Impact of Multiple Cues on Facial Emotion Recognition and Visual Attention of Individuals with ASD compared to Individuals without ASD

Anne Bloem
Industrial Design
Eindhoven, University of Technology
The Netherlands
a.c.bloem@student.tue.nl

Inge Hootsmans
Industrial Design
Eindhoven, University of Technology
The Netherlands
i.m.hootsmans@student.tue.nl

Lena Opheij
Industrial Design
Eindhoven, University of Technology
The Netherlands
l.m.opheij@student.tue.nl

Romain Toebosch
Industrial Design
Eindhoven, University of Technology
The Netherlands
r.h.a.toebosch@student.tue.nl

ABSTRACT
Due to an increase in the use and development of Artificial Intelligence and social healthcare robots, the demand for understanding of the human skill of emotion recognition is rising. Individuals with Autism Spectrum Disorder process sensory information differently, and designing intelligent technologies for them requires further analysis in the way they perform emotion recognition. In this study, 22 participants were shown videos of actors expressing the six basic emotions with or without the auditory cue removed, and then asked to indicate what emotion they thought to have recognized. During the entire test their gaze was tracked. In this paper we conclude that the inclusion of the auditory cue does impact the emotion recognition, but has no or minor impact on the visual attention, regardless of participants having ASD or not.

Author Keywords
Emotion recognition; Multiple cues; Visual and auditory cues; Autism Spectrum Disorder; Gaze tracking.

INTRODUCTION
With the increase of the use and development of Artificial Intelligence (AI) and social healthcare robots, the need for improvements in facial emotion recognition is rising. To get a better understanding of how to design personalized technologies that use emotion recognition, more information on the human skill of facial emotion recognition are needed.

Where many computer systems focus solely on the use of visual inputs to recognize emotions, humans have multiple senses they can rely on. This multimodality impacts our emotion recognition skills (De Silva, Miyasato, & Nakatsu, 1997). However, not all humans process this information in the same way: individuals with Autism Spectrum Disorder (ASD) process sensory cues differently than individuals with typical development (Van der Smagt, van Engeland, & Kemner, 2007). Mostly, persons with ASD only focus on one sense at a time (Lovaas et al., 1971). This is proven to be relevant to the problems of language and emotion processing shown by subjects with ASD (Van der Smagt et al., 2007). Therefore, it is interesting to look into the differences between the ways individuals with and without ASD recognize emotions.

Since most emotions are recognized by looking at specific subtle changes in facial muscles (Tian, Kanade, & Cohn, 2001), it could be of interest to investigate the differences in visual attention between individuals with and without ASD. Moreover, we want to find out whether single or multiple modalities, as this could tell more about the way people with or without ASD recognize emotions when having either multiple or single cues.

In this paper, a quantitative user study with 11 adults with ASD and 11 typical adults is presented. The focus of this study will be on the impact of multiple cues on visual attention and emotion recognition of individuals with compared to without ASD. More specifically, the difference between visual and audio-visual cues in facial emotion recognition is measured.

This study could help further research on AI and social healthcare robots, but could also impact the way products are designed for individuals with ASD.

RELATED WORK
A number of factors play a role in emotion expression and recognition. Primarily, these factors are: face (De Silva et al., 1997), body (Shan, Gong, & McOwan, 2007) and voice (De Silva et al., 1997). Shan et al.

(2007) researched the features of facial and body gestures in human emotion expression. The extraction of data for facial emotion recognition and bodily emotion recognition differs immensely: in bodily emotion recognition the body and hand movements are tracked (Castellano, Kessous, & Caridakis, 1973); for facial emotion recognition, the focus is predominantly on the Action Units (AU) of the face. Most of these AUs are related to the contractions of particular facial muscle groups (Tian, Kanade, & Cohn, 2001).

Research concerning AI facial emotion recognition has often relied on one modality at a time. However, it is important to take multimodality into consideration as well, because humans process auditory and visual information in parallel to recognize emotions (De Silva et al., 1997). Research conducted by Busso et al. (2004) has concluded that emotions are recognized more accurately with facial expressions, rather than just with speech. When these sensory cues are combined, the emotion recognition improves noticeably. However, these papers do not discuss whether having multiple cues impacts the way the other cues are perceived. Eye tracking could be used to see if the inclusion of audio impacts visual attention in emotion recognition.

Emotion recognition and multiple cues in combination with ASD has been a field of interest for a while, since the diagnostic criteria of ASD include emotional impairments (Begeer et al., 2008). ASD is connected to multiple sensory disturbances in the original description of autism. More specifically, individuals with ASD especially fall short in speech, in the processing of temporal multiple cues (Stevenson et al., 2014).

Although before mentioned studies have shown that multiple cues do impact the emotion recognition skill (De Silva et al., 1997; Busso et al., 2004), none of them show the impact of the removal of a cue on the way the other cues are perceived. A study from Rutherford and Towns (2008) showed no significant difference between gaze from adults with ASD and typical adults, but did not include sound in their test. In this paper, the effect of sound on the emotion recognition and visual attention will be discussed.

METHODS

To investigate the impact of audio on the visual attention and emotion recognition skills, a study setup was created. With this setup multiple adults either with or without ASD were tested. The test consisted of two segments of different videos in which actors display one of the six basic emotions. These videos were shown to the participants in random order. After each video, the participant answered in a multiple choice questionnaire what emotion they thought to have seen. During the display of the videos an eye tracker tracked their gaze. The information was separated for both with/without sound (n/s) and group A/T, as shown in Table 1. In this paper when referring to adults or participants without ASD, the typical adult group (T) is meant.

	No Sound	Sound
Group A: Adults with ASD	An	As
Group T: Typical adults	Tn	Ts

Table 1. Different category names.

Measurements

During the test multiple videos from the RAVDESS dataset (Livingstone & Russo, 2018) were shown, either with or without sound. This verified dataset contains videos that show the face of an actor acting out a particular emotional expression. The specific videos were selected by researchers of the University of Hamburg. At this university, research is carried out by neuroscientists and computer scientists, who showed the videos to their subjects. They concluded that the exaggerated videos were more easily perceived by the participants, than the non-exaggerated videos. This corresponds with research done by Castellano, Kessous, and Cardakis (1973). To stay in line with their research and in order to provide the base ground for future collaboration, the decision was made to use the same videos in this study. Videos concerning the emotion ‘Surprise’ were removed by the researchers from Hamburg, since the participants had difficulties recognizing this emotion. Next to the other emotions, the emotion ‘Surprise’ is added from the RAVDESS dataset to this study, to correspond with the six basic emotions defined by Ekman and Keltner (1997).

