Mini-project 2: Biometrics in UX research

Date of Report: 25.03.23

Date of Test: 10.03.23

Location of Test: University of Agder

Prepared by: Edvin Sæverås

Phone Number: +47 940 82 140

Email: edvins@student.uia.no

Prepared by: Georgi Ivanov

Phone Number: +47 458 11 767

Email: georgi18@student.uia.no

Table of content

1. Executive Summary

This report goes into the usability evaluation of the website https://www.graviditetskollen.nu/.
Graviditetskollen offers all products needed for pregnancy - Before, During, and After. This also includes a large assortment of products for babies and children. A total of 3 participants were recruited, and biometrics data was collected using iMotions software and Biometrics hardware. The goal of the test was to uncover and understand the potential usability shortcomings of the website.
Overall, the tasks given were performed successfully, and all users could complete all the different tasks. During the usability test, a consensus was formed regarding the website's outdated design, and a non-critical error was uncovered within the checkout.

2. State-of-the-art

Biometrics and UX evaluation are two distinct fields that can be fused to enhance the user experience of a product or service. Biometrics refers to the measurement and analysis of physical or behavioral characteristics of individuals, such as fingerprints, facial recognition, eye-tracking, EEG (electroencephalogram), and GSR (galvanic skin response)**.** On the other hand, UX evaluation involves assessing and analyzing the usability, accessibility, and satisfaction of a product or service from the user's perspective. By integrating biometric data with traditional UX evaluation methods, designers and developers can gain deeper insights into user behavior, emotions, and cognitive processes, allowing them to create more user-friendly, engaging, and effective products. When combining the two fields, the data gathered can function as quality assurance in the analysis and lay the foundation of a critical review, as the data often backs up claims regarding the product in question. In this way, biometrics can play a critical role in improving the overall user experience (Kula et al., 2017).

According to Kula et. Al (2017), traditional measures of usability and user experience problems can be confirmed using eye-tracking data. An example of this is an investigation from 2017 of users' self-perceptions and task perceptions on a mobile app using a self-report questionnaire. In addition, it recorded eye fixations, interface action fixations, fixation duration, and pupil size. The findings revealed that gaze data was significantly and positively related to users' self-and task perceptions. Moreover, eye tracking can provide data that is superior to traditional measures. It has also been used for investigating how players'

performance in a game interface is affected by motion and background complexity. The results indicate that fixation time data provided a more accurate measurement of the impact of background complexity than task performance alone.

However, there are some cons to eye tracking. For example, head movements and gestures, such as shaking and forehead scratches, can cause tracking problems. In addition, users that must use glasses are unsuitable for eye tracking (Kula et al., 2017).

Other authors evaluated two versions of a government website using GSR and conventional usability techniques. One version had an animated face to encourage user involvement and offer navigational assistance. The second offered comparable information but lacked an animated face. Users' interactions with the animated-face website increased arousal and emotionality, according to usability and GSR data. According to the authors, the reason is that the animated face version requires less effort and navigational interactions.

GSR is useful in assessing emotional intensity and arousal overall. However, it does not say much or anything about whether those emotions are good (like excitement and enthusiasm) or bad (e.g., stress and fear) (Kula et al., 2017).

3. Methodology

A group of two master students facilitated a usability test of the webshop www.graviditetskollen.nu, a Swedish webshop focusing on the market within pregnancy, supplements, and clothing. The test was completed at the Usability Lab of the University of Agder. During the evaluation process, the participants were asked to carry out tasks and answer pre-test and post-test questions fitting the stakeholder's perspective within approximately 30 minutes. The conducted usability test was structured as follows.

1. Pretest data gathering of the test participants to complete a background check. Gathering initial impressions of the web shop early in the test.
2. Completing given tasks within the webshop to outline the journey from a stakeholder's perspective.
3. Completing post-test questions regarding the overall satisfaction of the webshop.

