Robustness of current reading models to incorporation of word-emoji substitution

P10 Project
Group 1086
Engineering Psychology
Aalborg University
6th of JUNE 2019

Title: Robustness of current reading models to incorporation of word-emoji substitution.
Semester: 10th semester
Semester theme: Master theses
Project period: 01.02.19 to 06.06.19
ECTS: 30
Supervisor: Nikita Kharlamov
Project group: PDP 1086

SYNOPSIS:

In this study the phenomenon of word-emoji substitution was explored with the intention of determining if common models of reading would be able to accommodate for word-emoji substitution. To that avail it was explored if emoji were subject to preview benefit the same as words, and if other position dependent artefacts could be found. Furthermore the results would be compared and discussed in concert with earlier studies pertaining the same subject. To this extent the two earlier studies were reviewed and two experimental paradigms were constructed. The first paradigm was a repeat of a paradigm from one of the earlier studies with the intention of reproducing the results. This paradigm was set up to ascertain the comprehension time of paired single words and emoji. The second paradigm was a reading paradigm in which short sentences with word-emoji substitution were read and questions about the sentences were answered afterwards. This paradigm served the purpose of determining positional artefacts and ascertain if emoji were subject to preview benefit.
The results from this and earlier studies show that emoji only seems to affect the comprehension time, where single emoji are more quickly comprehended than words, and otherwise seem to express the same tendencies as words when it comes to reading speed, backtracking, fixation time and preview benefit.

Alexander Flyvholm Povlsen

Kasper Horslev Ravnkilde

Number of pages: 130
Appendices: 88

By accepting the request from the fellow students who uploads the study group's project report in Digital Exam System, you confirm that all group members have participated in the project work, and thereby all members are collectively liable for the contents of the report. Furthermore, all group members confirm that the report does not include plagiarism.

Contents

1 Introduction

Written language consists mainly of characters that are found within the alphabet of the language in question and when these letters are used in conjunction with each other words can be formed to relay information from one person to another. However it can be difficult to convey certain things with the written language, such as sarcasm. Scott E. Fahlmann tried to remedy this by making a smiley face out of a colon, a dash and a right parentheses to denote that a forum post was not to be taken serious by the reader [Fahlman, 2007]. This was the birth of the first emoticon which was later translated into small pictures called emoji. Throughout the years this evolution has continued and from emoji mostly depicting faces they now encompass all sorts of different things including sports, foods and flags. The Library containing all of these emoji are updated regularly and maintained by the Unicode Consortium [Unicode.org, 2019].

In October 2011 and July 2013, iOS and Android respectively, released their emoji keyboard, which then started an incredible growth in the usage of emoji on social media platforms, i.e. instagram [Engineering, 2015]. This growth in use of emoji can be seen on Figure 1.2 where the release of the iOS emoji keyboard resulted in an explosive increase in the use of emoji, and from the introduction of the Android emoji keyboard, this trend has been continually growing. This can also be seen on the fact that Apple introduced a function on their emoji keyboard in 2016 [?], to effortlessly substitute words and emoji with a single tap, see Figure 1.1.

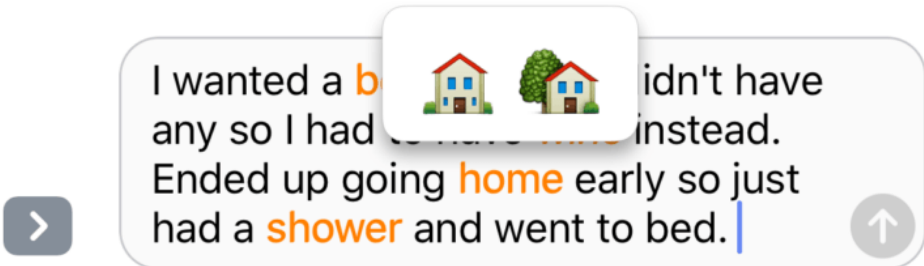

Figure 1.1: Word-emoji substitution as enabled by iOS 10 [Bell, 2016].

This increase in the use of emoji, and change in how it is used as either an accent to a sentence or replacements for words has lead to it being incorporated more and more into the vocabulary of written language on social media and the internet.

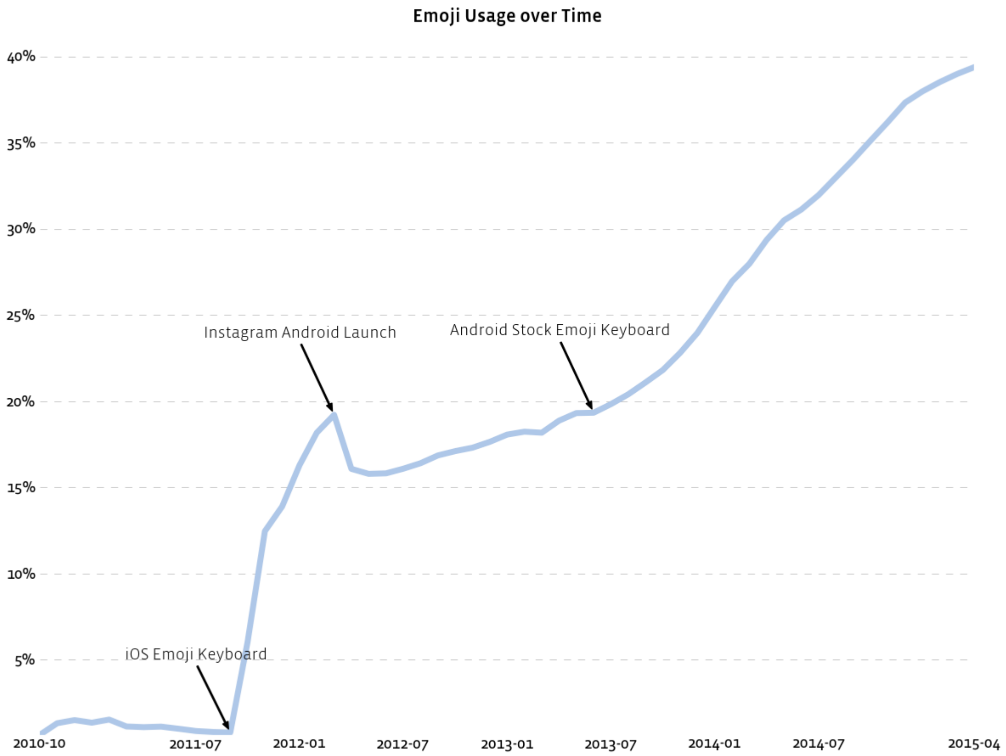

Figure 1.2: Emoji usage on instagram from 2010 to 2015 [Engineering, 2015].

As a testament to how used emoji are it can be seen that both marketing firms and commercials are beginning to use them as a tool [EDUCBA, 2016]. However not always as effectively or as successfully as they could be, see example below in Figure 1.3. Here Chevrolet used emoji to replace all words in a press release rendering it almost incomprehensible which necessitated that Chevrolet released a video to explain what it meant [EDUCBA, 2016].

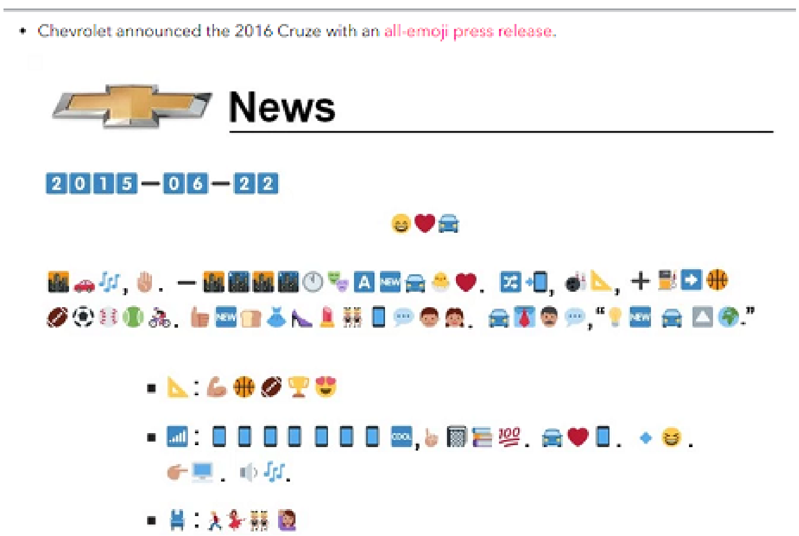

Figure 1.3: Chevrolet press release on the Cruze 2016 launch in all emoji [EDUCBA, 2016].

As seen on the figure above, Figure 1.3, emoji have a hard time standing alone without text to support it. As such it could be said that emoji only count as half-words that can be used to support normal words but will obstruct meaning if too many are used at the same time [EDUCBA, 2016]. Another example of how marketing are using emoji, though a bit more successfully is in email subject lines, as seen in Figure 1.4, where according to EDUCBA [2016] emoji in the subject line can boost the opening rates of the emails.

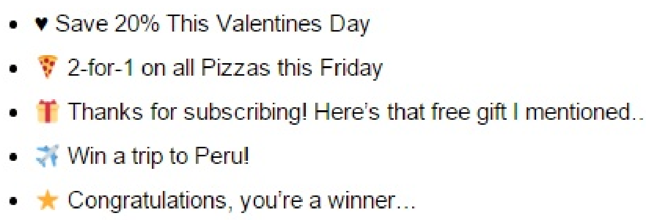

Figure 1.4: Emoji usage in email subject lines [EDUCBA, 2016].

Since emoji usage have been on the rise for the last decade, and is climbing its way into marketing, it inspires curiosity into the way emoji are read in sentences and if current models of reading can contain the reading of emoji in sentences.

2 Eye movements and models of reading

This chapter will introduce the basic concepts of eye tracking and eye movements, such as fixation and saccade, together with three models of how reading functions and what the eye movements during reading can tell about the cognitive processes.

2.1 Basics of eye tracking

When looking at eye movements and eye tracking there are some basic concepts that are important to know. First off the terms saccades and fixations which describe the two states the eye can have. Saccades are the movements the eyes make when a persons focus is shifted towards something new [Weaver and Holmes, 2012], this movement typically takes 25-50 ms on average. During saccades no visual information is taken in due to saccadic suppression, which means that all information is obtained during fixations [Rayner et al., 2012,p.91]. Fixations are the part of the eye movements where the eye doesn't move and takes in information about what it is looking at, this typically takes 200-250 ms on average [Weaver and Holmes, 2012].

Next, visual span is a key component of understanding how and why eye movements function as they do. The visual span of the human eye can be divided into three different regions called the fovea, the parafovea and the periphery [Rayner et al., 2012,p.9]. The fovea is the area with the highest visual acuity and spans about 2° around the center of focus. The parafovea extends from the fovea and out to about 10° around the center of focus and lastly the the periphery extends from here and beyond. Visual acuity drops off the farther you get from the fovea into the parafovea and the periphery [Rayner et al., 2012,p.9], see Figure 2.1 for an illustration of the three visual regions.

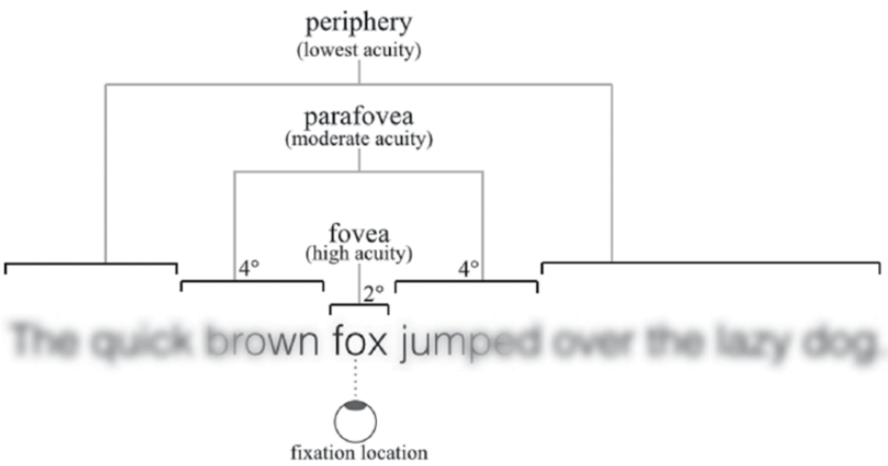

Figure 2.1: Illustration of the three regions in visual span [Rayner et al., 2016]

When looking at reading specifically the visual span is skewed heavily in the reading direction such that it extends from the start of the currently fixated word, by some equated to 3-4 character spaces to the left [Rayner et al., 2016,p.10], and about 15 character spaces to the right of it, this is called the perceptual span [Rayner et al., 2012,p.115].