Within the test there was refrained from implementing a fixed order of the audio and the non-audio segment. This is to prevent that the participants’ emotion recognition skills are influenced by the dataset shown beforehand. Therefore, half of the participants first saw a set

of videos with sound, followed by a set of videos without sound. The other half of the participants saw videos without sound first, and the set of videos with sound second, as shown in Table 2. Each participant got one of the two orders pseudo-randomly. The order of the videos within the two sets were played at random as well. Next to that, the videos from the sets with and without sound differ from each other. This was mentioned to the participant before taking the test, to prevent participants from trying to remember the videos.

Order	First	Second
Group 1	Sound	No sound
Group 2	No sound	Sound

Table 2. The two orders in which the test was conducted.

During the test, the participant selected which emotion they thought to have recognized by clicking on one of the options displayed on the computer monitor. The options from which the participant could choose are: 'Anger', 'Disgust', 'Fear', 'Happiness', 'Sadness', 'Surprise', and 'I don't know'. The 'I don't know'-option is added for when the participant does not know which emotion is displayed. Since this study was done by Dutch participants all the options were written in Dutch ('Boos', 'Afschuw', 'Bang', 'Blij', 'Verdrietig', 'Verrast', and 'Ik weet het niet'), to avoid any misunderstandings. The videos are in English, but since the sentences were meaningless and had no emotional charge, this should not prevents the participant from recognizing the emotions. The sentences were clarified by the researchers before the participants took the test. The eye gaze of the participants was simultaneously tracked throughout the test, by an eye tracking device.

Next to this, questions and remarks from the participants during the tests will be noted. It is not mandatory for the participants to comment.

The eye tracking data, together with the answer data, was aggregated in Python and analyzed in SPSS (SciPy developers, 2019; IBM Corp., 2017). The differences between the groups was analyzed with a Multivariate Analysis of Variance. The difference within subjects (between the tests) was performed using a Wilcoxon Matched-Pairs Test.

Participants

In this study two groups were created. Group A consists of eleven adults with a form of ASD. Some participants of group A also had a different disorder, such as ADHD, or a physical constraint. Group T consists of twenty typical adults, meaning they do not have any mental developmental disorders. Although people from different cultures are able to recognize the six basic emotions (Ekman & Keltner, 1997), only participants with a Dutch nationality were used in this study.

The gender distribution is the same in both groups. All participants are in the same age group, between 18 and 30 years old. According to article 233 from Burgerlijk Wetboek Boek 1 in the Netherlands a person is considered an adult when they turn 18 (Overheid.nl, 2019). This means that they can give consent to participate in the study themselves, instead of their guardians. This also made this age group accessible for the researchers, which meant that a higher number of participants was feasible. All participants have received the same compensation for taking the test.

Measurement instruments

To measure the visual attention of people, a Tobii Pro X2-60 Eye tracker was used in combination with the Tobii Studio software. The eye tracker was placed underneath a 24-inch 1080p display, that showed videos from the RAVDESS dataset (Livingstone & Russo, 2018). From this dataset a total of 36 videos with and 36 videos without sound were shown to the participants. Which videos have sound and which do not was selected at random.

During the study, the participants selected an answer by clicking one of the seven options in the questionnaire shown on the screen in between the videos. The answers were saved in Tobii Studio. The computer on which Tobii Studio ran, saved all the gaze and answer data, which was exported in .tsv files.

Context

The study took place in three different neutral rooms. All rooms were closed off, meaning the participant was not influenced by distractions from their surroundings during the study. When taking the test there were two or three researchers and one participant present in the room, for both groups A and T.

The first location of the study was in a room in Atlas at the Technical University in Eindhoven (Figure 1). The

second one was at Stichting OOK Begeleiding in Geldrop. This is an independent organization that specializes in guiding individuals and families with autism (Stichting OOK Begeleiding, n.d.). The study took place at their institution, as this was convenient for the participants. The third location is an office space in the public library of Alkmaar. All test locations are situated in the Netherlands.

Figure 1. Picture of test set-up in Atlas, one of the three neutral rooms where the study took place.

The study was carried out during a three-week period. The duration of the test was approximately 20-30 minutes, depending on how long the participant took their time answering the questions.

Study Protocol

In this study, multiple procedures took place to ensure that the data was gathered in the same way. First, the participant was welcomed in the set environment, was asked to fill in the consent form, and was then seated in front of the monitor. A researcher explained the calibration of the eye tracker. When the participant had no further questions, the eye tracker was calibrated. After calibration succeeded, an explanation of the test was given by a researcher and the introductory example test was started. This example test was added to check whether the participant understood what they needed to do and for the researchers to confirm if the eye tracker was calibrated properly. The results from the example test were not taken into account for the results of the findings in this paper. Next, the real test started. After each video the participant filled in the questionnaire, then the next video was played. At the end of the test the participant was thanked for their cooperation and received an appropriate compensation.

FINDINGS

For the analysis of the test the results of all eleven ASD participants' were used. To keep an equal amount of participants with and without ASD, a selection from the participants without ASD was made before reviewing the results. This selection is based on similarities in the participants ages and sexes, so the participants matched as closely as possible. For the visual attention only nine participants' results within each group will be used. This is due to the eye tracker not following the eyes of two participants with ASD properly. To keep the results accurate, two matching participants without ASD have also been removed from the gaze results.

Furthermore, a human error was made while selecting the videos with sound for the test. Two videos were duplicated, this concerns the emotions 'Angry' and 'Sad'. This resulted in 'Angry' having one more duplicated video, 'Sad' having the same amount (although having a duplicate video), and 'Surprised' having one video less. This was taken into account when doing the analysis.

Answers test

To process all the answer data, gathered from the questions in between the videos, data was aggregated in Python using the SciPy libraries (SciPy developers, 2019): the percentage scores per participant were calculated for each test (n/s), as shown in Table 3. Then this aggregated data was analyzed in SPSS (IBM Corp., 2017). In this entire study a significance level of $\alpha = 0.05$ was chosen.

Participant	Group	Score test n	Score test s
ID	A/T	Score between 0 and 1	Score between 0 and 1

Table 3. Structure of the data aggregated in Python.

Using SPSS descriptive statistics have been calculated, as shown in Table 4.

Group		Mean	Std. Deviation	N
Sound Answers Mean Score	no ASD	0.93131300	0.059694177	11
	ASD	0.91414136	0.069671291	11
	Total	0.92272718	0.063918324	22
No Sound Answers Mean Score	no ASD	0.82575764	0.099451269	11
	ASD	0.77020209	0.100991300	11
	Total	0.79797986	0.101857341	22

Table 4. Descriptive statistics of the answer data.