3.1 Test participants

Three students participated in the usability evaluation. They were recruited through the university. There is a clear difference between the stakeholders in this case, but as the group test facilitators, we can only test the perspective from a buyer's viewpoint. None of the participants had used the web shop before, so the test of this website was a test of first encounters with the system and its design. Two participants have well-developed computer literacy, and the third has little fundamental knowledge of computers or browsers. All the test participants had the following characteristics:

Audience type

Buyer	3
Total Participants	3

Weekly computer usage

10 – 20 hours	1
20 – 30 hours	2
Total Participants	3

Age

20 – 24	1
24 - 28	1
28 – 30	1
Total Participants	3

Gender

Male	3
Total Participants	3

3.2. Study

The study was conducted in a controlled usability lab at the University of Agder, where the participants were asked to complete six independent tasks. The group got the participants' permission and recorded them for later analysis.

Software:

This study used Imotions software, the world's leading provider of human insights software. It is used for analyzing human behavior data (Imotions, n.d.). Moreover, Imotions collaborate with Affectiva's AFFDEX, a technology that goes deeper into human emotional reactions and is used for detecting facial expressions (Imotions, n.d.).

Hardware:

To see what the participants gazed on, it was used device named Tobii. It is placed on the

screen, and after manual calibration, it catches the user's eye movements. Lastly, the team collected data using Shimmer3 GSR. It is a tool that monitors skin conductivity by placing sensors on the fingers and one of the ears.

4. User Tasks

To facilitate a usability test of the natural path for a customer within a website such as the one tested, we have decided to break up the test into different pieces, which are as follows: Navigation, product pages, Checkout process, search functionality, and overall user experience. The user tasks are as follows:

1.	Find supplements that help women get pregnant.
2.	Find "Storpack," a bag of pregnancy tests.
3.	Locate a Clearblue Fertilitetsmonitor without using the search bar.
4.	Find Lube that enhances fertility.
5.	Find Squitos 24 pack, a mosquito plaster for kids.
6.	Add Squitos 24 pack to the cart and follow through with checkout.

Table 1 User tasks

These tasks aim to determine if the website is easy to navigate and test its overall usability. Each participant completed the intended tasks.

4.1 Preliminary test questions:

Before the conduct of the test, each of the participants was asked these pre-test questions:

1. Have you ever used the platform before?

- None of the participants had used the platform before the test.

2. How old are you?

- All the participants are between the age of 24 – 30.

3. How many hours a week do you spend in the browser?

- The usage of browsers ranges between 10 – 40 hours a week.

4.1.1 Summary of preliminary questions

All three test participants had never used this website before but had proficiency within browsers and knowledge regarding webshops in general. However, with a general proficiency, there is significant fluctuation in time spent in browsers, which results in different times spent performing given tasks in the following parts.

4.2 Post-test questions:

To gather first impressions of the usability of the webshop, the participants were asked to answer questions regarding navigation, product pages, the checkout process, search functionality, and the overall user experience.

Moreover, the participants had to answer three questions determining their overall satisfaction with the website. The questions were in the Net Promoter Score (NPS) and Likert Scale scope. NPS is a customer satisfaction measurement of how likely people are to recommend a product (in this case, a website) to others (Hotjar, 2023). The following figures represent the participant's satisfaction with Graviditetskollen:

.

On a scale from 1 to 10, how likely are you to recommend this websit
3 svar

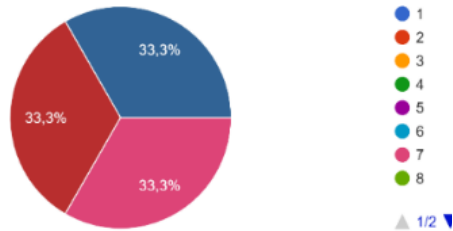|	
●	1
●	2
●	3
●	4
●	5
●	6
●	7
●	8

1/2 ▼

Figure 1 Recommendation of the website (NPS)

The participants' answers we got were 10, 9, and 7. According to Hotjar (2023), people who answer 9-10 are seen as promoters that help the brand grow. On the other hand, people who answer 7-8 are not likely to recommend the brand, but neither are they likely to spread the bad word about it (Hotjar, 2023).