The fixations and saccades are the primary measures used by reading models to predict and estimate cognitive processes. However when actually recording these phenomenon using eye tracking there is only one measure and that is point of gaze or POG. Most remote systems use a method of detecting the POG called pupil centre corneal reflection (PCCR) [Guestrin and Eizenman, 2006; Farnsworth, 2019; TobiiPro, 2019], which uses the estimated centre of the pupil and one or multiple light reflections on the cornea to estimate the visual axis of both eyes, see Figure 2.2, and in combination triangulate the POG onto an object in 3D space [Guestrin and Eizenman, 2006]. The light used to create the reflections is usually infrared or near infrared as this doesn't distract the person and makes it so the sensor used to record the data can be sensitive to a specific wavelength exclusively and doesn't get influenced by ambient light [Farnsworth, 2019].

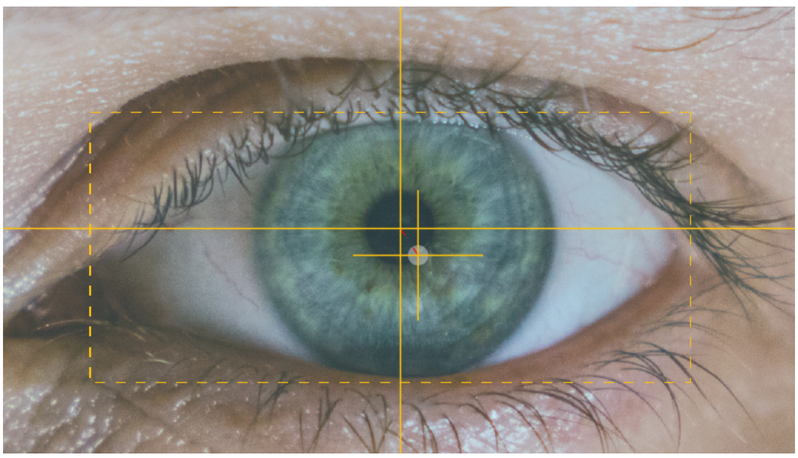

Figure 2.2: Pupil centre corneal reflection (PCCR) illustration [Farnsworth, 2019]

2.2 Models of reading

To better understand the act of reading, different models of eye movement will be looked at in this section. First off the E-Z reader model will be taken into account, as this is one of the most influential models of eye movements in reading [Rayner, 2009a,1474]. Furthermore the SWIFT and the EMMA model will be examined, but as to avoid repetition these two will be described in light of the E-Z reader model and only points where they differ will be mentioned.

2.2.1 E-Z reader model

The E-Z reader model have two core assumptions. First, that reading is a serial event, which means that words must be processed one at a time in an order determined by the reading direction of the specific language such that word N is processed before word N+1 [Reichle, 2013,p.775]. Secondly, that the process of reading necessitates that attention is focused on the word that is currently being processed, as to encode the word as a word object that can be processed for its orthographic, phonological and semantic values[Reichle et al., 2005,p.5]. While processing of a word necessitates that attention is focused on that word it does not imply that the eye is fixated on that same word. This is because saccading is decoupled from shifts in attention, such that the fixation of the eye can remain on word N while attention moves on to word N+1. This division of attention and fixation is what allows for preview benefit which makes it possible to start the processing of word N+1 before the eye moves to fixate on it, see Figure 2.3. This can be an explanation for why the last words in sentences often aren't fixated, however Rayner et al. [2012] states that first words often aren't fixated as well. A possible explanation for this could be that attention shifts covertly leftward of the first fixation (second word) in the new line and takes advantage of the visual span to the left of fixation to also process the first word [Rayner et al., 2012,p.115].

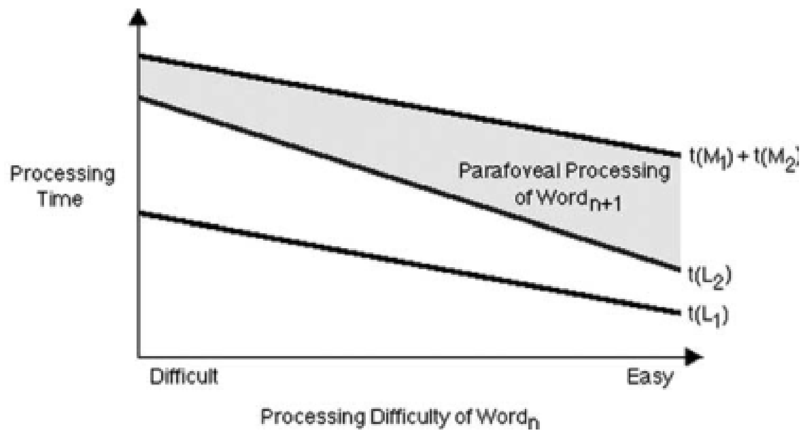

Figure 2.3: A graphical representation of the relative time course of lexical processing and saccadic programming [Reichle et al., 2005,p.9]. As saccades are initiated from the familiarity check(L1) the time for saccade completion runs parallel to this graph. Parafoveal preview benefit is thus only affected by the processing rate of lexical access.

The process of handling the word objects mentioned earlier are divided into two steps, first the familiarity check(L1), point 4 on Figure 2.4, and afterwards the lexical access(L2), point 5 on Figure 2.4, both of which operate with high spatial frequency information, point 3 on Figure 2.4. The familiarity check(L1) functions as a rapid assessment of the word, based on its predictability in the sentence structure and the frequency at which it occurs in text media. If the word is familiar to the reader a signal is sent to begin planning of a saccade as lexical access for that word is imminent [Reichle et al., 2005,p.6]. This means that a saccade can begin to be programmed in preparation for the eye to move on to the next fixation. After the completion of the familiarity check the word object is processed for its orthographic, phonological and semantic values to achieve lexical access(L2) and when this happens the attention can be moved on to the next word in line. A factor that can prolong the time L2 takes to finish is if the word is long or far from central vision, both factors diminishing the visual acuity of the word being processed and as such making the processing slow or impossible which would prompt a refixation at a more optimal viewing position. The last stage is the integration of all the words that have been processed so far, into a coherent sentence [Reichle, 2013,p.776], see point 6 on Figure 2.4. This is a post-lexical process and as such does not affect attention or fixation. However in cases with integration difficulty it can result in either longer fixations or regressing saccades (Backtracking). This is in most cases brought on by word N+1 being identified before the integration of word N has concluded, misreading a word or a word being semantically incoherent with the rest of the sentence [Reichle, 2013,p.776].

For the programming of saccades the E-Z reader model operates with a two step process that consists of a labile program(M1), point 7 on Figure 2.4, and a non-labile program(M2), point 8 on Figure 2.4, [Reichle et al., 2005,p.9]. The labile program(M1) consist of two sub stages of which the first is a general preparation of the oculomotor system to make it ready for saccade programming, point 7a on Figure 2.4. In the second sub-stage the coordinates for the targeted word are translated into an eye movement, point 7b on Figure 2.4. The labile program(M1) takes about 100 ms on average to complete, during which the saccade can be overwritten at any point [Reichle et al., 2005,p.9]. The non-labile program(M2) cannot be overwritten or cancelled when it has started and as such the saccade will always be completed if this step has been reached, this takes 25 ms on average [Reichle et al., 2005,p.9]. This means that the act of completing a saccade from start to finish takes approximately 125 ms plus the 50 ms of eye-to-mind lag, giving a constant time for saccade execution of 175 ms on average [Reichle et al., 2005,p.9].

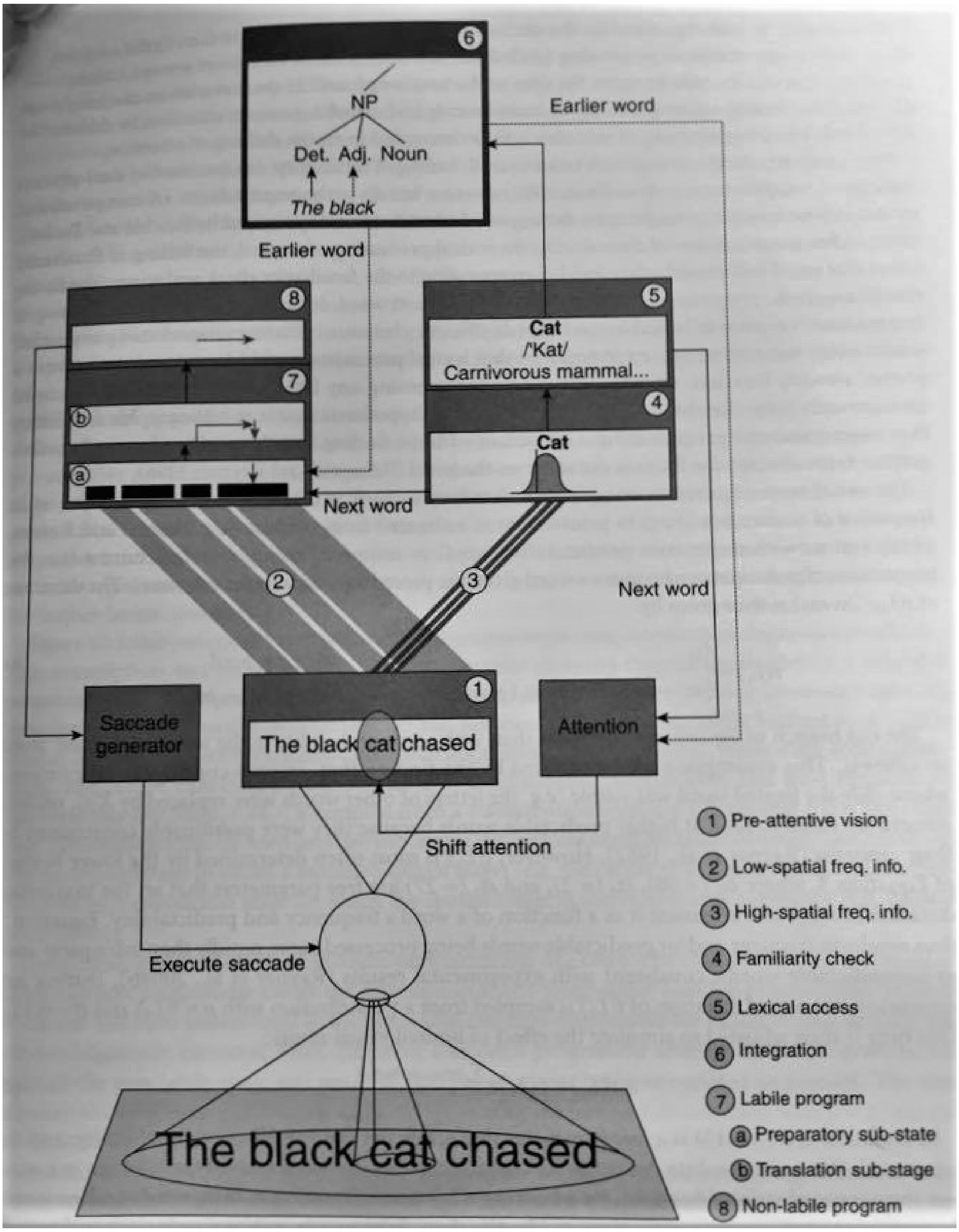

Figure 2.4: Graphical representation of the E-Z reader model [Reichle, 2013,p.773]

2.2.2 SWIFT model

When looking at the SWIFT model there are two key differences from E-Z reader. First off E-Z reader is categorised as a serial attention model or SAS in short, whereas the SWIFT model is categorised as a gradient by attention guidance model or GAG [Rayner, 2009b]. This means that the SWIFT model allows for parallel lexical processing where E-Z reader handles lexical processing serially. GAG models handles this parallel processing by assuming that there is a distribution of lexical processing rate across all the words that are currently in the region that is fixated on [Engbert et al., 2005]. The SWIFT model specifically proposes that up to four words can processed at a time and that the distribution of lexical processing rate is highest on the the word fixated on at the moment and decreases on word N+1 and N-1, as well as some lexical processing also taking place on word N+2 [Engbert et al., 2005].

Secondly, the SWIFT model operates under the assumption that saccades are generated autonomously [Engbert et al., 2005], where the E-Z reader model says that the eye movements through the text is driven by lexical processes [Rayner, 2009b]. To account for longer or more difficult words the SWIFT model employs an inhibitory process to delay the movement of the eye and by doing so extending the processing phase [Engbert et al., 2005].

2.2.3 EMMA model

When looking at the EMMA model (Eye Movements and Movements of Attention) it is very similar to the E-Z reader model described earlier, however it's base assumptions have been simplified to make the model as minimalistic as possible [Salvucci, 2001,p.205] [Salvucci, 2001,p.208]. However, there are some key differences between the two models. First off, the EMMA model is a domain independent model that doesn't specialize specifically in reading as the E-Z reader model does. This is expressed through the EMMA model's use of degrees of visual field to indicate saccade lengths and measures between objects where E-Z reader uses character spaces [Salvucci, 2001,p.208]. Secondly, in the EMMA model the processing of the word and the programming of the saccade to the next target starts at the same time [Salvucci, 2001]. The process of programming a saccade in E-Z reader is only initiated when the familiarity check has completed which means that the processing of the word is ongoing when saccade programming initiates [Reichle, 2013] [Salvucci, 2001].