First the between subject analysis of the answers was done, between group A and T. Since the test had two continuous outcome variables (independently observed) and two categorical predictor variables, the best suited test would be a Multivariate Analysis of Variance (MANOVA) (O'brien & Kaiser, 1985). Before performing this test, several assumptions have to be met: the data should be normally distributed; all dependent variables should have a linear relationship; there should be an equality of covariance matrices.

To test the normality, a Q-Q plot was made, as shown in Figure 2. If the points of the results approximate the line y=x, the data is normally distributed.

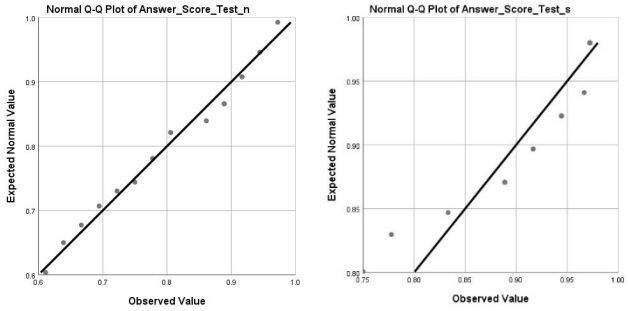

Figure 2. Q-Q plot of the scores of the answers of test n (left) and test s (right).

To test the linearity of the results, a scatter plot was made between the scores of test n and s (Figure 3). If no non-linear relationships, such as exponential or parabolic relationships are found, the data is considered to have either a linear or no association (Walker, 2017).

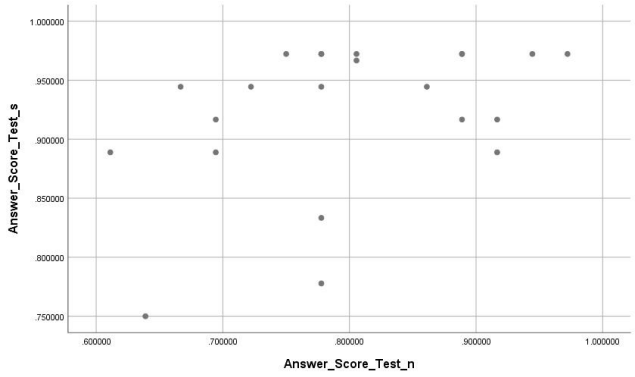

Figure 3. Scatter plot of the scores of the two tests (n/s).

To test the homogeneity of covariance matrices, Box's test of equality of covariance matrices (BTECM) and Levene's test of equality of error variances were performed (LTEEV). If the outcome of these tests are insignificant, the assumption is met. The result for BTECM is p=0.960 and the results for LTEEV are p_{sound}=0.756 and $p_{no\ sound}$=0.874.

The results of the MANOVA test show whether there is a significant difference between the results of both groups. Wilk's Λ shows difference between the results of group A and T are insignificant, $\Lambda = 0.459$, $F(2, 19) = 0.811$, $\eta^2 = 0.079$. MANOVA tests were also performed per emotion, these all showed insignificant results, as shown in Table 5.

Emotion(s)	Wilk's Λ
All	0.459
Angry	0.854
Disgust	0.154
Fearful	0.224
Happy	0.781
Sad	0.789
Surprised	0.558

Table 5. Results of the MANOVA tests per emotion.

Second, the within subject analysis of the difference between test n and s was performed for both groups. Since the compared variables are ordinal, independent, and repeated measures of the same populations, a Wilcoxon Matched-Pairs test was performed. If the results are significant, there is no proven difference between the results of both tests. Results for both groups are shown in Table 6.

Emotion(s)	P-value group A	P-value group T
All	0.005	0.013
Angry	0.015	0.225
Disgust	0.058	0.317
Fearful	0.012	0.011
Happy	0.783	0.763
Sad	0.011	0.011
Surprised	0.575	0.497

Table 6. Results of the Wilcoxon Matched-Pairs test per emotion for both groups.

To give more insight in these statistics, several confusion matrices have been created of the data grouped per

category An/As/Tn/Ts. The confusion matrices show the relationship between the emotions expressed and the answers given by the participants. The data has been normalized within each column for easier comparison. The number shown in each square is the percentage of participants that gave the corresponding answer to that emotion. These confusion matrices can be seen in Figure 4.

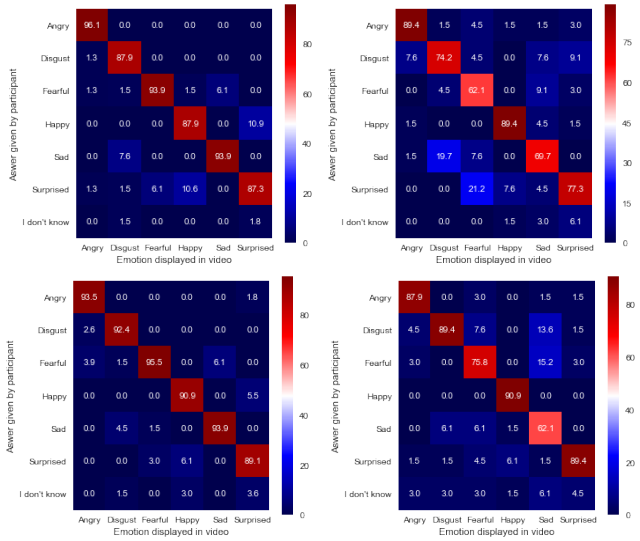

Figure 4. Confusion matrices. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

Gaze

Throughout the test, gaze data was tracked by the eye tracker. The Tobii Studio software processed the data to calculate several new variables, of which the location of the gaze based on screen pixels and the fixation points were used. This data was exported from the software and aggregated in Python using the open source SciPy ecosystem (SciPy developers, 2019).

To get an indication whether the visual attention shifted between the categories, heatmaps were created to visualize the gaze of the participants. Since it is difficult to perceive the exact gaze positions on a heatmap, each point was given a radius of 50 pixels, to more easily perceive where the visual attention of the participant was. In Figure 5 heatmaps with and without radius increase are shown respectively. All the following heatmaps will use the same radius.

Figure 5. Heatmaps. Left: Heatmap of a set of gaze points without added radius of 50 pixels. Right: Heatmap of set of gaze points with added radius of 50 pixels.

The heatmaps represent the screen of 1920 by 1080 pixels the participant viewed, with the RAVDESS video in the middle. Since the video itself has a resolution of 1280 by 720 pixels, the video was not shown fullscreen. The following heatmaps in Figure 6 will be cropped onto the face to increase visibility.