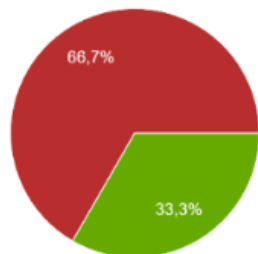On a scale from 1 to 10, how satisfied are you with the websi
3 svar

Figure 2 Satisfaction with the website (NPS). Answers: 8, 9, 9

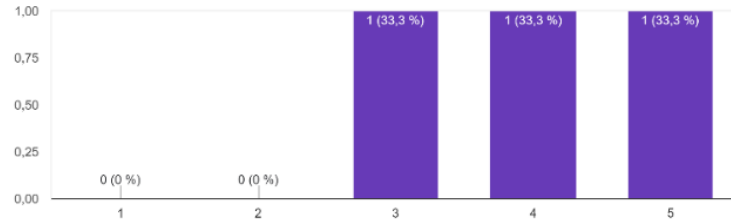I find the website easy to navigate
3 svar

Figure 3 Ease of use (Likert scale). (1- Strongly disagree, 5- Strongly agree)

5. Data

5.1 Areas of interest (AOI)

To gather relevant metrics from the participants, we decided to specify three areas of interest (AOI); Country (part of the checkout), Logo, and Menu. We wanted to research how many participants were using the logo to navigate to the front page and the menu to find the hierarchical path to the products mentioned in the tasks. As a last area of interest, we wanted to see whether the participants spent time looking at the information given to resolve the problem when trying to finish the checkout.

5.1.1 Gaze-based metrics

⊞ AOI metrics •••	Country	Logo
Gaze based metrics		
Respondent count	· 3	3
Respondent ratio (%)	100	100
Revisit count	4.3	4.3
Dwell time (ms)	3215.2	4835.6
Skip count	0	0

Figure 4 Gaze-based metrics

All three respondents gazed at the defined AOIs, and the average of how often they looked at the first two AOIs was 4.3. On the other hand, Menu AOI has been looked at six times more often than the previous two. Moreover, the results reveal that Country AOI has gazed at 3.215 sec and on average, followed by Logo with 4.835 sec. and Menu with 68.128 sec. Finally, none of the participants gazed at the AOIs without a detected fixation. The most gazed AOI is the Menu.

5.1.2 Fixation-based metrics

Fixation based metrics		
Dwells with fixations	3.7	
Fixation count	10	
TTFF AOI (ms)	2900.1	
Dwell time (ms)	3354.4	
Dwell time (%)	25.5	
Fixation duration (ms)	317.7	

Figure 5 Fixation-based metrics

We can see through the metrics in figure 5 that the menu is the component of the website that stands out with the highest dwells with fixations, fixation count, and dwell time. The participants mainly use the menu as navigation, not the search bar. We were surprised to see that the participants did look at the AOI that covered the section for selecting the country of residence without registering it and changing it to the correct country of residence.

5.1.3 Mouse-based metrics.

Mouse based metrics		
Mouse click count	1.3	

Figure 6 Mouse-based metrics

The average of all detected mouse clicks on Country AOI is 1.3, followed by 4.3 for Logo AOI and 5.3 for the Menu AOI.

5.2 Eye tracking and facial recognition

Facial recognition and eye tracking were used in the testing to capture data on the participant's emotional responses and natural browsing behaviors. This information can help designers and developers optimize website design to meet users' needs and preferences better. With the data gathered, we can provide the metric foundation for claims regarding the web shop's flaws.