3 Review of earlier studies

In this chapter previous studies made by the authors about the same phenomenon of word-emoji substitution will be reviewed, as to make a comparison between all the studies later.

3.1 6^{th} Semester bachelor project

The study Ravnkilde and Povlsen [2017] was a first attempt at figuring out if replacing words in text with emoji affected the reading of the text, and as such it lays out some of the ground work and exploration of what an emoji is and how it fits into current reading models. The experimental setup was done with the natural environment of emoji in mind, and therefore it was chosen to present the stimuli in a chat window.

The Hypotheses in Ravnkilde and Povlsen [2017] is:

1. Sentences with no emoji takes longer to read than sentences with words replaced with emoji.
2. Sentences with emoji as accents will take longer to read than sentences with words replaced with emoji.
3. Sentences with emoji as accents will be faster than sentences with no emoji.

All of the hypotheses assumed that there would be a change in the reading speed of sentences with and without emoji, and none were found. However it was found that when answering questions about the sentences, the ones with emoji in had a significantly faster answer time than the ones with no emoji. However this was probably due to the sentences and questions being available simultaneously. This could however indicate that emoji facilitates visual search to a greater extent than plain text, which makes sense as emoji contrasts the other words greatly by being visually different in many aspects.

3.2 1^{st} Semester master project

The study Jensen et al. [2017] was another attempt at exploring what effects word-emoji substitution had on reading. As the first study [Ravnkilde and Povlsen, 2017] was based on the EZ-reader model, the same model was used for this study as well. For this study 3 experimental setups were carried out, first, a task to ensure that the subjects could correctly identify what item an emoji was representing (i.e. a naming task), second, a sorting task with a subset of the emoji used in the naming task and third, a longer reading task consisting of 1500 character texts with word-emoji substitutions implemented.

The hypotheses in Jensen et al. [2017] are as follows:

1. The substitution of words with emoji will slow down the reading speed of the subjects.
2. The substitution of words with emoji will result in backtracking around the emoji.
3. The substitution of words with emoji will result in increased fixation time.
4. There is no difference in comprehension time between words and emoji.

The naming task was mainly used as a preliminary test to explore if the subjects had the same understanding of the emoji stimuli as the experimenters anticipated. This understanding would be critical for the validity of the data in the sorting task because right and wrong answers were predetermined. The naming task yielded 0.9% wrong answers which indicated that the participants understanding of the emoji was close to the same as the experimenters anticipation. To answer the 4$^{\text{th}}$ hypothesis a sorting task was made and consisted of two different categories of emoji: Fruits and Animals. The subjects were asked to indicate if the emoji or word they were presented with were part of the target group (Animal or Fruit) by answering Yes or No. For the Animal targets the amount of wrong answers were **2%** and for the Fruit targets it was **7.5%**. This was an early indication that the Fruit targets were a little different.

For the Animal targets, the median for the emoji comprehension time was **706 ms** with a **95% CI [703, 737]** and for the text stimuli the median was **740 ms** with a **95% CI [736, 771]**, this data is illustrated on Figure 3.1. To compare the two conditions a Wilcoxon signed rank test was used yielding a result of **p-value** = **0.018** which indicates that there was a significant difference. This significance warranted the calculation of an effect size, in this case using Cliff's delta, yielding **d** = **-0.565** which indicates that there was a large difference between the emoji and text stimuli with the Animal targets.

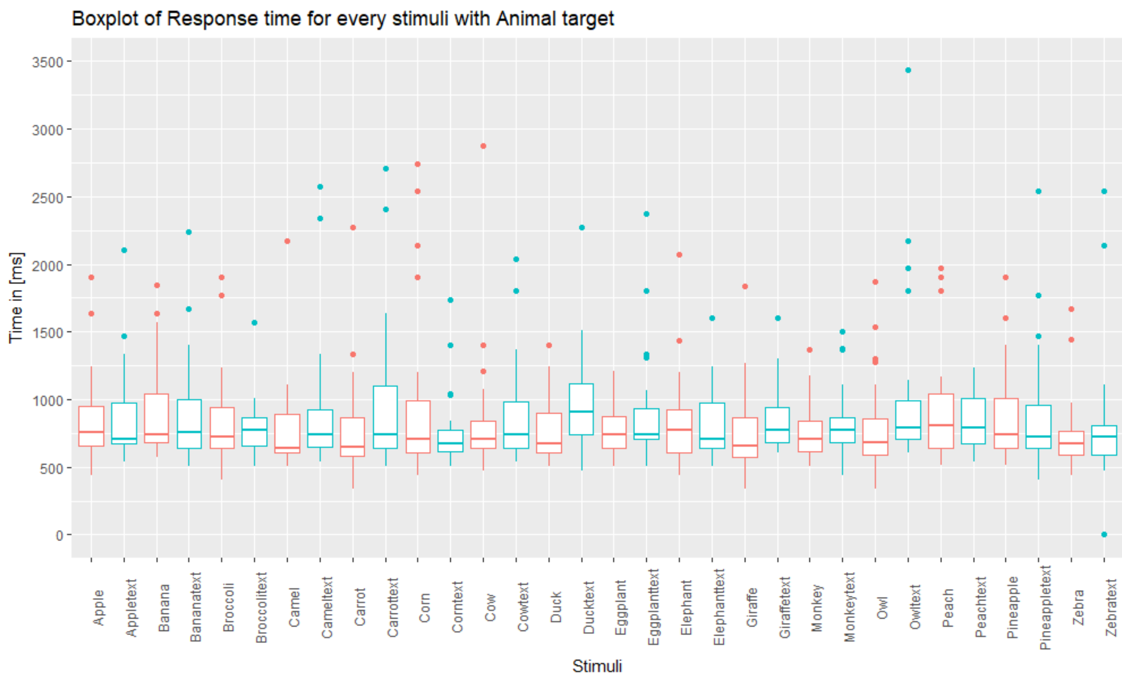

Figure 3.1: The boxplot shows data from all of the stimuli from the Animal target task [Jensen et al., 2017]

For the Fruit targets, the median for the emoji comprehension time was **906 ms** with a **95% CI [870, 950]**, and for the text stimuli it was **940 ms** with a **95% CI [905, 1004]**, this data is illustrated on Figure 3.2. Here a Wilcoxon signed rank test was also used to compare the two conditions yielding a result of **p-value = 0.258** which indicates no significant difference between the comprehension times of words and emoji. This again indicated that there was a difference between the two conditions in the setup.

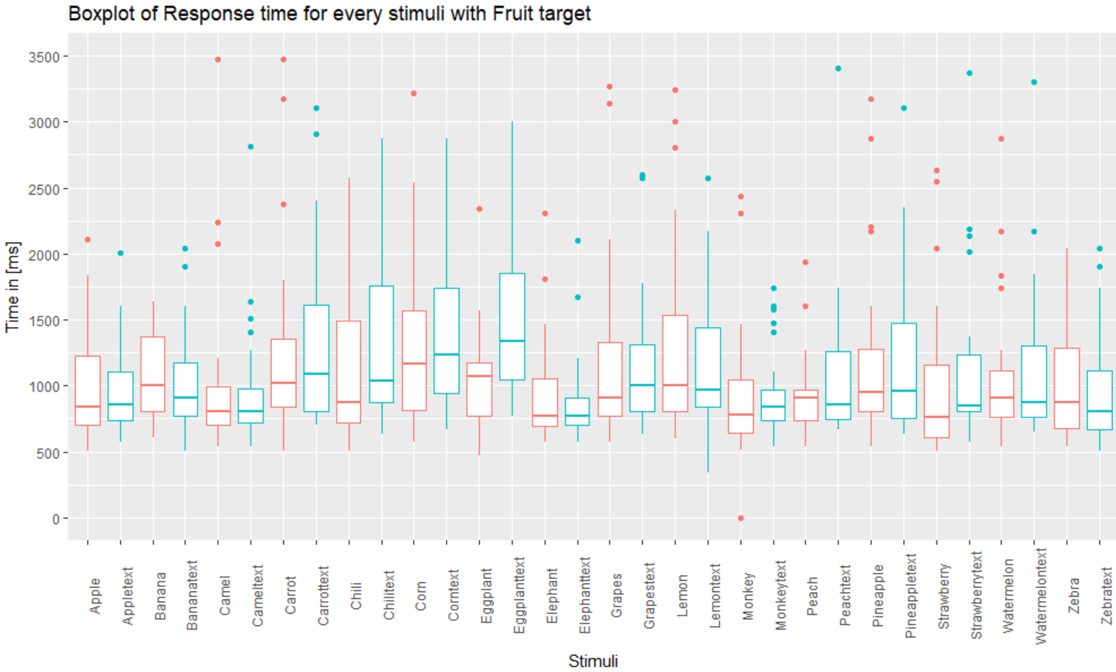

Figure 3.2: The boxplot shows data from all of the stimuli from the Fruit target task [Jensen et al., 2017]

The difference between the Fruit and Animal conditions were attributed to the distractors used in both of the tasks. In the Fruit task the distractors of vegetables where more visually similar to the target group, than they were in the Animal task. In the end the results of the Animal targets, seen on Figure 3.1, showed that there was a tendency for the comprehension time for the emoji being faster than the one for text. This result then both rejects, for the Animal targets, and confirms, for the Fruit targets, the 4th hypothesis of there being no difference in comprehension time between text and emoji. However the comprehension times for for the Animal targets was probably a better representation of a precise comprehension time as the Fruit targets were influenced to a greater extend by the distractors due to visual similarity.

To test hypothesis 1, 2 and 3 a reading task was made during which the subjects were presented with a 2666 character pretext for establishing reading speed, and 4 texts with about 1500 character as stimuli. Two of these texts were made to contain 5 word-emoji substitutions and the last two with no alterations. Each of the texts were followed by 5 questions to measure the subjects comprehension of them.

The average pretext comprehension rate was 56% and was used as the base line for comparison with the stimuli. For the texts with emoji in there was a 53% comprehension rate. A Shapiro-Wilk test was made for the reading speed and yielded a **p-value = 0.016** with a median of **199 WPM** and a **95% CI [177, 283]**. For the texts without emoji a 61% comprehension rate was found. The Shapiro-Wilk test for these stimuli yielded a

p-value = **0.041** and a median of **197 WPM** with a **95% CI [172, 324]**. Reading speeds were compared with a Wilcoxon signed rank test and with a **p-value** = **0.742** showing no difference between the two conditions which results in a rejection of hypothesis 1.

To try and answer hypothesis 2 the eye tracking data was perused for backtracking, where after means were found for all the stimuli. For the texts without emoji the mean amount of backtracks found were **M** = **198.5** and with emoji it was **M** = **234**. However this increase does not seem to happen specifically around the emoji and therefore the hypothesis is rejected. Another curiosity discovered through the eye tracking data was that emoji seemed to support preview benefit, however no statistical evidence was found for this as it wasn't the focus of the test setup.

To test hypothesis 3 a random subset of words from the stimuli were chosen to compare with the emoji. A Shapiro-Wilk test was made yielding a **p-value** = **0.479** for the accumulated fixation time on the emoji with a mean of **M** = **322 ms**, **SD** = **150** and a **95% CI [271, 402]**. Likewise a Shapiro-Wilk test was made for the accumulated fixation time of the subset of random words yielding a **p-value** = **0.429** with a mean of **M** = **323 ms**, **SD** = **217** and a **95% CI [188, 458]**. For the comparison of the two conditions a Welch two-sample t-test was used yielding **t** = **0.171**, **df** = **13.2** and a **p-value** = **0.867**, which rejects the hypothesis.

All of these results indicate that emoji show some of the same tendencies as words, for example by exhibiting tendencies towards preview benefit. Tendencies also show that the stimuli with emoji performed as good or slightly better than the control group without.

3.3 Review conclusion

In review it was seen for both the setup in section 3.1 and section 3.2 that there was no measurable difference in the reading time when the stimuli contained emoji. Furthermore it was found that word-emoji substitution did not contribute to an increase in backtracking or fixation time. Lastly the results showed a tendency towards comprehension time for emoji being significantly faster compared to the one for words.

Due to these findings a new working hypotheses was formulated to explore other aspects of the phenomenon.

3.4 Hypotheses

For this project three areas of focus was chosen. First off, to explore if the reading process breaks down from the introduction of a logographical element such as an emoji as the first part of the sentence as earlier studies have shown little to no effect from the addition of such elements when the reading process is ongoing [Jensen et al., 2017]. Secondly to explore if emoji are subject to preview benefit, as such tendencies have been observed previously [Jensen et al., 2017]. Lastly, the sorting task from Jensen et al. [2017] will be reproduced to see if similar results can be obtained.

With this in mind it is hypothesised that:

1. Introduction of an emoji as the first part of a sentence will result in an increase of fixation time.
2. Emoji introduced by word-emoji substitution will facilitate preview benefit.
3. Comprehension times for emoji will be faster than the comprehension time for single words.

