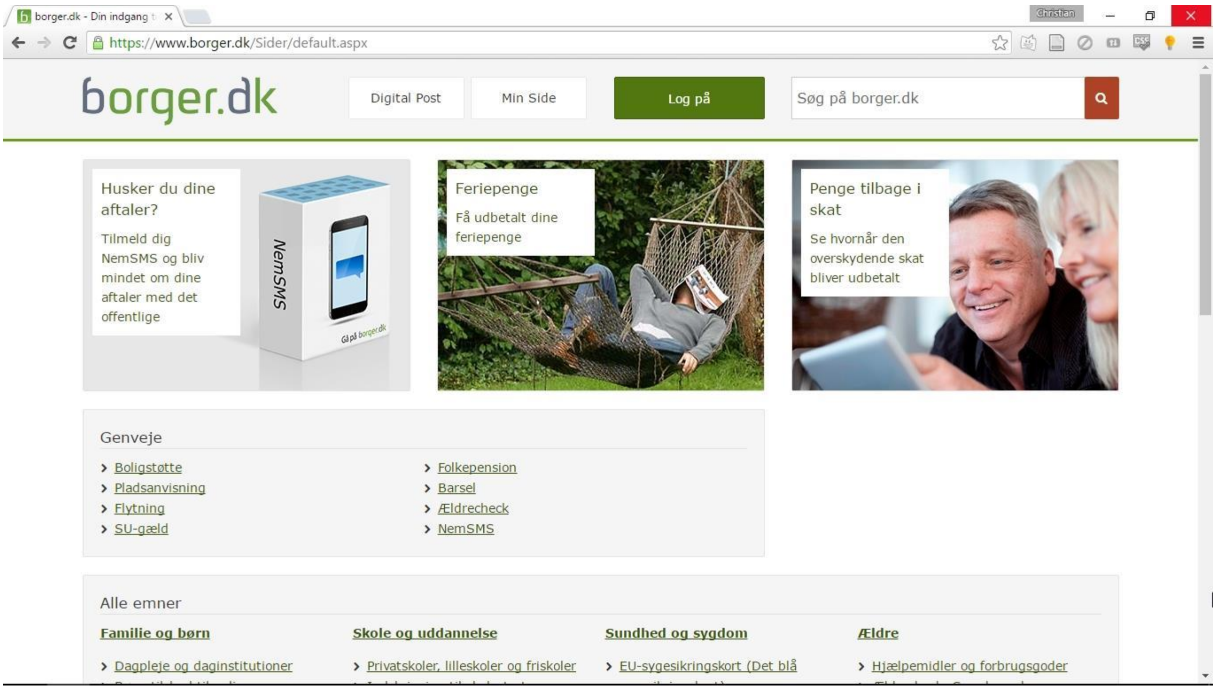


3340

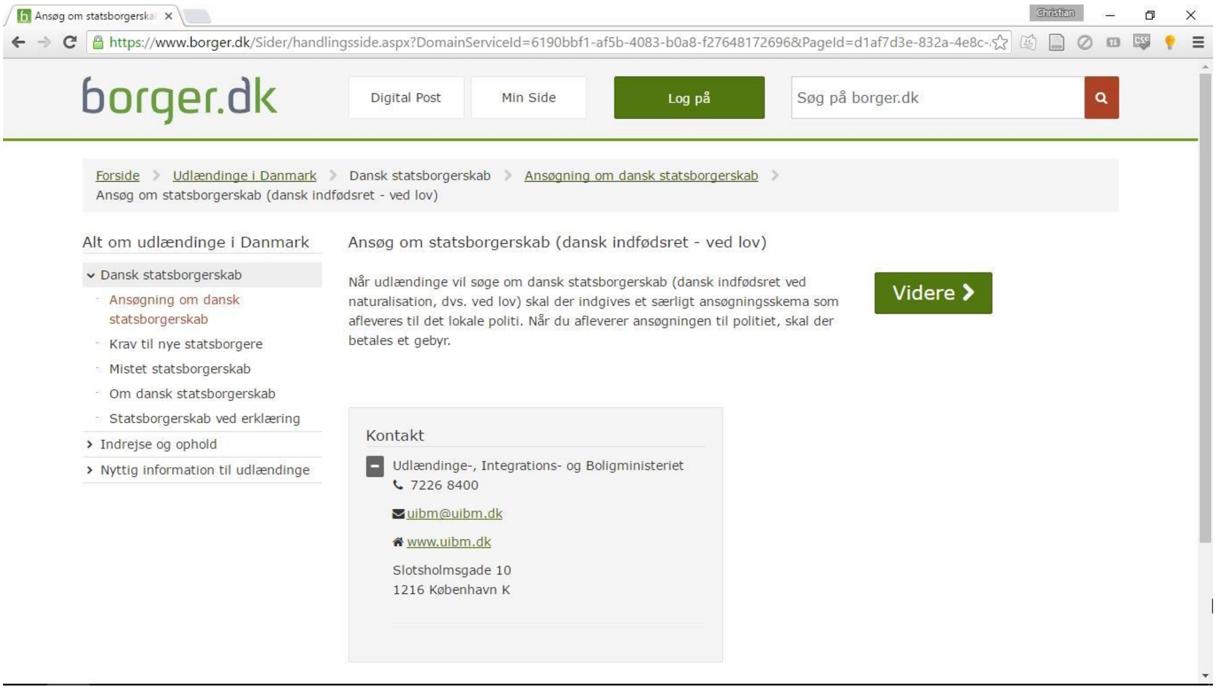