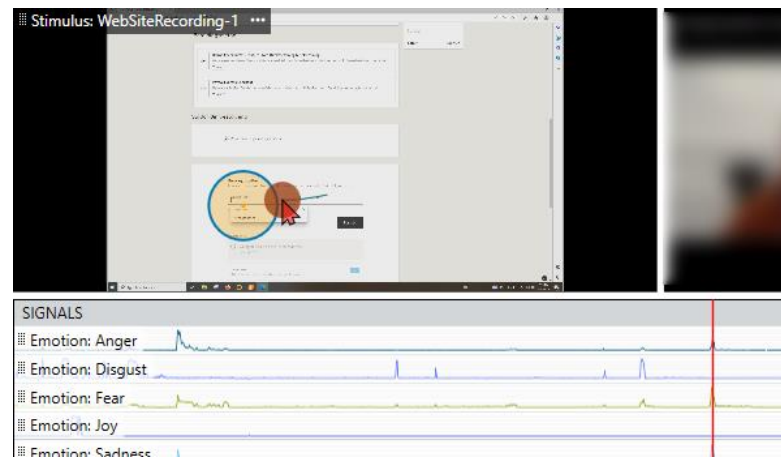### 5.2.1 Anger & Fear & Disgust

Figure 7 Emotion of fear and anger

During the completion of the last user task, the participant finds it challenging to checkout. As shown in figure 7, the levels of anger and fear both fluctuate at the same time.
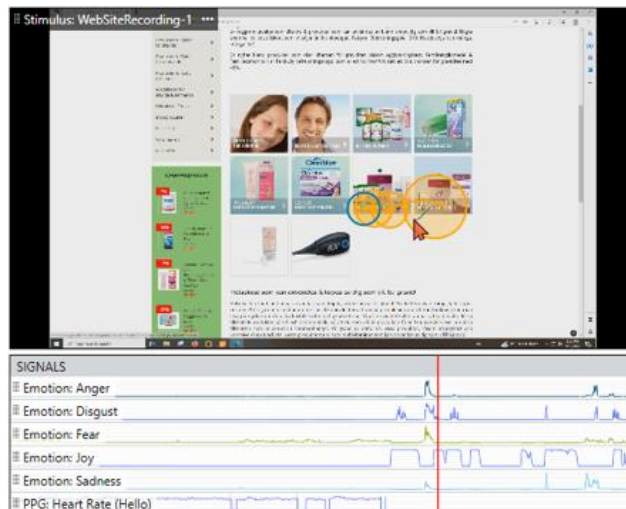
Figure 8 Emotions of fear, anger, and disgust

In this example, the participant feels fear, anger, and disgust because he cannot complete the task without using the search bar. However, some seconds later, he found the correct item and felt happy, as shown in figure 8.
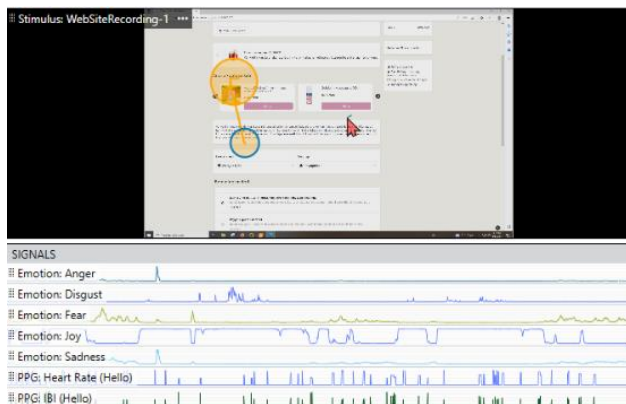
5.2.1 Joy

Figure 9 Emotion of joy

In this example, the participant feels joy when he finally can order the product.

6. Findings and recommendations

6.1 Interface

After completing the tasks, all participants mentioned that the website interface feels outdated. An outdated interface can make the website difficult to use and unappealing, leading to user frustration and decreased engagement. With a consensus regarding the outdated interface of the website, there is a potential that the design could damage the website's credibility.