4 Test design

The test presented in this report will be divided into two separate paradigms. The first, will be a repeat of a comprehension task from an earlier study by the study group [Jensen et al., 2017] similar to the procedure used by Balota and Chumbley [1984] in the first part of their experiment to explore comprehension times of words dependent on their frequency of occurrence in writing. The second part will be a reading paradigm that will be used to explore if there are any sentence position related artefacts from preview benefit or lexical breakdown when substituting words with emoji. Both paradigms will be explained in greater detail below.

4.1 Participants

As this study concerns itself with reading in general, there is no specific requirements for the participants other than have danish as their first language and normal or corrected sight. **20 students** from Aalborg University participated in the experiment, **12 male** and **8 female** with a **mean age = 24.5** and a **SD = 2.2** and a **range from 22 to 32**. Due to unforeseen circumstances 2 of the participants have been excluded from the eye tracking data.

4.2 Apparatus

Both tests were set up in the PsychoPy 3.0.7 builder and all eye tracking data where gathered using the SMI Experiment Center and analyzed with SMI BeGaze. PsychoPy is an extension of the python programming language, in which it has been made very easy to build test setups with the GUI provided by Peirce and MacAskill [2018]. SMI is a company that sells eye tracking equipment with software solutions to assist in the data collection and processing of the data, these include SMI Experiment Center and SMI BeGaze. The SMI Experiment Center is SMIs own software for programming eye tracking experiments. The SMI BeGaze is the analytic tool used in cooperation with the SMI Experiment Center for statistical analysis.

For hardware the SMI RED250n eye tracker was used in combination with an SR research head support and an external monitor with a screen width of 476 mm and a height of 269 mm, with a resolution of 1920x1080. A QWERTY keyboard was used as input.

4.3 Materials

In this section the materials used during both part 1, see subsection 4.3.1, and part 2, see subsection 4.3.2, of the experiment will be explained. All materials will be made in danish as participants will all be native danish speakers.

4.3.1 Comprehension task

As mentioned earlier this will be a repeat of a sorting experiment comparing comprehension times for words and their corresponding emoji. The premise of the experiment is for the participants to sort the basics of words and emoji into their superordinate categories [Rosch et al., 1976] making it possible to get a baseline for the comprehension of both words and emoji. This experiment is originally from an earlier study made by members of this project group in 2017, see sorting task [Jensen et al., 2017]. For this study a similar setup was used but on a different population.

For the word-emoji pairs used in this test, a set consisting of almost the same emoji as the ones used in the earlier experiment will be used [Jensen et al., 2017]. This set consisted of 8 emoji depicting fruits, 8 emoji depicting animals and 5 distractor emoji depicting vegetables, as seen in Figure 4.1. This set were split into two tasks one which had fruits as a target group and one which had animals. For these tasks all 8 emoji from the target group, together with their corresponding words, would be used as targets and distractors would be a mixture of 4 word-emoji pairs from each of the two other categories.

A few changes have however been made to this setup by replacing a word-emoji pair and adding two additional pairs. First off the strawberry have been replaced with a kiwi, as earlier trials have shown that some participants found it weird to categorise strawberry as a fruit as it isn't one. By exchanging the pairs this should no longer be a concern as kiwi should be easily categorised as a fruit. Secondly the two additions were necessary as the emoji in this experiment will be used as a familiarisation for the reading experiment later. The original experiment had 16 pairs of words and emoji, but for the reading task in this study there will be 18 sentences with familiarised emoji the participants will be asked to read, see subsection 4.3.2. As two more emoji were needed an emoji were added from each target group, one for animals and one for fruits. The animal that was added were the pig and the fruit were the cherry. The set of emoji used can be seen on Figure 4.1.

Figure 4.1: All of the emoji used in the comprehension task. First line is the Fruit target condition, second line the Animal target condition and the third line is vegestable distractos

4.3.2 Sentence position

When conducting a reading study mostly one of two options are chosen. Either it is chosen to use naturally occurring texts such as newspaper articles as stimuli [Kliegl et al., 2006], [Kennedy and Pynte, 2005] or it is chosen to construct texts specifically tailored for the purpose of the study [McDonald and Shillcock, 2003], [Schilling et al., 1998], it all depends on what suits the study best. One of the key arguments for using naturally occurring texts is that the stimuli will be more ecological and as such the results will also be closer to describing reading in everyday life. On the other hand this method isn't very useful for studying very particular singular partitions of the reading process, as many of these such as preview benefit and lexical breakdowns are dependent on very specific conditions. In these instances the constructed texts are chosen as they can be manipulated to present the exact conditions that are being tested for. As this study revolves around word-emoji substitution a phenomenon almost exclusively found in instant messaging and as such only in shorter more informal sentences, it has been chosen to construct the sentences as stimuli would be very difficult to obtain otherwise.

Furthermore as the purpose of this test is to look for artefacts from sentence position, such as preview benefit and lexical breakdown. The sentences will be constructed to contain a target emoji as either (a) the first word of the sentence, (b) in the middle of the sentence or (c) as the last word of the sentence. Examples of the 3 different types of sentences can be seen on Figure 4.2.

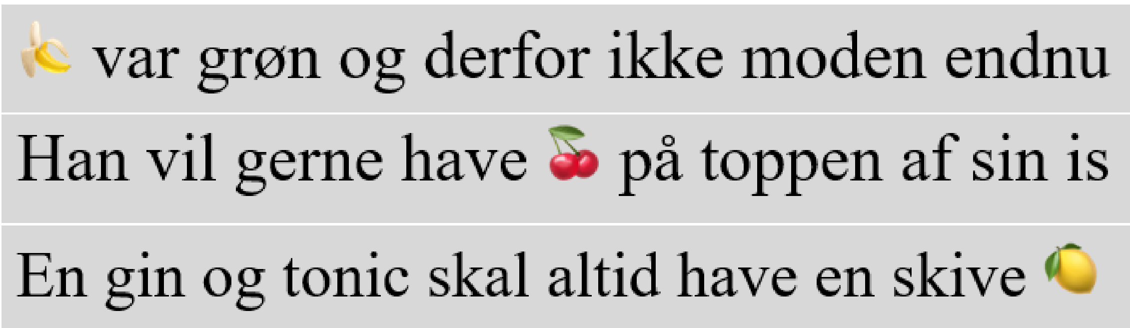

Figure 4.2: Examples of the 3 different stimuli types. In order (a), (b) and (c)

Based on these parameters 27 sentences were constructed, 9 for each target location. Sentences ranged from 5 to 11 words with a **mean = 7.8** and a **SD = 1.6**. Furthermore it was decided to use the emoji from the comprehension task as some of the targeted emoji in the sentences. In 6 of the sentences from each target location these emoji were used. The other 3 sentences were made with emoji that the participants had not been familiarised with, see the familiarised emoji on Figure 4.1 and the non-familiarised emoji on Figure 4.3 below.

Figure 4.3: The 9 non-familiarised emoji added to the subset from the comprehension task

The reason for using the emoji from the comprehension task is to make sure that all participants had been presented with the specific emoji at least once. The non-familiarised emoji were included to see if the familiarisation to the other emoji have an effect on the strategies used when reading them. All of these sentences will be accompanied by an easily answerable question to engage the participants in a task and give them a sentiment for reading the sentences.

4.4 Procedure

All participants will be greeted by the instructor of the experiment and will be asked to sign a declaration of consent, see Appendix B, before participating in the actual experiment. After this the participant will be sat in front of a screen with an eye tracker attached underneath and a keyboard connected. The instructor will be sat behind a screen to the left of the participant, controlling the setup from a laptop, see Figure 4.4 below.

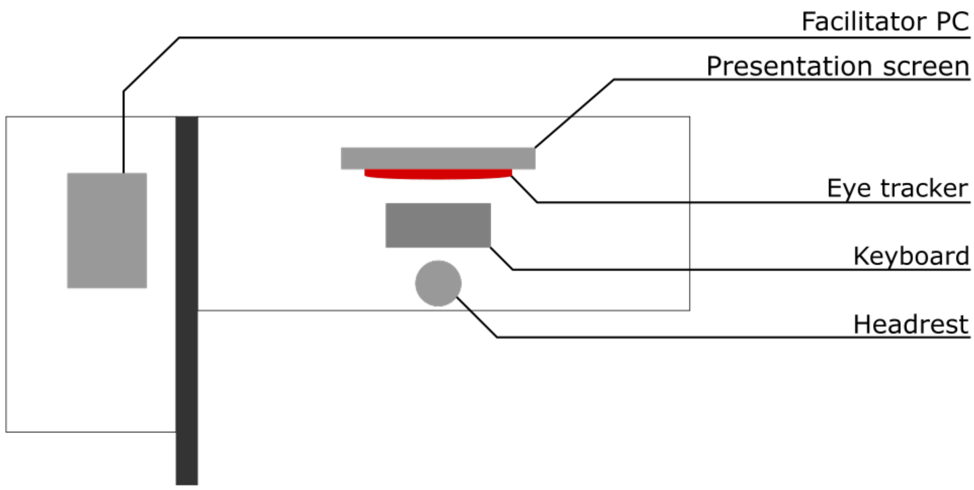

Figure 4.4: Graphical representation of the test setup.

For the the comprehension task the procedure will be as follows. Firstly the participants will be shown an introduction screen with a description of the task they are going to be doing and instructions on how they are supposed to do it. The task they will be asked to carry out is to assign the different emoji and words to the targeted superordinate category. Next a familiarisation round is initialised to give the participants the possibility to get to know the controls. Afterwards the real task begins, the graphical setup for this can be seen on Figure 4.5. When the participants have categorised the first 18 words and 18 emoji the targeted group will be switched and they will be presented with another 18 emoji and 18 words. After this the participants will be allowed a small break before starting on the reading task.

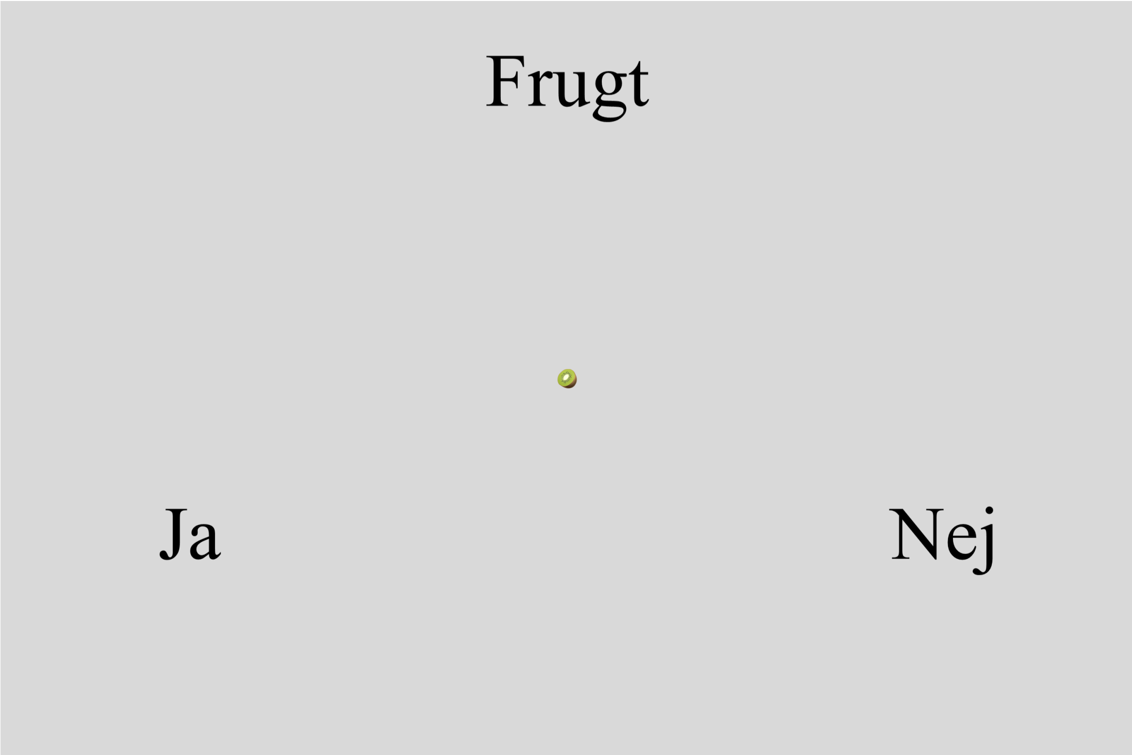

Figure 4.5: Graphical setup for the comprehension task.

This task will start with a calibration of the eye tracker in which deviation of 0.8° or above will result in re-calibration, or rejection of the participant if re-calibrations does not succeed. After this the task starts out like the comprehension task with an introduction screen giving information about the structure of the experiment and the participants task. In this part the participants are asked to read a short sentence and when they're done press the spacebar to continue, the sentences can be found in Appendix A and C. On the next page they will be presented with a question about the sentence they have just read accompanied by 2 possible answers to that question. They answer by using the right and left arrow keys on the keyboard, see Figure 4.6. Before each sentence a fixation cross will appear on the screen in the position of the first character of the sentence. This is done to try to eliminate chasing saccades in the beginning of the eye tracking and by doing so increasing the accuracy of the tracking on the first characters. As in the comprehension task a familiarisation round will be initialised before the actual task to give the participants the possibility to get to know the controls beforehand.