Slide 3



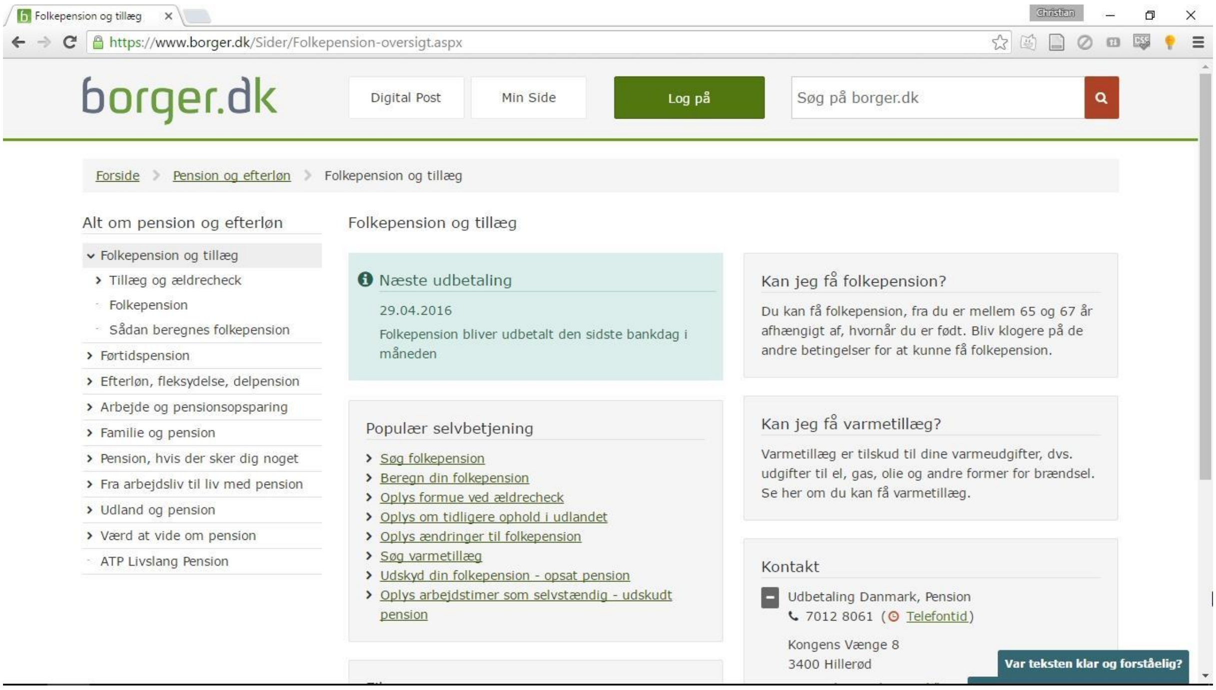
Slide 4



Slide 5



Slide 6



3355