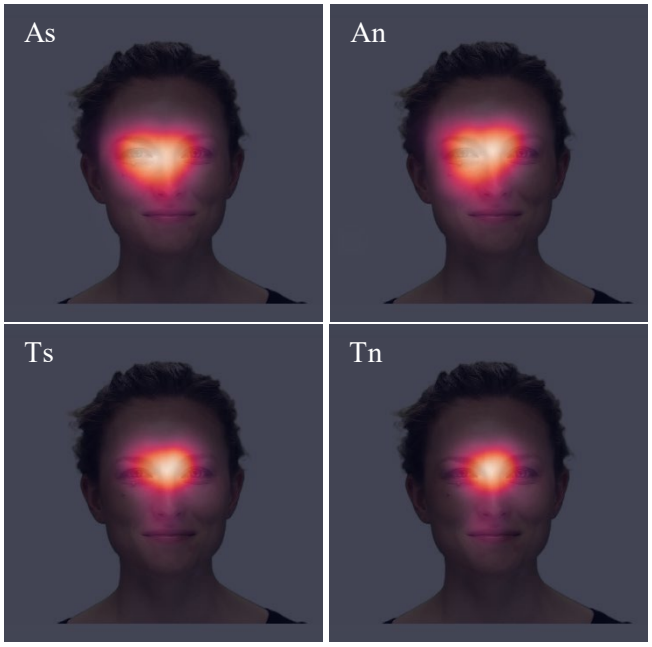

Figure 6. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

Heatmaps were also made per emotion. These can be found in the appendix (see Appendix A).

Observations by the researchers

Observations were made during the testing phase. Before, between, and after the tests the researchers asked the participant if they had any questions or wanted to give comments. A selection of notable comments regarding this study was made. The comments were given in Dutch and are translated in English. Several participants with and without ASD gave similar comments on the videos without sound, with one participant from group T saying: “The videos without sound are

more difficult [than videos with sound], one focuses on other aspects [of the video]". Another participant from group T commented "I watched the videos with sound less actively". A remarkable comment made by an individual from group A was: "With sound is sometimes easier, without sound you really had to pay attention to facial expressions".

There were comments on how the emotions were perceived. Some participants remarked "A lot of emotions look the same". Others mentioned that the difference between certain emotions was difficult to see, as an individual from group T said: "Surprise leans towards indignantly in some videos". Another participant from group A commented: "The difference between disgust and sad is hard to see, but it was easier to recognize with sound".

DISCUSSION

In this section the results from the tests are discussed in relation to the research question and theoretical background. First, the interpretations from the researchers and implication of the results are discussed. Second, the limitations of this study and suggestions of possible future research on this topic are given.

Interpretation of findings

Emotion recognition

When comparing the results from group A and T, the MANOVA test showed there was no significant difference between the participants with and without ASD. This is not in line with the literature, which suggests people with ASD should perform worse on average (Van der Smagt et al., 2007). This discontinuity could have multiple reasons: first, it could be possible the amount of participants was too low; second, it could be the difference is unnoticeable when showing the six basic emotions instead of mixed or more subtle emotions; third, it could be that there is a discontinuity because the literature is mostly about children with ASD (Begeer et al., 2008; Lovaas, et al., 1971; Stevenson et al., 2014), and our participants were all adults.

However, when comparing the results with and without sound using the Wilcoxon Matched-Pairs test, a significant difference is found. In general, both groups significantly performed worse in the test without sound. This is reinforced by the confusion matrices and the descriptive statistics, showing more mistakes being made in the tests without sound. This is however not always the case when looking at the results per emotion. For the participants of group T, only the emotions 'Fearful' and 'Sad' showed significant difference between the two tests. For group A, 'Angry', 'Fearful', and 'Sad' showed significant difference between the two tests, and disgust approaches significance.

Literature shows 'Fearful' and 'Sad' are mostly auditory dominant emotions (De Silva et al., 1997), and this is in line with the Wilcoxon Matched-Pair test results of both groups A and T. In the confusion matrix it can be seen that for group T these emotions went from best answered with sound to worst answered without. Participants of group A also had a significant difference within the emotion 'Angry', and came near significance within the emotion 'Disgust', according to the Wilcoxon Matched-Pair test. This could mean that these emotions are more auditory dominant for people with ASD, but for 'Angry' it could also be the result of the human error that was described in results, which, even if the results were normalized, could have an influence on the ranking within the test.

In the confusion matrix it also seems that group T has more difficulty to recognize 'Sad' without sound than group A does, but this is not proven by the statistical tests. Further it seems that group T confused 'Happy' and 'Surprised' more often with each other with sound compared to without sound, but the difference is small and could be insignificant.

Visual attention

When comparing the heatmaps of group A and T some differences stand out. For instance, in the general heatmaps can be seen that the gaze has a larger spread for group A in comparison to group T, both with and without sound. This trend continues when comparing the heatmaps of the separate emotions (see Appendix A). Participants of group T mostly focus on the eyes of the actors. When group T shifts their gaze to the mouth, it is focused on the center of the mouth. However, the shift to the mouth and cheek region is far less common and/or spread when compared to the group A. A slight preference for the left side of the face can be seen in the heatmaps of group A. Research conducted by Rutherford and Towns (2008) concluded that people with and without ASD mostly look at the eyes instead of the mouth. This is also the case in this study when looking

at the heatmaps. However, the same research by Rutherford and Towns (2008) suggests there is no significant difference between the gaze of individuals with and without ASD, which is not the case when looking at the data from the findings in this study.

Comparing the sound and no sound heatmaps, differences are less drastically than the difference between the groups. It seems that group T focused their gaze slightly more on the eye region during the no sound test, but this could be insignificant. The gaze of group A is mostly the same with or without sound, if not spread slightly more at the test without sound.

Even though the visual attention differs between group A and T, this does not seem to effect the emotion recognition. Where the emotion recognition does show a difference between the tests with and without sound for both groups, the visual attention stays mostly the same. This could imply that the exclusion of one of the cues (sound) does not affect the way others are perceived (visual attention). However, this would need further investigation.

Implications

For the field of AI and emotion recognition, this study has confirmed the importance of sound for emotion recognition. It has also shown where the visual attention of people lies when trying to recognize emotions both with and without sound, which can potentially be used to optimize emotion recognition AI. Furthermore, confirming which emotions are more visually and which emotions are more auditory, AI researchers can possibly optimize AI to differ in its approach per emotion.