Figure 4.6: Graphical setup for the question in the reading task.

5 | Results

In this chapter the data from the two experiment setups, see chapter 4, will be presented and a statistical analysis will be made.

5.1 Comprehension task

First off the error rate for the comprehension task is calculated to assure the validity of the data. The overall error rate were **4.3%** and individually for each target group the rates were, **6.9%** for fruit target and **1.7%** for animal targets.

When looking at the comprehension times for both the emoji and the words in this task a Shapiro-Wilk normality test was first made to check for normality in both data sets. As all the times recorded will be akin to reaction times the data was positively skewed, see Figure 5.1.

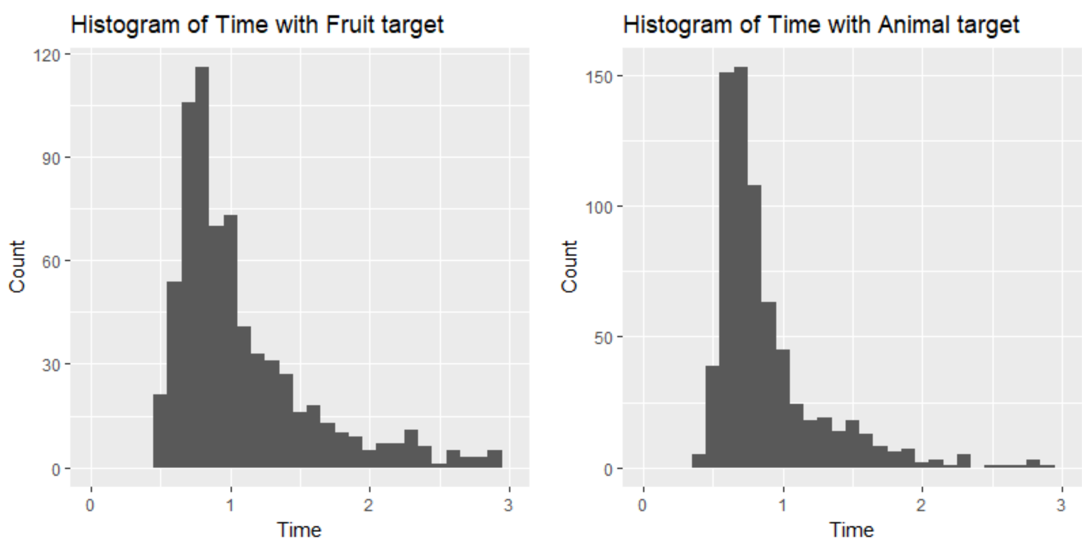

Figure 5.1: Histograms of data distribution for fruit targets(left) and animal targets(right), time is plotted on the x-axis and the frequency is plotted on the y-axis.

The test for fruit targets yielded a $\mathbf{W = 0.90733}$ and a **p-value** $<$ **2.2e-16** and the animal targets yielded a $\mathbf{W = 0.88415}$ and likewise a **p-value** $<$ **2.2e-16**. As both p-values showed a significant difference it was seen that the data is not normally distributed.

As a consequence of this the median was used to report the average of the different groups, together with a median absolute deviation(MAD) and confidence interval(CI) as a measure of the spread of the data. For the fruit target emoji the **median = 918 ms**, **MAD = 369** with an **95% CI [850, 968]** and for the words the **median = 950**, **MAD = 345** with an **95% CI [900, 984]**, for the animal target emoji the **median = 733**, **MAD = 221** with an **95% CI [700, 767]** and for the words **median = 767**, **MAD = 199** with an **95% CI [734, 816]**. For comparison of the word-emoji sets a Wilcoxon signed rank test was used and yielded a **p-value = 0.715** for the group with fruit targets and a **p-value = 0.005** for the group with animal targets. As the animal targets showed a significant difference an effect size was calculated using Cliff's Delta [Macbeth et al., 2011], because of nonparametric data, yielded **d = -0.132**. Cliff's delta gives a number between +1 and -1, where at 0, distributions are overlapping completely yielding no effect and at ±1 the distributions don't overlap at all yielding a big effect [Macbeth et al., 2011]. Given that Cliff's delta in this circumstance yielded a value close to 0 the effect size is negligible. To visualise the data from all individual stimuli box plots were made, see Figure 5.2 and Figure 5.3.

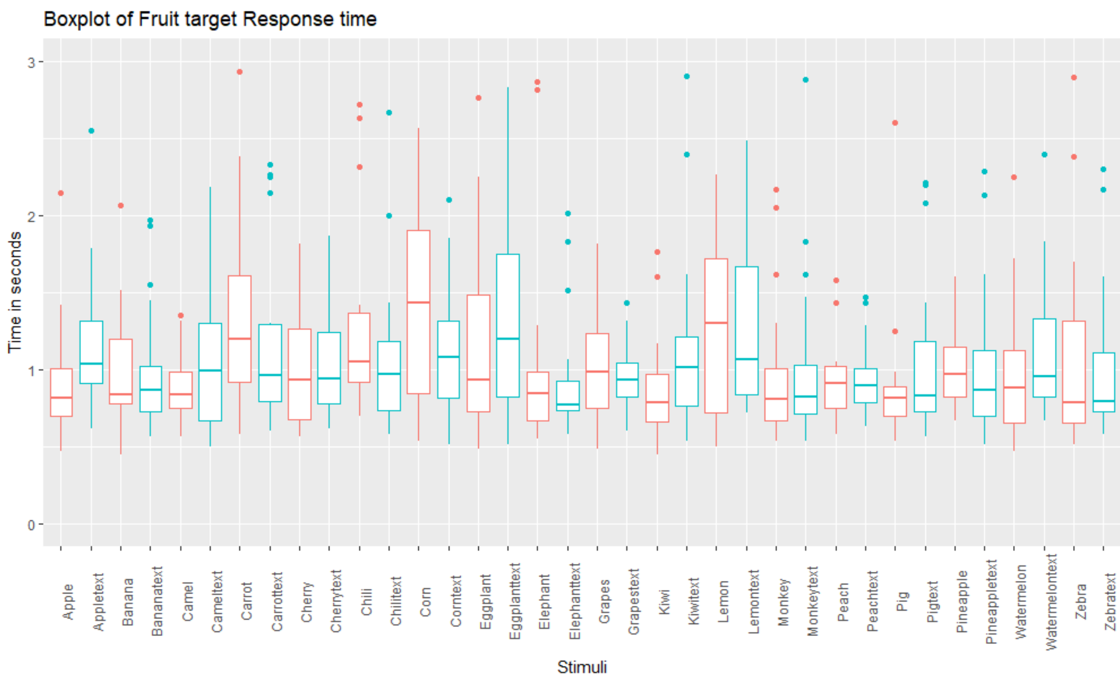

Figure 5.2: Box plot of fruit targets, with stimuli on the x-axis and comprehension time in seconds on the y-axis.

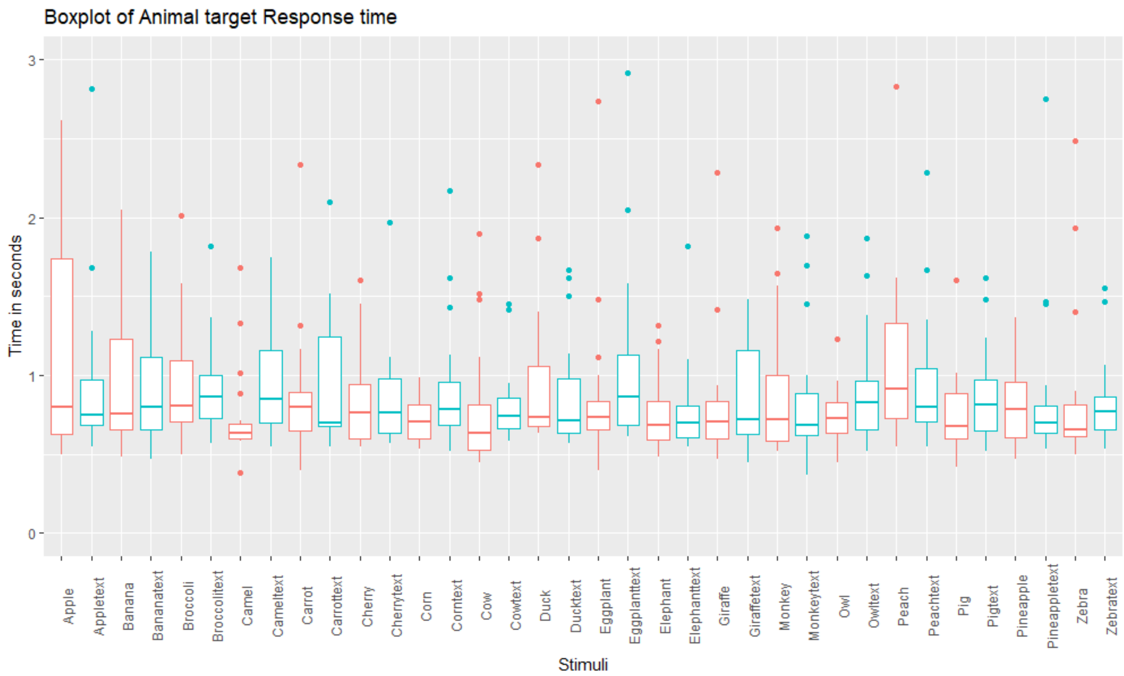

Figure 5.3: Box plot of animal targets, with stimuli on the x-axis and comprehension time in seconds on the y-axis.

5.2 Reading task

As with the comprehension task the error rate was calculated for all the stimuli of the reading task, the overall error rate were **0.4%** with 1 error on a front target(F) and one error on a middle target(M). Both errors were on sentences with non familiarised emoji.

As all eye tracking data were gathered as a screen recording and all stimuli were shown in a randomised sequence the first part of the analysis were mapping of the eye tracking data for all participants to the stimuli it belonged to. After this had been done heat maps were made to show what parts of the stimuli sentences that drew most attention on average across all participants. When looking at the three different emoji positions in the different stimuli some tendencies emerged. First off, when looking at the sentences with end targets(E), it was seen that 5 out of 9 sentences displayed an increased amount of attention on or in the near vicinity of the emoji. This tendency was also seen on the sentences with front targets(F), in 6 out of 9 cases, and for the sentences with middle target(M), in 5 out of 9 cases, examples of this can be seen on Figure 5.4, 5.5 and 5.6 below.

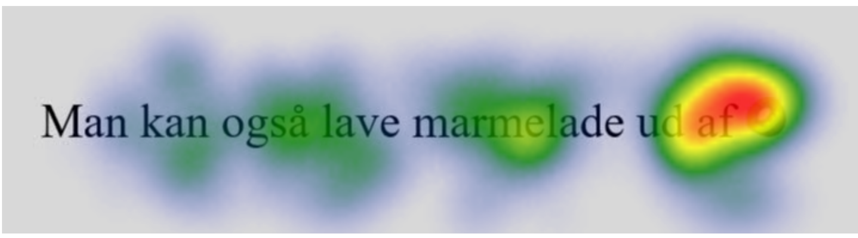

Figure 5.4: Heat map of stimuli sentence E3.

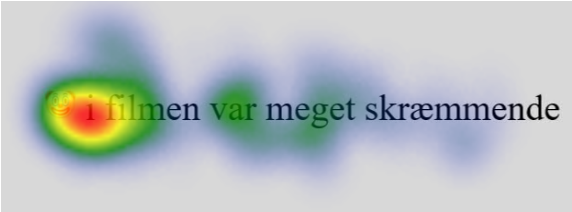

Figure 5.5: Heat map of stimuli sentence F1x.

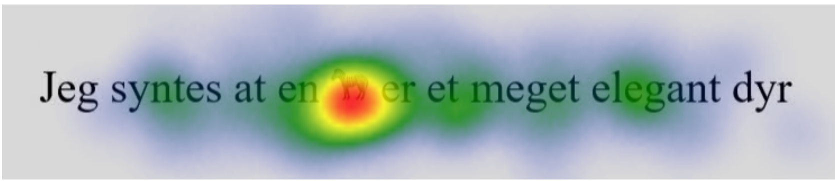

Figure 5.6: Heat map of stimuli sentence M2.

When looking at the remaining sentences there are either more than one hot spot or the the attention is more evenly distributed throughout the sentence, see Figure 5.7 and 5.8.

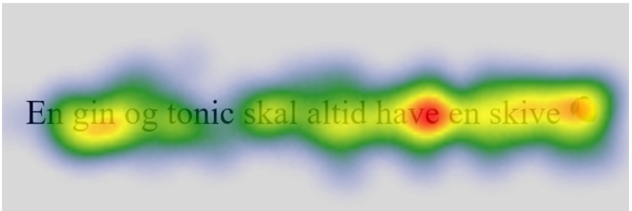

Figure 5.7: Heat map of stimuli sentence E4.