6.1.1 Recommendation

Updating the interface can improve the user experience, increase satisfaction, and improve the website's success. Our recommendations regarding the interface are to use less stock photography and strive for a more consistent picture style. As seen in the main pictures on the landing page, the two pictures on the left are characterized by retro colors, and the two on the right are more crisp, clean, modern, and blue-hinged. Moreover, the main critique lies in the design not being up to date regarding trends. This can be seen primarily in the use of colors. Within pregnancy and supplements, a blue color would be more appropriate as it is a good combination of health and seriousness. The primary green color could therefore be exchanged for blue color. We also see that the green color in the logo is not perceived as the same as the primary color and contributes to the cheap impression. Altogether the color vocabulary can be perceived as slightly messy.

6.2. Checkout error

During the last task, to buy a product, all participants experienced difficulties checking out. When entering the user's postcode, an error appeared. (see figure 9) To gather information regarding this error, we set the part of the checkout section that provided errors as an AOI of a not valid postcode (See figure 1).
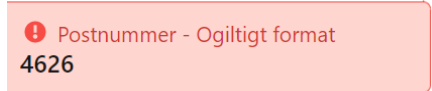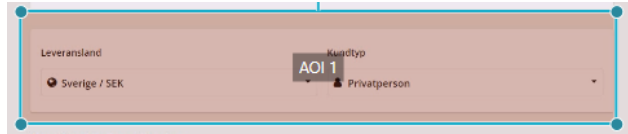
Figure 10 Postcode error


Figure 11 AOI of country

6.2.2 Recommendation

To prevent users from encountering errors when purchasing products due to incorrect country of residence settings, we recommend that the website implement an automatic IP address detection system. This system can map a user's IP address to their physical location and automatically set their country of residence accordingly. The moment of checkout in eCommerce is critical and is also where most customers leave. The checkout should therefore be as optimized as possible and not have any functions that cancel the flow of the customer.

Figure 12 Choice of the delivery country

6.3 Messy checkout section

One participant mentioned the checkout section within the webshop to be messy and poorly designed. The test participant mentioned that there are many unnecessary elements and that it could have a more straightforward path. This is set apart from the error previously mentioned.
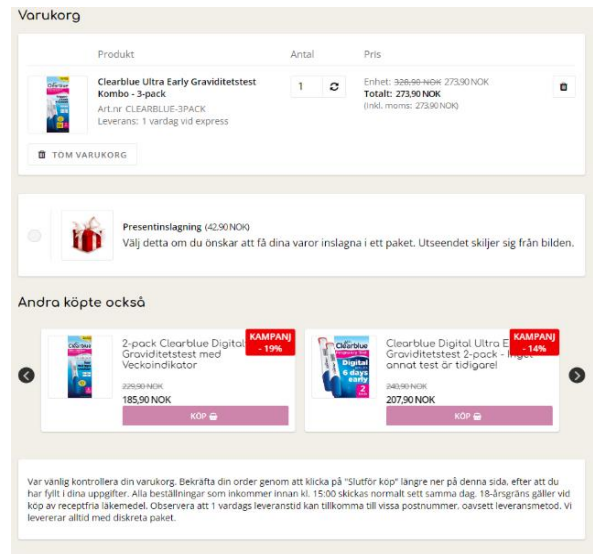
6.3.1 Recommendations

As well as setting the country or residence to the appropriate one using an automatic IP solution, we also would recommend that there is implemented a more step-by-step path for the customers trying to check out. The web shop's design seems messy within the check-out section as it does not provide a step-by-step path and includes elements that are not often very interesting, such as gift wraps and products other people bought

in combination with what you already have in your basket.