For research concerning ASD, this study has shown how sound influences the emotion recognition skills of people with ASD, and which emotions benefit most of sound in order to be recognized correctly by people with ASD. In this study a discontinuity was found between which emotions were more auditory based between people with ASD and without, but this needs further investigation. It has also given insight into visual attention of people with ASD, and the implications of sound on their visual attention. Lastly, this study has given an overview of which emotions are confused most with each other, with sound or without sound.

For the design of intelligent technologies dedicated to individuals with ASD, this study has shown where their visual attention lies. This can be useful information when designing solutions for emotion recognition training or social healthcare robots. Furthermore, it can be beneficial for designers trying to incorporate multimodality into their design for people with ASD.

Limitations

During this study, certain limitations were found. By working with a niche group, namely people with ASD, the sample size of participants is rather small and from a select range of ages. This could result in less reliable results, especially since the test was self-designed. Furthermore, the education level of the participants were not taken into account in this study. This could have influence on their ability to recognize emotions and thus have influenced the results. Next to this, gaze was treated qualitatively in this study. To more accurately interpret the gaze, the gaze data should be quantified and analyzed statistically.

There is a chance that the selection of videos to the sound or no sound category had influence on the difficulty of the set, even though this was done at random. This could be improved by creating a second test were the videos with sound are shown without sound and vice versa. Lastly, instead of mixing the sound and no sound videos, the videos were grouped together, which may have given a different result.

Another thing that could conflict with the reliability is that the testing was done in a controlled environment. This means that in real life the outcome could differ, as participants are not necessarily actively trying to recognize emotions. During this research all the participant only had to focus on the six basic emotions, while knowing their gaze was being tracked. In real life it is more likely that a person expresses a more complex or mixed emotion than a filtered version of the six basic emotions.

Even though Ekman says that all six basic emotions are universal (1999), there could be a (minor) difference in the expression of an emotion by North Americans (where the videos were made) to how an emotion is expressed by Europeans or Dutch people, which could have a (minor) influence on the results of the test.

Finally, it is common for people with ASD to have a combination with other disabilities such as ADHD. Its influence on this study is uncertain.

Suggestions for further research

For the research to be more applicable to the real world, testing can be expanded by adding more complex or mixed emotions than only the six basic emotions. Another interesting perspective would be if the testing would happen in real-time, thus being face to face with an actor. And how the real life recognition would differ from monitor recognition.

A recommendation for this kind of tests would be the development of custom software, which also tracks the actor's face to create a more accurate map of where the visual attention of the participants lies. This would eliminate the movement of the face out of the heatmaps and could also help improve accuracy of the analysis, as more specific parts of the face could be analyzed. This way the perception of specific Action Units by the participants can be monitored.

Lastly, it would be interesting to do joint research between visual attention and neurological activity when comparing emotion recognition with and without sound, to see if the brain reacts differently to this multimodality.

CONCLUSION

Concluding, the auditory cue does impact the emotion recognition, but has no or minor impact on the visual attention, regardless of participants having ASD or not.

The recognition of emotions is generally more difficult when there is no sound for both typical adults and adults with ASD. No significant difference between persons with and without ASD is found in the emotion recognition. The recognition of the emotions 'Fearful' and 'Sad' in the test without sound is hard for both participants with and without ASD, but participants with ASD also struggled slightly more with 'Angry' and 'Disgust'.

There is a difference in the visual attention between the typical adults and adults with ASD. The gaze of people with ASD was spread broader than for people without ASD, for both with and without sound. However, this area needs further investigation.

ACKNOWLEDGMENTS

We want to thank: all the participants, for participating in our test; our coach Emilia Barakova, who guided us through the process, and wrote and provided helpful comments on previous versions of this document; Pablo Barros and Matthias Kerzel from the University of Hamburg, for providing feedback; our squad coach Daniel Tetteroo for advisement; Inge van den Broek from Stichting OOK Begeleiding and Moos Wittebrood, for helping us find participants; Matthias Rauterberg for advice on experimental design.

REFERENCES

1. Begeer, S., Koot, H. M., Rieffe, C., Terwogt, M. M., & Stegge, H. (2008). Emotional competence in children with autism: Diagnostic criteria and empirical evidence. *Developmental Review*, *28*(3), 342-369.
2. Busso, C., Deng, Z., Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., ... & Narayanan, S. (2004, October). Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 205-211). ACM.
3. De Silva, L. C., Miyasato, T., & Nakatsu, R. (1997, September). Facial emotion recognition using multi-modal information. In Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat. (Vol. 1, pp. 397-401). IEEE.
4. Ekman, P. (1999). Basic emotions. *Handbook of cognition and emotion*, 45-60.
5. Ekman, P., & Keltner, D. (1997). Universal facial expressions of emotion. Segerstrale U, P. Molnar P, eds. Nonverbal communication: Where nature meets culture, 27-46.
6. IBM Corp. Released 2017. *IBM SPSS Statistics for Windows*, Version 25.0. Armonk, NY: IBM Corp.
7. Kessous, L., Castellano, G., & Caridakis, G. (2010). Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis. *Journal on Multimodal User Interfaces, 3*(1-2), 33-48.
8. Livingstone, S. R., & Russo, F. A. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PloS one*, *13*(5), e0196391.
9. Lovaas, O. I., Schreibman, L., Koegel, R., & Rehm, R. (1971). Selective responding by autistic children to multiple sensory input. *Journal of abnormal psychology*, *77*(3), 211.

10. O'brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychological bulletin*, *97*(2), 316.
11. Overheid.nl. (2019, January 29). *Burgerlijk Wetboek Boek 1*. Retrieved May 18, 2019, from wetten.overheid.nl: https://wetten.overheid.nl/BWBR0002656/2019-01-29
12. Rutherford, M. D., & Towns, A. M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of autism and developmental disorders*, *38*(7), 1371-1381.
13. SciPy developers. (2019). *SciPy.org*. Retrieved June 1, 2019, from scipy.org: https://www.scipy.org
14. Shan, C., Gong, S., & McOwan, P. W. (2007, September). Beyond Facial Expressions: Learning Human Emotion from Body Gestures. In *BMVC* (pp. 1-10).
15. Van der Smagt, M. J., van Engeland, H., & Kemner, C. (2007). Brief report: can you see what is not there? Low-level auditory–visual integration in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(10), 2014-2019.
16. Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience*, *34*(3), 691-697.
17. Stichting OOK Begeleiding. (n.d.). *Stichting OOK Begeleiding*. Retrieved May 19, 2019, from ookbegeleiding.nl: https://www.ookbegeleiding.nl/ook.html
18. Tian, Y. I., Kanade, T., & Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *IEEE Transactions on pattern analysis and machine intelligence*, *23*(2), 97-115.
19. Walker, J. A. (2017, June 3). *Reading scatterplots - MathBootCamps.* Retrieved June 7, 2019, from https://www.mathbootcamps.com/reading-scatterplots/

APPENDIX A

Heatmaps of the gaze within the four categories: An, As, Ts, and Tn, per emotion.