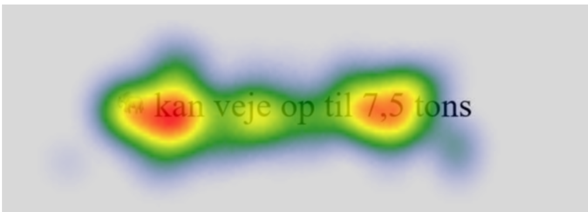

Figure 5.8: Heat map of stimuli sentence F2.

The next step in the analysis were the application of *Areas of Interest* also called AOIs to get more specific data. This is done by the AOIs functioning as a filter that only takes in to consideration the data from within the specified area and thus giving results such as fixation time based solely on the stimuli contained in the AOI. AOIs were created around each word and emoji in the sentences to get mean fixation times, when the word or emoji was fixated compared to the rest of the sentence, *Revisitors* to the word or emoji and how many of the participants actually fixated on it. As all stimuli were presented in a single line it was chosen to extend all AOIs below and above the words they belonged to, to counteract potential inaccuracies on the Y-axis. Furthermore the AOIs were used to create sequence charts for both raw data and fixations, to graphically illustrate the sequence in which emoji and words were looked at, an example can be seen on Figure 5.9 and the rest of the charts in Appendix C.

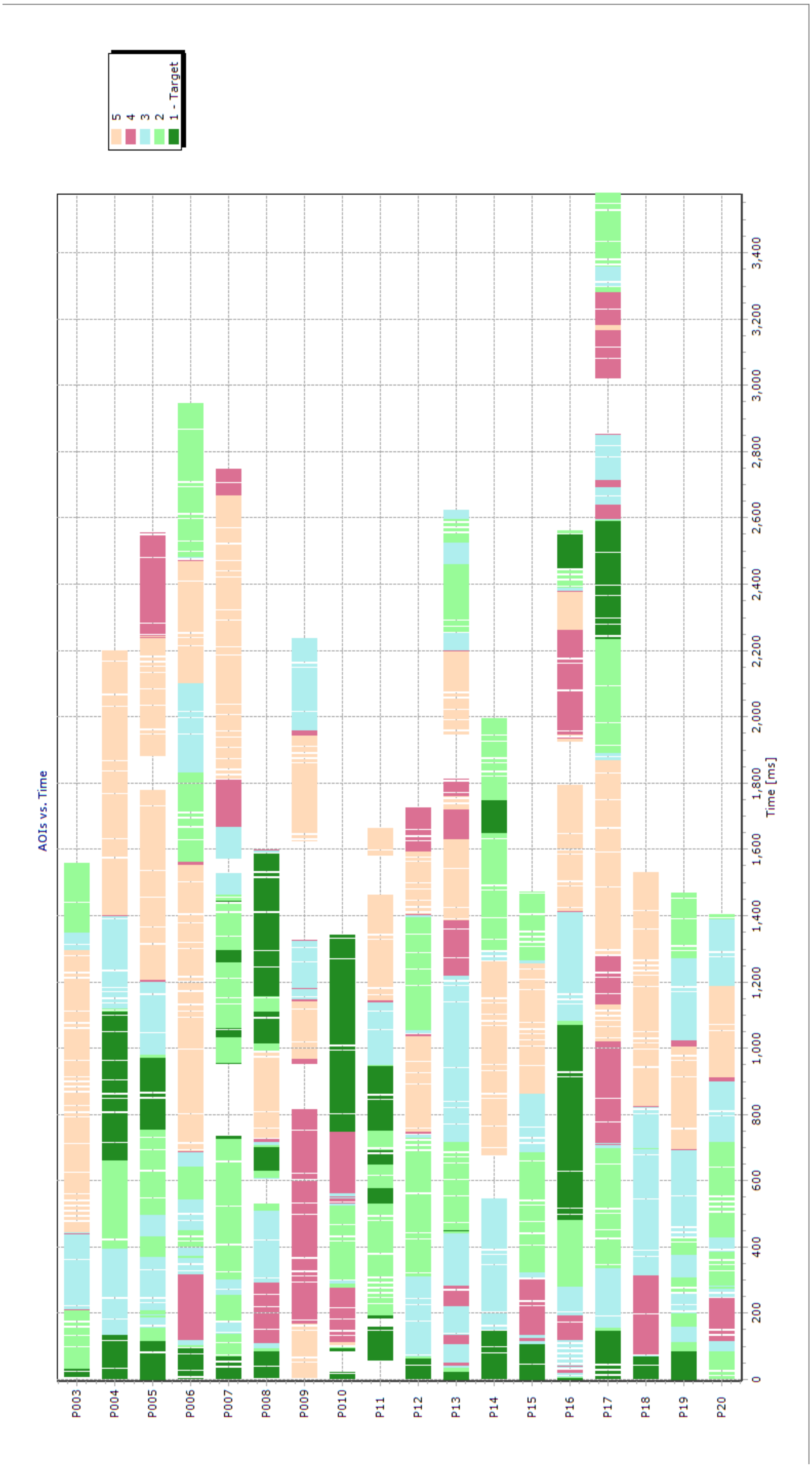

Figure 5.9: Graph showing raw data in their respective AOIs. X-axis = reading time in ms and Y-axis = participants. In the legend all AOIs are named with a number corresponding to their respective placement in the sentence with the emoji AOI marked with the word *Target*.

For the fixations the detection parameters used were a maximum dispersion of data points by 2° at a distance of 60 cm to the eye tracker and a minimum duration of 80 ms, as these are the standard settings used by BeGaze.

When looking at the fixation times the BeGaze metric *First Fixation* was used as it was observed and can be seen in the sequence charts in Appendix C that many participants did secondary scans of the sentences. The *First Fixation* metric gives a mean of all participants first fixation on the given AOI and as such this gave the most accurate measure of how long it took to process the words and emoji. Furthermore all 0-value entries have been removed in the calculation of mean first fixations as these values otherwise would positively skew the data. This was chosen to to get a more accurate median of the fixation time on the emoji by excluding the participants that didn't look at it at all. The metrics for how many of the participants actually looked at the emoji can be found in Table 5.1 under *Hit Ratio*.

Stimuli	Emoji	Median First Fixation Time (ms)	Hit Ratio (X/18)	Revisitors (X/Hit Ratio)
E1	🍑	518	6	0
E1x	🍺	314	9	1
E2	🐪	504	10	0
E2x	☕	414	8	1
E3	🥝	370	9	0
E3x	🦇	445	10	0
E4	🍋	348	6	0
E5	🐒	396	8	0
E6	🦆	372	8	0
F1	🍌	185	3	1
F1x	🤡	351	5	2
F2	🐘	207	6	1
F2x	🦋	740	6	2
F3	🍍	208	6	1
F3x	🐈	631	6	1
F4	🐄	472	9	1
F5	🦒	165	5	1
F6	🥕	380	2	1
M1	🍒	211	12	4
M1x	🍕	190	11	3
M2	🦓	222	10	7
M2x	🥛	209	13	2
M3	🐖	278	13	0
M3x	⚽	196	9	4
M4	🦉	343	12	8
M5	🌶️	220	14	3
M6	🥦	267	12	5

Table 5.1: Key metrics for emoji targets in all sentences, all data is made for fixations. Hit Ratio is the number of participants out of all participants that have fixated within the AOI, *Revisitors* is the number of participants out of the Hit Ratio on the specific AOI that have fixated within that AOI more than once. X's indicate emoji that participants have not been familiarised to.

To visualize the data from Table 5.1 and add information about quartiles and outliers a box plot was made, see Figure 5.10.

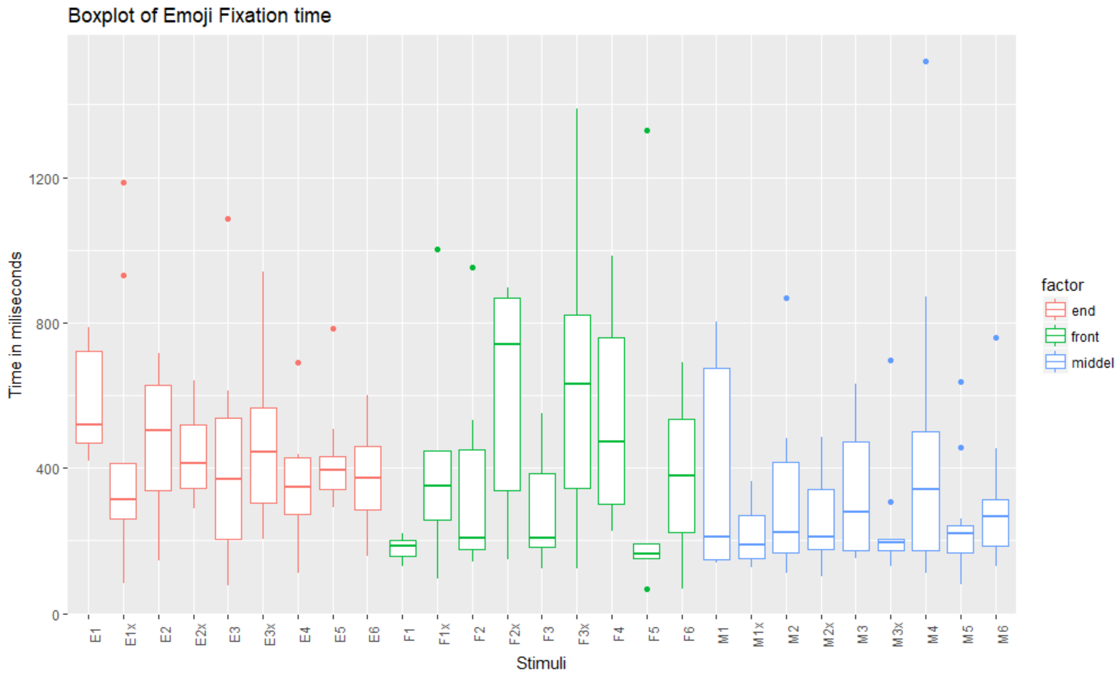

Figure 5.10: Box plot of emoji fixation times across all participants. This plot is based on the Hit Ratio of Table 5.1, and as such some of the stimuli have very few data points.

Next medians were calculated for the fixation times across the 3 different types of stimuli to see if there were some notable differences. All of the following groups tested significant with the Shapiro-Wilks test and are therefore represented with medians. First off, for the end targets(E) the median was calculated to be **median = 411 ms**, **MAD = 183** with an **95% CI [368, 457]**. For the front targets(F) the median was calculated to be **median = 306 ms**, **MAD = 244** with an **95% CI [208, 481]** and lastly for middle targets(M) the median was calculated with the result, **median = 222 ms**, **MAD = 110** and an **95% CI [195, 268]**. As the different conditions were missing data points from some of the participants, they were of unequal sample sizes and therefore a Wilcoxon signed rank test was not conducted to make a statistical comparison.

In continuation of this, medians were calculated for the fixation times of stimuli that the participants had been familiarised with, and non-familiarised stimuli. All of these can be found in Table 5.2 below together with MAD values and CI for all the conditions.

Stimuli	Median	MAD	CI low	CI high
E fam	421	173	368	510
E non-fam	381	159	288	494
F fam	220	134	185	472
F non-fam	448	483	504	638
M fam	245	126	195	282
M non-fam	199	93	177	308

Table 5.2: Summery for fixations in familiarised and non-familiarised sentences

As a further analysis of the data means were also calculated for the Hit Ratio and the number of *Revisitors* for the different emoji positions and for familiarised and non-familiarised groups in the data, see Table 5.3.

Stimuli	Mean Hit Ratio	Mean Revisitors
E	8.2	0.2
E fam	7.8	0
E non-fam	9	0.7
F	5.3	1.2
F fam	5.2	1
F non-fam	5.7	1.3
M	11.8	4
M fam	12.2	4.5
M non-fam	11	3

Table 5.3: Mean Hit Ratios and *Revisitors* across all stimuli, as well as for familiarised and non-familiarised

As it can be seen in Table 5.3 above the hit ratio for the front and end positions were quite low with less than half of the participants actually looking at the emoji in both cases. This combined with the theory stating that front and end words not being fixated in many circumstances, prompted a calculation of mean hit ratios for these words in the stimuli sentences. To explore the counterpart to the *Hit Ratio* of emoji a table for *Hit Ratio* of the first and last word in each sentence can be seen below on Table 5.4.

Stimuli	First word	Last word	Hit Ratio (X/18) First word	Hit Ratio (X/18) Last word
E1	Af	🍑	8	6
E1x	Vil	🍺	1	9
E2	Jeg	🐪	3	10
E2x	Jeg	☕	1	8
E3	Man	🥝	8	9
E3x	En	🦇	2	10
E4	En	🍋	5	6
E5	Vi	🐒	2	8
E6	Til	🦆	0	8
F1	🍌	endnu	3	12
F1x	🤡	skræmmende	4	15
F2	🐘	tons	6	6
F2x	🦋	luften	6	11
F3	🍍	frugt	6	7
F3x	🐈	solen	6	10
F4	🐄	drøv	9	7
F5	🦒	pattedyr	5	13
F6	🥕	A-vitaminer	2	16
M1	Han	is	4	5
M1x	Han	toppen	7	12
M2	Jeg	dyr	1	3
M2x	Jeg	morgenmad	1	13
M3	Jeg	dag	2	9
M3x	Bo	eftermiddag	3	15
M4	Syv	Danmark	10	9
M5	Der	stærk	3	7
M6	Kina	verden	11	11

Table 5.4: Comparison of the Hit ratio for first and last words and emoji across all stimuli. Hit Ratio is the number of participants out of all participants that have fixated within the AOI.