Figure 13 The check-out section within graviditetskollen is perceived as messy

Again, the design should strive for a more suitable color selection and use it to build a consistent design. As shown in figure 14, we recommend that Graviditetskollen strive for a more hierarchical flow like that have done in Komplett.no
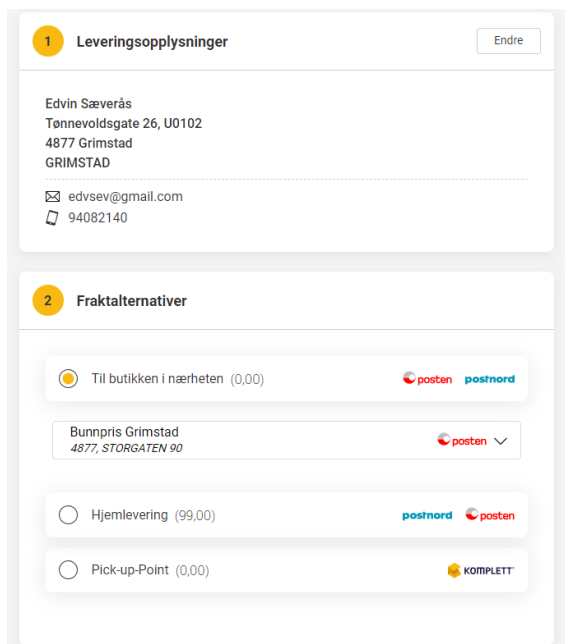

Figure 14 The checkout section within Komplett

6.4 Deviance in performance

	participant			Time on task	AVG
	1	2	3		
Task 1	47,0	41,0	55,0	143,0	47,7
Task 2	50,0	52,0	77,0	179,0	59,7
Task 3	44,0	42,0	64,0	150,0	50,0
Task 4	54,0	59,0	82,0	195,0	65,0
Task 5	54,0	51,0	67,0	172,0	57,3
Task 6	87,0	82,0	113,0	282,0	94,0
SUM	336,0	327,0	458,0	1121,0	
AVG	56,0	54,5	76,3	186,8	

Figure 15 Deviances in time on task, numeric values represent seconds.

In the section of the test where the participants were to carry out different tasks regarding using the website, the results showed that the participant who spent the least time weekly in browsers had the most challenging time completing the given tasks. This can be attributed to the participant's general lack of experience with browsers and the web. With a lack of knowledge, the third participant stands out with significantly higher time on tasks than the other participants who spend more time on browsers.

Instead of using the search bar as a means of navigation, only one of the test participants ever used the search to find one specific item. To navigate to the products mentioned in the tasks, the participants mainly used the menu to navigate. We believe this contributes to a longer time on task for the participants and represents, therefore, a potential area to be improved.

6.4.1 Recommendations

The results highlight the importance of designing an intuitive and easy webshop for users with varying experiences with browsers and the web. By doing so, the webshop can provide a seamless user experience for all users, regardless of their level of expertise. To minimize the time

spent on tasks, we recommend placing the search bar in a more visible area and presenting it with a more apparent contrast to the surrounding elements. In order to tie the design together and decrease the messy color vocabulary, they should beforehand work on the primary and complementary colors. After that, design the search bar with the same colors.

As shown below (figure 16) from www.komplett.no, they have placed the search bar closer to the logo, which functions as a part of the navigation. Not only that, but the design also uses the company's logo as a contrast. This is something we would recommend graviditetskollen also does. Perhaps also include a more explicit instructional placeholder text in the search bar, as shown in figure 16.

Figure 16 correct use of search bar design

Figure 17 Graviditetskollens' search bar

Only one participant ever used the search bar while locating a specific product. We believe that implementing the recommendations mentioned above will make the search bar easier to spot for a potential customer and consecutively decrease time spent on tasks.

Resources

Hotjar. (2023). What is net promoter score (NPS)? How to calculate & interpret. https://www.hotjar.com/net-promoter-score/

Kula, I., Branaghan, R. J., Atkinson, R. D., & Roscoe, R. D. (2017). Assessing User Experience via Biometric Sensor Affect Detection. IGI Global EBooks, pp. 123–139. https://doi.org/10.4018/978-1-5225-2639-1.ch006

https://imotions.com/

https://komplett.no/