[A.1] Heatmaps of the gaze for the emotion 'Angry'

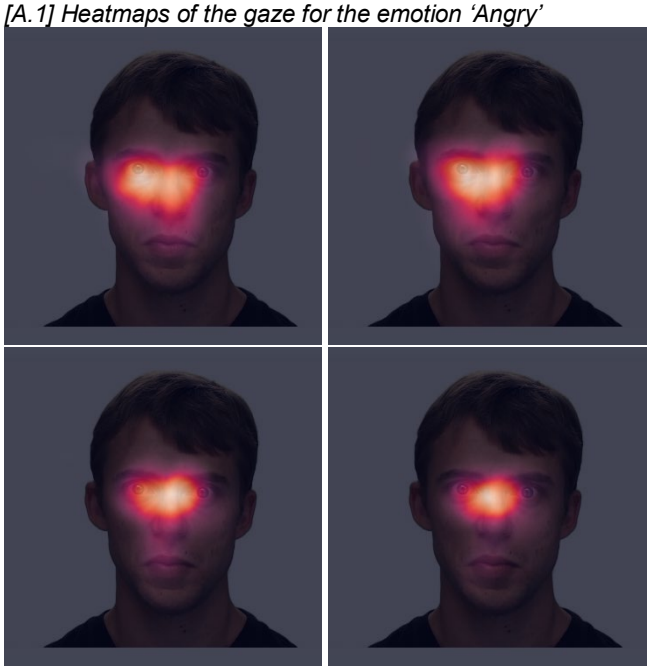

Figure 4.1. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

[A.2] Heatmaps of the gaze for the emotion 'Disgust'

Figure 4.2. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

[A.3] Heatmaps of the gaze for the emotion 'Fearful'

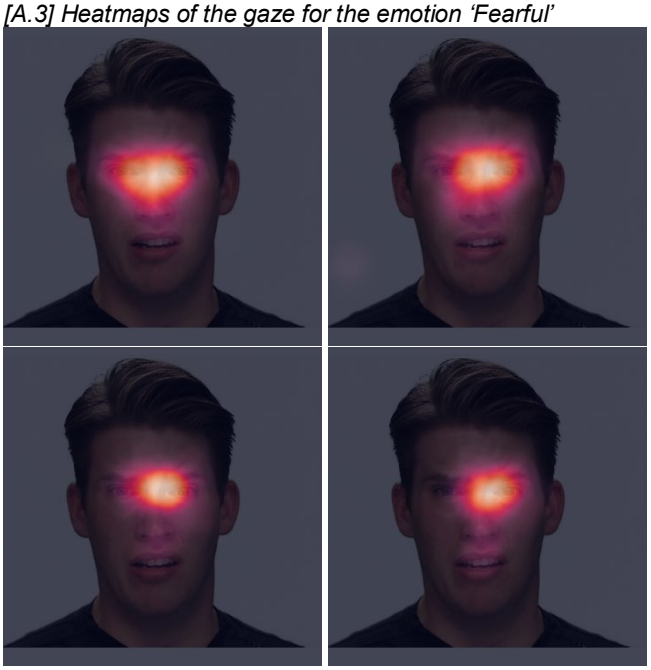

Figure 4.3. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

[A.4] Heatmaps of the gaze for the emotion 'Happy'

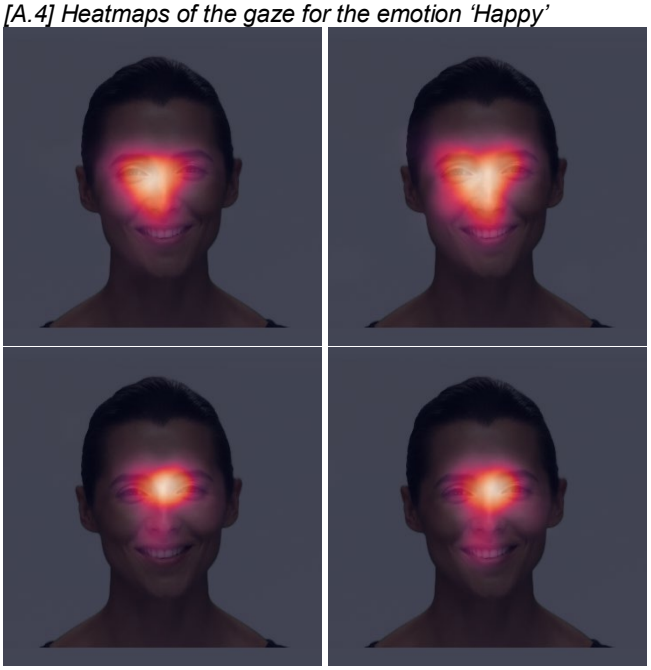

Figure 4.4. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

[A.5] Heatmaps of the gaze for the emotion 'Sad'

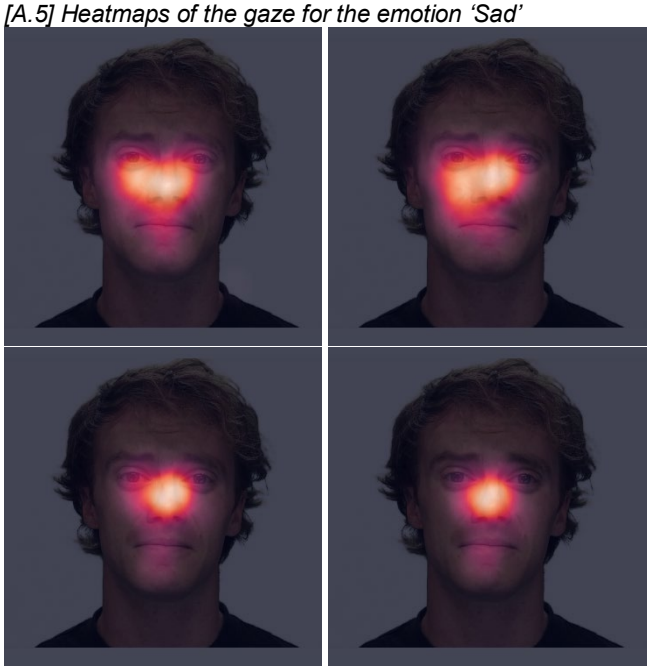

Figure 4.5. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

[A.6] Heatmaps of the gaze for the emotion 'Surprised'