To be able to compare first and last words with front and end emoji, means were calculated across the stimuli that did not have an emoji in the position in question. As such first words were calculated on middle target(M) and end target(E) and yielded a **mean** = **4** and last words were calculated on front target(F) and middle target(M) and yielded a **mean** = **10**.

6 Discussion

6.1 Comprehension time of words compared to emoji

When the data from the comprehension task is explored it can be seen that there is a clear difference between the median comprehension time of both the words and emoji of the Fruit targets, **word median = 950 ms** and **emoji median = 918 ms** respectively, compared to the median comprehension time of the Animal targets, **word median = 767 ms** and **emoji median = 733 ms**. This tendency of a 200 ms difference can most likely be attributed to the fact that the Fruit targets shared a visual similarity to the Vegetable distractors to a greater extent than the Animal targets did, see Figure 4.1.

This visual similarity could also explain the fact that there were found no significant difference, with a **p-value = 0.715**, between words and emoji when looking at the Fruit targets. The opposite was found when looking at the Animal targets where the **p-value = 0.005** indicates a clear significant difference between the words and emoji. Furthermore the effect size, **d = -0.132**, show that there is an effect of this difference however, that effect is only negligible which means that there is a great deal of overlap between the two samples but they are still far enough from each other that there is a significant difference. With this in mind it is argued that the comprehension time for the Animal targets are a closer representation of a real comprehension time than the comprehension times of the Fruit targets.

Lastly it can be seen on the MAD values that the internal deviation of the data sets are pretty much the same for emoji and words in the two conditions. However, the data points in the Fruit condition are largely more spread out, with a **MAD = 369** for emoji and **MAD = 345** for words than the points in the Animal condition, with a **MAD = 221** for emoji and **MAD = 199** for words. This again could be attributed to the visual similarity between distractors and targets for the Fruit targets as mentioned above.

When looking at these results it is important to keep in mind that they are not true representations of comprehension times, as there is a lag between the actual comprehension, the decision on what to answer, the fingers pressing the button on the keyboard and the computer registering the key press. With this in mind all words and emoji comprehension times will be longer than the actual comprehension time. However all stimuli should be affected equally by this delay.

6.1.1 Comparison with earlier comprehension results

Looking back at the results from Jensen et al. [2017] and comparing them with the results from this study it can be seen that they are almost identical in regards to the deviation between comprehension times for words and emoji in both target groups. Specifically the emoji and word deviation for the Fruit targets between studies are only **32 ms** in this study and **34 ms** in Jensen et al. [2017]. The same goes for the Animal targets where the deviation is exactly **34 ms** in both studies.

An aspect where the two studies deviate slightly is on the actual comprehension time where this study displays longer comprehension times on both emoji and words compared to the earlier study. Here the Fruit targets take **12 ms** longer for emoji and **10 ms** longer for words and the Comprehension time for the Animal targets has increased with **27 ms** for both emoji and words. These increases are however so small that they are negligible and could easily be attributed to an increase in lag time from keyboard to computer. This is due to the earlier setup using a keyboard directly connected to the computer and the setup in this study using a keyboard connected to the computer via a USB on the secondary screen.

Another difference between the two studies is the effect size found within the Animal target condition, where the effect shown in Jensen et al. [2017] is large with $\mathbf{d} = \mathbf{-0.565}$ and the effect found in this study is negligible with $\mathbf{d} = \mathbf{-0.132}$. This difference in effect size, despite there being almost no difference in the medians of the two samples, can probably be contributed to the difference in sample sizes, as this study has a sample size of 20 and Jensen et al. [2017] had a sample size of 30.

When looking at the setups for the two studies another difference can be seen as Jensen et al. [2017] familiarised all emoji the participants would be shown in the sorting task, where as this study did not. This could have had an influence on the results of this study as participants could have stronger associations with some emoji than they had with others and in fact such a tendency can be seen. When looking at the quartiles on for the comprehension time of the individual stimuli in both the Fruit targets and the Animal targets, see Figure 3.1, 3.2, 5.2 and 5.3, it can be seen that there is a larger difference in the dispersion of the data points between individual stimuli in this study than in Jensen et al. [2017], the difference in sample sizes could also contribute to this disparity. However as the median comprehension times of the two studies are so similar it could indicate that familiarisation of the emoji doesn't impact the results greatly and as such familiarity with the specific emoji might not have that big an influence on the comprehension time.

6.2 Fixation tendencies in the reading task

First off, when looking at the heat maps the general tendency is that there is either a single hot spot on or around the emoji or two or more hot spots centred on key information in the sentences, including the emoji as well. This tendency can most likely be explained by the experimental setup where the participants are asked to read the sentence with the intent of answering a follow-up question. This in combination with the participants not being restricted to only one read through of the text makes it likely that on second takes only the key information parts of the sentences are fixated upon resulting in longer accumulated fixations on these specific areas.

When looking at the different conditions in this task it can be seen that for the middle targets(M) the fixations depicted on the heat maps are focused directly on the emoji, where as when looking at the front and end emoji the fixations don't align completely. For the front targets(F) the fixations are situated just after the emoji, on the next word in line, and for the end targets(E) the same tendency can be seen, just with the fixations situated on the previous word instead. This tendency is explained by both Rayner et al. [2012] and Engbert et al. [2005] in two different ways. Rayner et al. [2012] says that attention will shift leftward from the fixation covertly and in doing so taking advantage of the perceptual span in reading that extends 3-4 character spaces to the left of fixation, processing the emoji, see subsection 2.2.1. Engbert et al. [2005] however states with the SWIFT model that multiple words around the fixation will be processed at the same time, this specifically including word N-1 which in this case is the emoji in question, see subsection 2.2.2. Both seems an equally likely explanation for the scenario.

6.2.1 Fixation time across placements

Moving on to the median fixation times for the emoji used in the reading task it can be seen that times differ by about **100 ms** between the three target positions, end targets(E) having the longest fixation time, followed by front targets(F) and lastly middle targets(M) having the lowest. It is known from section 2.1 that the theoretical average fixation time is somewhere between **200-250 ms** and only the fixation time for the middle targets(M) fits this with a **median = 222 ms**. Both front targets(F) and end targets(E) have longer fixation times with medians of respectively **306 ms** and **411 ms**. This could mean that the processing of these emoji is difficult and as such take longer to finish.

The E-Z reader model, subsection 2.2.1, states that a prolonging factor for the completion of the familiarity check(L1) is the word's predictability in the sentence structure and its frequency of use in text media and as an emoji can be said to be highly unpredictable and maybe infrequently used, depending on the emoji, this could be an explanation for the longer fixation times. The SWIFT model, subsection 2.2.2, has a similar feature where saccades are inhibited to allow for more processing time on the current fixation. As such both models could describe the phenomenon in question, however, as the fixation time on the middle emoji is within the boundaries of the theoretical average, this might not be the case.

Lastly, when looking at the fixation times for the familiarised compared to the non-familiarised emoji, no clear tendencies arise. As such the differences in the data could very well be due to individual and unforeseen factors with the specific emoji or in the sentence structure.

6.2.2 Hit ratio and revisitors

A tendency that emerges when looking at Table 5.3 is that when the emoji is situated in the middle of the sentence the mean hit ratio is higher than the hit ratios of both the end targets(E) and the front targets(F). The middle targets(M) have an average ratio of **11.8/18** whereas the end targets(E) have an average of **8.2/18** and the front targets(F) have **5.3/18**. This supports the finding from the heat maps mentioned earlier where fixations are situated just to the right of the front targets(F) and just to the left of end targets(E). As such these findings support the notion that emoji facilitates preview benefit. To support this statement the mean hit ratio of the first and last words can be taken into account, see Table 5.4. Here it can be seen that the means of the emoji and words match up with almost no difference, as the mean for start words is **4** and the mean for front emoji is **5.3** and the mean for last words is **10** and the mean for last emoji is **8.2**. The biggest difference is between the last words and emoji with a difference of nearly two persons, however this difference could be explained by word length as the end words are **6** characters long on average, with a **SD** = **2.6**, ranging from **2** to **11** character. This means that the words most likely wouldn't be fully processed by preview benefit, thus necessitating a fixation.

When looking at the fixation data for F6 in Table 5.1 it can be seen that this is the stimuli with the lowest hit rate, however if the sequence chart for raw data is explored it can be seen that most participants briefly have their gaze on the target, **16/18** participants in fact, see Figure C.90. This tendency can also be seen on the other stimuli with a low hit rate, such as F1,F1x and F5, however not with the same magnitude. An explanation for this tendency could be found in the setup of the task where all trials start with a fixation cross on the first character of the sentence and as such the gaze registration could just be residuals from the eye not having moved yet. This would support the statement of attention being shifted covertly to the left on first fixations in sentences, see subsection 2.2.1, and processing being spread out to word N+1 and N-1 from fixation, see subsection 2.2.2.

A similar tendency to the hit ratio can also be seen when looking at the average number of revisitors where the middle targets(M) receives 3.3 times as many revisitors on average than the sentences with front targets(F). The middle targets(M) receive **4** revisitors on average compared to the **1.2** revisitors for the front targets(F). When the data from the middle targets(M) are compared to the end targets(E) it can be seen that they have a little over 18 times the average revisitors than these sentences that only have an average of **0.2** revisitors. This further substantiates the notion that emoji facilitates preview benefit, by taking into account that even when the participants do second takes of the sentences they only rarely fixate the emoji. This tendency correlates with the fixation times on the different target positions where a lower fixation time equates to a higher number of revisitors. As such the amount of revisitors could also be an expression of the targets needing more processing to achieve integration. However as the amount of revisitors is as low as it is for all positions, it could also be due to individual factors for the different participants.

7 | Conclusion

On the foundation of the data found in chapter 5, the earlier results displayed in chapter 3 and the joint discussion in chapter 6 the following is concluded.

That the application of word-emoji substitution to a text does not result in any measurable difference in reading time of said text. Furthermore it does not contribute to either increases in fixation time or backtracking. It can also be seen that sentences beginning with emoji does not seem to express tendencies towards lexical breakdowns with either increased fixation times or revisits by participants. On the contrary, the addition of emoji as both the first and last word of a sentence, seem to show results consistent with preview benefit. Lastly, all results showed a tendency towards comprehension time for emoji being significantly faster compared to words.

In chapter 1 the question of whether current models of reading can contain the reading of emoji in sentences is asked and on the basis of above conclusions it is found that current models of reading does encompass word-emoji substitution without modifications. However for this and the two previous studies the focus has been on comprehension and fixation times and not on saccadic structure which would also have to be studied to draw any final conclusions.

Bibliography

Balota and Chumbley, 1984. David A. Balota and James I. Chumbley. *Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage.* 10(3), 340–357, 1984.

Bell, July 2016. Killian Bell. *How to swap words ofor Emoji in iOS 10.* Blog, 2016. URL `http://www.cultofmac.com/439952/how-to-swap-words-for-emoji-in-ios-10/`.

EDUCBA, April 2016. EDUCBA. *10 Effective Ways to Use Emojis Everywhere in Marketing.* Blog post, 2016. URL `https://www.educba.com/emojis-everywhere/`.

Engbert et al., 2005. Ralf Engbert, Antje Nuthmann, Eike M. Richter and Reinhold Kliegl. *SWIFT: A Dynamical Model of Saccade Generation During Reading.* Psychological review, 112(4), 777–813, 2005. doi: 10.1037/0033-295X.112.4.777.

Engineering, April 2015. Instagram Engineering. *Emojineering Part 1: Machine Learning for Emoji Trends.* Blog post, 2015. URL `https://engineering.instagram.com/emojineering-part-1-machine-learning-for-emoji-trendsmachine-learning-for-emoji-trend`

Fahlman, 2007. Scott Elliott Fahlman. *Smiley Lore :-),* 2007. URL `http://www.cs.cmu.edu/~sef/sefSmiley.htm`.

Farnsworth, April 2019. Bryn Farnsworth. *What is Eye Tracking and How Does it Work?* Website, 2019. URL `https://imotions.com/blog/eye-tracking-work/`.

Guestrin and Eizenman, June 2006. Elias Daniel Guestrin and Moshe Eizenman. *General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections.* IEEE, 53(6), 2006.

Jensen et al., 2017. Mikkel B. R. Jensen, Karolis M. Nielsen, Kasper H. Ravnkilde and Alexander F. Povlsen. *The Effect of Word-Emoji Substitution on Readability of Text.* 2017. URL `https://projekter.aau.dk/projekter/files/268053705/The_Effect_of_Word_Emoji.pdf`.