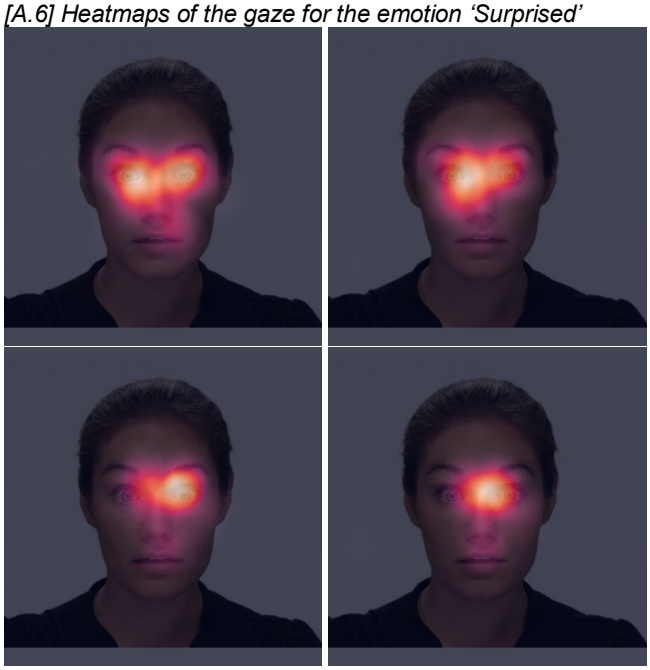

Figure 4.6. Heatmaps of the gaze. Top left: Category As. Top right: Category An. Bottom left: Category Ts. Bottom right: Category Tn.

APPENDIX B

Here you can find our personal reflections.

Anne Bloem - 1238059

After choosing this project, emotion recognition, I was interested in the kind of direction it could take. However, I still was hesitant at the start. This was because of when I received the briefing on the project, I came to the conclusion that it was not a design research project, but a research project. I had my doubts because after having had the course 'Design<>Research', what was given to prepare us for this project, this project felt like a completely different direction. But after having thought about it and discussing it with my fellow teammates, I realized that this new direction could create new and different learning opportunities I may not have had in a design research project.

From the start of the project I found that our team had a good atmosphere, not only a good working atmosphere but a social one as well. This helped us in reflecting upon the teams and each other's work and giving constructed criticism where it was fitted. Furthermore, we tried to work structured and productive by creating agenda's, "to-do" tasks and a planning for the entire project. Since planning can be a weaker area for me, I found this way of creating an overview very helpful. However, this structure did not always work in practise in our group. I believe it was due to having multiple long meetings in a week and not being able to maintain the concentration for such a long time. I noticed this in myself but also in my other teammates. I believe we could benefit in having more structured small breaks, when we are to continue with this project.

In the midst of the project we did a more elaborate reflection session, to be able to tell each other their strengths and weaknesses in the project so far. As strengths, I received that I have an active role in the group, I create a good atmosphere, can be relied on as a group member and give good input/remarks in important meetings. My weaker areas were that I can sometimes talk too much on other subjects and that I could be more active in choosing my tasks. In the continuing of the project I tried to improve upon this and take more initiative when choosing tasks. At the end of the project my teammates told me how I made progress in this area after the feedback I received. However, I have more room for growth and want to be more aware

of this weakness and search for improvements in future projects.

Later in the project we got to the phase of testing our participants. Here we had to gather the participants, explain our intentions of the research test and conduct the tests. All of this had to be done in a way that all participants received the same kind of treatment. For this I was active in both the explanation role as in the conducting the tests. This way of controlled testing made me more aware of how to explain something consistently and what kind of words to use or not to use.

After having conducted the user tests, we had to analysis the gathered data. Before this I had little knowledge of data analysis and the different kinds of techniques used, I only had the experience of the course 'Data Analytics for engineers'. Because of this, I wanted to try, help and improve myself in this area. Firstly, I took the task of making the confusion matrices of our gathered data. Even though I did not know when starting this task what a confusion matrix was, with the help of tutorials and tips from my fellow teammate Romain, I was able to understand and complete de confusion matrices. Later, after the meeting with Matthias Rauterberg, the group got the advice to use spss to check the reliability and visualize our test results. Using the program spss was a new experience for my teammates and myself. However, after looking at the tutorials together with my teammate Romain, I found the program understandable to use and apply to our reliability check and visualisation of the data. However, I could not comprehend all the terms that were used for and in spss. This is something I want to do more research in, in order to get a better understanding of the program. This especially could be helpful if I choose to continue this project next year.

In conclusion, in this project I worked in a team that collaborated well, learned how to conduct a proper research and I was able to grow significantly in my personal and professional development.

Inge Hootsmans - 1252208

This reflection is going to be about my experiences with my research project in the "Health" squad DPB220. It will go on about the writing of a research paper, organization and planning, teamwork and communication, my learning points, and integration of expertise areas.

Starting with writing a research paper. I already had some prior knowledge on writing a research paper from the course Design <> Research DDB100. However, having paper writing workshops and peer reviews, provided by our squad, helped me improve my writing skills. My most important learning point is; to keep everything structured in your paper. I find that having a structure helps keep my mind at ease and it gives me a good overview of all the tasks at hand. What also helped is the way the workload was divided between the team members in my group, two people were assigned to a specific section. Once a section was 'finished' the other two would leave comments and remarks, then the first two would adjust the section and finish it. When all sections were checked, we went through the whole paper with the group and made sure it was consistent. Even though we maybe could have done a better job at time management, as we had a lot to do in the last few weeks, I find this way of working very efficient. Doing a peer review and receiving feedback was really helpful. Receiving feedback from outsiders proved to be valuable, as they gave a honest opinion on whether the paper was well written and understandable. I will remember to let outsiders read and review my reports for future projects.

The teamwork and communication within my group was good. As I prefer to keep things organized in a project, I set up a planning, this gave the project structure. We also held productive meetings, since we clarified what we were going to discuss beforehand. This way of meeting was efficient, as less time was wasted on unnecessary points. After the midterm we held an evaluation meeting, which improved the group dynamic. We started to work with 'tiers', which means we divided tasks in: 'most important' (first tier) to 'nice to have' (last tier). This gave us a big overview of the tasks at hand and made it easier to plan and divide tasks.

From the evaluation meeting I also learned some things about myself. I am usually the one who will take on extra work, when someone in the group mentions they are busy. Even though, I do not mind working on a project I like, my group reminded me that by doing so, some team members have little to no work. From then on we agreed that I should say 'no' to too many tasks. Looking back, I think I improved on this as I took a sufficient amount of tasks upon myself in the weeks after the meeting. Next, I also got feedback on being more

honest towards the group. I can be quite reserved when it comes to saying how I feel, as I do not want to make anyone feel bad. However, this resulted in the group guessing on how I felt. By being more open, I experienced the group becoming more flexible and understanding of what I want. Last, from time to time I can be a bit impatient. This means I sometimes tend to rush through things, and might miss some important details. I think this is still a point I can improve upon.

I also improved in different expertise areas. I had set a goal for myself; get a better understanding of different user tests, connecting to the User & Society area. During this study I have been involved in making the test, arranging participants, and conducting the test. A learning point was the ethical responsibility you have as a researcher when working with a group of individuals who have a special condition. Within this project I definitely improved my data analytics and statistical skills. I learned more about data analyzation with Python and how to work with the data analyzing program SPSS. This connects to the expertise area Math, Data, and Computing. Since my team wants to continue with this project in the upcoming semester, I do think I should invest time in getting an even better understanding on how statistics work. In this project I wanted to improve my presenting skills in the form of posters and flyers. I learned how to present a research project on a poster, making sure enough information is given and keeping it visually attractive. This connects to the area Creativity and Aesthetics.

Concluding, key learning points for me were; to not take too many tasks upon myself, be more open on how I feel about certain points, and be more patient, so that I do not miss any important details. To keep a task division more structured I learned of the 'tiers'-technique, I will definitely use this technique more often for future projects. I experienced that having evaluations from time to time with the team can lead to more productive meetings. In future projects, I will let outsiders give feedback on my reports, think of my ethical responsibility as a researcher and designer, and know how to present a visually attractive research project. For the near future, I want to improve my statistical analyzing skills.

Lena Opheij - 1232242

This is a reflection for the course DFP004: Health, research project. I will address my goals, personal development and future progress.

Before starting this course there were some goals I set up during the second design project and the course Design <> Research.

First, I wanted to learn how to identify a target group correctly. What are the specifics and more importantly how to approach them. Second, I want to be able to find a niche in already existing research and design more easily. Third, I want to learn how to visualize gathered data well.

The group was a pretty well functioning machine during the research project. Throughout the project, we planned reflection sessions with the group to be able to discuss what was going well, what was not, and how we could improve it. This helped to keep me motivated and allowed for a feedback moment not just at the end of the project, when it is too late to change something. During one of these feedback sessions I was advised to be more active in searching for tasks rather than waiting to have them assigned to me. At the end of the projects, the meetings were longer and more frequent. This meant that we were meeting for at least one entire day in the week, sometimes even two. With my schedule it was impossible to spend another (half a) day on the project. Not being able to spend as much time on the project than the others made me feel not as invested as the rest.

By writing the related work section all by myself, I have developed my second goal immensely. During this project, two group members often did not agree with each other. This took down the efficiency of the meetings. I often felt like we were wasting our time, because the entire reason for the discussion was insignificant in the first place. So, I have gotten confirmation that I easily take up the mediator role in a group.

Based on the knowledge I gained from this research project I have set up some goals for the future. First, I want to learn more about statistics. Over the course of this project, especially during the data analysis, I noticed that understanding why you have to perform a certain test or visualize data in a specific way is really important. This is not only make the validity of the project

stronger, but also to fully understand why you do what you do. I want to improve this during my internship. Second, I did not get the chance to visualize the data. So, that is an aspect that I intend to develop during the internship as well. Third, I struggle with receiving feedback on tasks that I have performed individually. This is an important aspect to develop more to be able to receive constructive comments.

Romain Toebosch - 1235294

Since the beginning of this Bachelor we have been thought to look ahead and start thinking about vision and where we want go with our professional identity. I thus laid out a plan to myself, and stuck to it as much as possible. The assumptions that research was boring to me and that I would not like anything health related were quickly made here, as I had little experience with both. This project proved me wrong on both.

Not only did the Health squad have plenty of topics that would interest me (and a lot of great people to work with), this research project turned out to be one of the most interesting projects I have done so far. I have been confronted before by the fact that my plan might be too rigid and that I did not give myself enough room for other possibilities, but this project truly showed me that I should better use this Bachelor as an opportunity to explore, and get a taste of as many different topics as possible, to truly get to know what I would like to do.

Next to this shift in personal development, I have further developed my skills and expertise areas as well. One of the more obvious developments was the *Math, Data & Computing* expertise area: one year after having had the basic course *Data Analytics*, the first real world application came where data analysis was needed; mistakes have been made during this analysis, but much has been learned as well. Firstly, a I have learned a lot about different (more complex) statistical tests and in what situations to use them, which has helped a lot in the understanding of data analysis in general. Secondly, the discovery of SPSS and how to use it was a pleasant one, as it helped a lot doing more structured analytical tests in a more efficient way. In the area of *User & Society* the most notable development has been done in how to work with specialty groups like people with ASD, but I have also learned a lot about user testing in general.

Furthermore, I have learned about conducting research in general. Although having had *Design <> research* in Q2, only now do I feel like I properly understand the basics of doing research and paper writing, having experienced both of them firsthand in a more elaborate manner. During this project I also set myself a goal to learn more about optimization of teamwork and how to work in a more structured manner, and immediately implemented it. I felt like this helped the team get certain things done better and quicker (even though we already were on the right track). Something which I felt helped a lot as well were the group reflection moments we planned in over the course of the semester, as this gave an opportunity for everyone to get feedback on how they worked within the group and what they could do better, while preventing (smaller) irritations to form.

There are however still many points of improvement for me. The most notable being that once I have made up my mind about something, I have a lot of difficulty letting it go. This can help in some cases, but most of the time it results in having long arguments about minor details. I have given myself the task to, whenever I come up one of these situations, wait for ten seconds and then re-evaluate whether the subject is worth the argument, and if I am thinking rationally. I hope this will help, and if not, I will try new strategies. Furthermore, I want to further investigate ways to enable teams to work better and more structured together, I have seen what some simple good habits can do to productivity, and would like to try more of these. This increase in productivity (and thus decrease in time needed to get work done) could help with a different goal I have set myself: look better after my own health. Too often have I caught myself working late, and consequently not sleeping enough. I should stop treating 8 hours of sleep a day as a luxury and more as a necessity to stay healthy and function better in everyday life. I have thus set myself a goal to improve my sleeping pattern during my internship period, and go to bed and get up at fixed hours.

Overall, I am happy to have had this project. It has been a real eye-opener, and helped a lot in my development on a personal and professional level. This could of course not have been possible without my team, and I thus want to thank them for that.