Kennedy and Pynte, 2005. Alan Kennedy and Joël Pynte. *Parafoveal-on-foveal effects in normal reading.* Vision Research, 45, 153–168, 2005. doi: 10.1016/j.visres.2004.07.037. URL `https://www.sciencedirect.com/science/article/pii/S0042698904003979`.

Kliegl et al., 2006. Reinhold Kliegl, Antje Nuthmann and Ralf Engbert. *Tracking the Mind During Reading: The Influence of Past, Present, and Future Words on Fixation*

Durations. Experimental Psychology: General, 135(1), 12–35, 2006. doi: 10.1037/0096-3445.135.1.12. URL `https://psycnet-apa-org.zorac.aub.aau.dk/fulltext/2006-01956-002.html`.

Macbeth et al., 2011. Guillermo Macbeth, Eugenia Razummiejczyk and Rubén Daniel Ledesma. *Cliff's Delta Calculator: A non-parametric effect size program for two groups of observations*. Universitas Psychologica, 10, 545–555, 2011.

McDonald and Shillcock, 2003. Scott A. McDonald and Richard C. Shillcock. *Eye Movements Reveal the On-Line Computation of Lexical Probabilities During Reading*. Psychological Science, 14(6), 648–652, 2003. doi: 10.1046/j.0956-7976.2003.psci_1480.x. URL `https://doi.org/10.1046/j.0956-7976.2003.psci_1480.x`. PMID: 14629701.

Peirce and MacAskill, 2018. Jonathan Peirce and Michael MacAskill. *Building Experiments in PsychoPy*. Sage Publications, first edition edition, 2018. ISBN 978-1-4739-9138-5.

Ravnkilde and Povlsen, 2017. Kasper H. Ravnkilde and Alexander F. Povlsen. *Readability of Emoji*. 2017. URL `https://projekter.aau.dk/projekter/da/studentthesis/readability-of-emoji(be50ad6c-b709-4ab4-838e-cb943b8e77e2).html`.

Rayner, 2009a. Keith Rayner. *Eye movements and attention in reading, scene perception, and visual search*. The quarterly journal of experimental psychology, 8, 1457–1506, 2009. doi: 10.1080/17470210902816461.

Rayner, 2009b. Keith Rayner. *Eye Movements in Reading: Models and Data*. Eye Movements Research, 2009.

Rayner et al., 2012. Keith Rayner, Alexander Pollatsek, Jane Ashby and Charles Clifton Jr. *Psychology of Reading*. Psychology Press, second edition edition, 2012. ISBN 978-1-84872-943-8.

Rayner et al., 2016. Keith Rayner, Elizabeth R. Schotter, Michael E. J. Masson, Mary C. Potter and Rebecca Treiman. *So Much to Read, So Little Time*. Psychological Science in the Public Interest, 17(1), 4–34, 2016. doi: 10.1177/1529100615623267.

Erik D. Reichle. Serial-attention model of reading, chapter 42, pages 773–776. Oxford university press, 2013.

Reichle et al., 2005. Erik D. Reichle, Alexander Pollatsek and Keith Rayner. *E-Z Reader: a cognitive-control, serial-attention model of eye-movement behavior during reading*. Cognitive Systems Research, 7, 4–22, 2005.

Eleanor Rosch, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson and Penny Boyes-Bream. Basic Objects in Natural Categories, pages 382–439. 1976.

Salvucci, 2001. Dario D. Salvucci. *An integrated model of eye movements and visual encoding*. Cognitive systems research, 1, 201–220, 2001.

Schilling et al., Nov 1998. Hildur E. H. Schilling, Keith Rayner and James I. Chumbley. *Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences*. Memory & Cognition, 26(6), 1270–1281, 1998. ISSN 1532-5946. doi: 10.3758/BF03201199. URL `https://doi.org/10.3758/BF03201199`.

TobiiPro, 2019. TobiiPro. *How do Tobii Eye Trackers work?* Website, 2019. URL `https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/`.

Unicode.org, May 2019. Unicode.org. *Emoji Versions, v5.0*, 2019. URL `https://unicode.org/emoji/charts/emoji-versions.html`.

Charles A. Weaver and A. E. Holmes. Psychology of Reading, pages 209–217. 2012. ISBN 978-0-08-096180-4.

A | Stimuli Sentences

Emoji position is anoted with *...*, Y = Correct answer, N = Wrong answer.

Emoji start
Bananen var grøn og derfor ikke moden endnu
Man kunne se at bananen ikke var moden fordi
Y:Den var grøn
N:Den var brun
Elefanten kan veje op til 7,5 tons
Hvor meget kan en elefant veje op til
Y:7,5 tons
N:6,5 tons
Ananas er en sød og tropisk frugt
Hvordan smager en ananas
Y:Sød
N:Salt
Koen ligger på marken og tygger drøv
Hvad foretog koen sig på marken
Y:Tygger drøv
N:Tager en lur
Giraffen er verdens højeste pattedyr
Hvad er det højeste pattedyr i verden
Y:Giraf
N:Elefant
Gulerod er rig på C-vitaminer
Hvilken type vitamin er gulerødder rige på
Y:C-vitamin
N:E-vitamin

Klovnen i filmen var meget skræmmende
Hvordan var klovnen i filmen portrætteret
Y:Meget skræmmende
N:Meget sjov
Sommerfuglen fløj let igennem luften
Hvordan fløj sommerfuglen igennem luften
Y:Let
N:Besværet
Katten tager sig en lur i solen
Hvor tog katten sig en lur

Y:I solen
N:I skyggen

Emoji middle
Han vil gerne have *kirsebær* på toppen af sin is
Hvad vil han gerne have på toppen af sin is
Y:Kirsebær
N:Jordbær
Jeg syntes at en *zebra* er et meget elegant dyr
Hvordan er zebraen beskrevet
Y:Som værende elegant
N:Som værende klodset
Jeg vil gerne have *gris* til aftensmad idag
Hvad ønskes der til aftensmad
Y:Gris
N:Ko
7 arter af *ugler* yngler i Danmark
Hvor mange arter af ugler yngler i Danmark
Y:7
N:9
Der skal godt med *chilli* på maden så den bliver stærk
Hvad skal der på maden
Y:Chilli
N:Peber
Kina er en af de største *brocoli* producenter i verden
Hvilket land er en af de største producenter af brocoli i verden
Y:Kina
N:Japan

Han kunne bedst lide *pizza* med ananas på toppen
Hvad for noget topping kunne han bedst lide på sin pizza
Y:Ananas
N:Hvidløg
Jeg drikker altid *mælk* til min morgenmad
Hvornår drikkes der mælk
Y:Til morgenmaden
N:Til aftenkaffen
Bo og John spiller *fodbold* hver onsdag eftermiddag
Hvem spiller fodbold hver onsdag
Y:Bo og John
N:Bo og Lars

Emoji end
Af alle frugter kunne hun bedst lide *fersken*
Hvilken frugt kunne hun bedst lide
Y:Fersken
N:Appelsin
Jeg har altid gerne villet have en *Kamel*

Hvilket dyr ønskes der
Y:Kamel
N:Pingvin
Man kan også lave marmelade ud af *kiwi*
Hvad kan man også lave ud af kiwi
Y:Marmelade
N:En sko
En gin og tonic skal altid have en skive *citron*
Hvad skal en gin og tonic altid have
Y:En skive citron
N:En oliven
Vi mennesker er nært beslægtede med *aber*
Hvad er vi mennesker nært beslægtede med
Y:Aber
N:Sæler
Til juleaften spiser min familie altid *and*
Hvad spiser familien altid til jul
Y:And
N:Boller i karry

Vil du med ud og drikke en *øl*
Hvad bliver du inviteret med ud og drikke
Y:En øl
N:En kop te
Jeg kunne virkelig godt drikke en kop *kaffe*
Hvad kunne personen virkelig godt drikke
Y:En kop kaffe
N:En Cola
En meget berømt tegneseriefigur er baseret på en *flagermus*
Hvad er den meget berømte tegneseriefigur baseret på
Y:En flagermus
N:En flodhest

Familiarization round
Du skal ikke række *tunge* af mig
Hvad skal du ikke gøre
Y:Række tunge
N:Stirre
Jeg tager lige et *bad* og så kommer jeg
Hvad gør personen
Y:Tager et bad
N:Vander blomster
Slanger er ikke slimede at røre ved
Hvordan føles slanger at røre ved
Y:Ikke slimede
N:Slimede
Det tætteste på en dinosaur i dag er en *krokodille*
Hvad er det tætteste man kommer på en dinosaur i dag
Y:En krokodille

N:En fugl
Det anbefales at man spiser *fisk* to gange om ugen
Hvor mange gange om ugen skal man helst spise fisk
Y:2
N:3

B Declaration of consent

I forbindelse med Kasper Horslev Ravnkilde og Alexander Flyvholm Povlsen og deres projekt på 10. semester omhandlende "The robustness of current reading models on incorporation of word-emoji-substitution" giver jeg hermed min tilladelse, til at indsamlet data må bruges som en del af den studerendes uddannelse ud fra følgende aftaler og specificeringer:

Jeg giver min tiladelse til,	JA	NEJ
- At indsamlet data må bruges i forbindelse med omtalte projekt		
- At indsamlet data må bruges i forbindelse med videre forskning		

Forudsætningen for denne samtykkeerklæring er, at alt materiale bliver opbevaret sikkert og fortroligt i henhold til Datatilsynets krav. Alt personfølsomt materiale bliver opbevaret indtil endt eksamen i juni 2019, hvorefter dette slettes. Alt ikke personfølsomt data vil blive gemt og anvendt til videre forskning indenfor feltet.

Navn: **Dato:**

Underskrift:

C | Data from eye tracking

C.1 End targets(E)

C.1.1 E1

Figure C.1: Stimuli E1

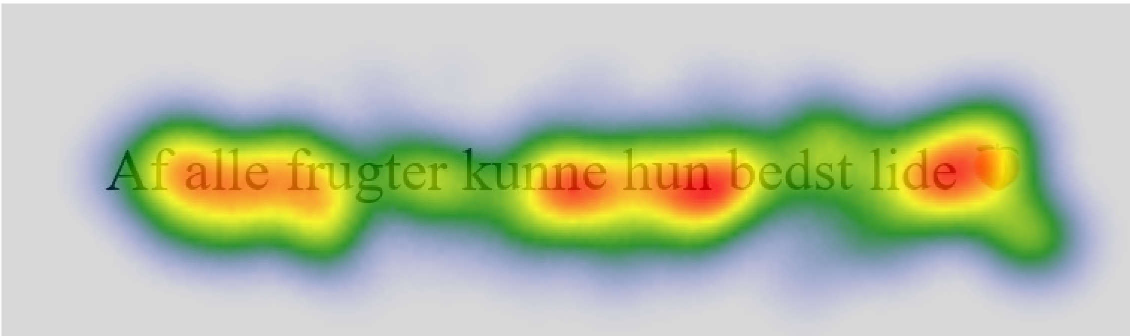

Figure C.2: Heat map E1

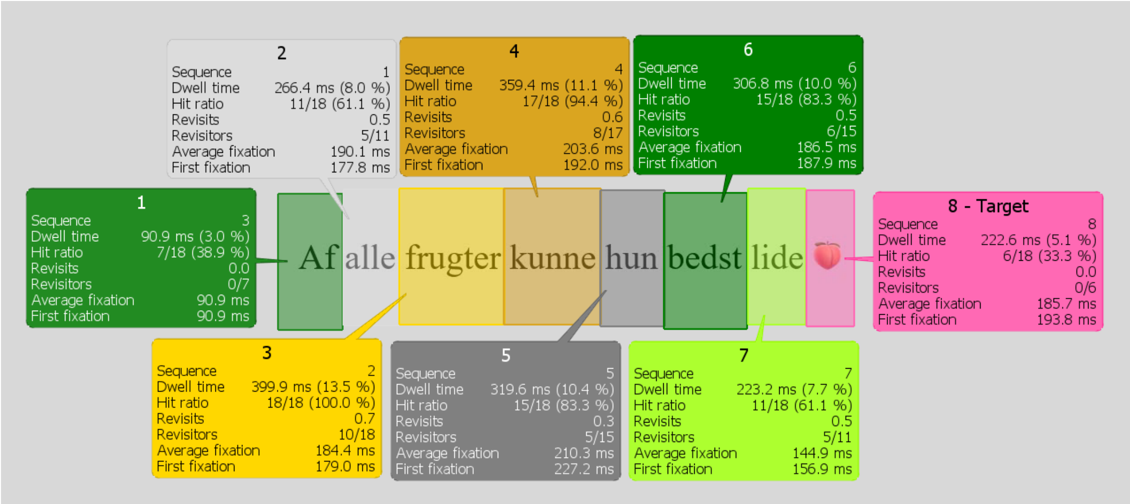

Figure C.3: Metrics E1

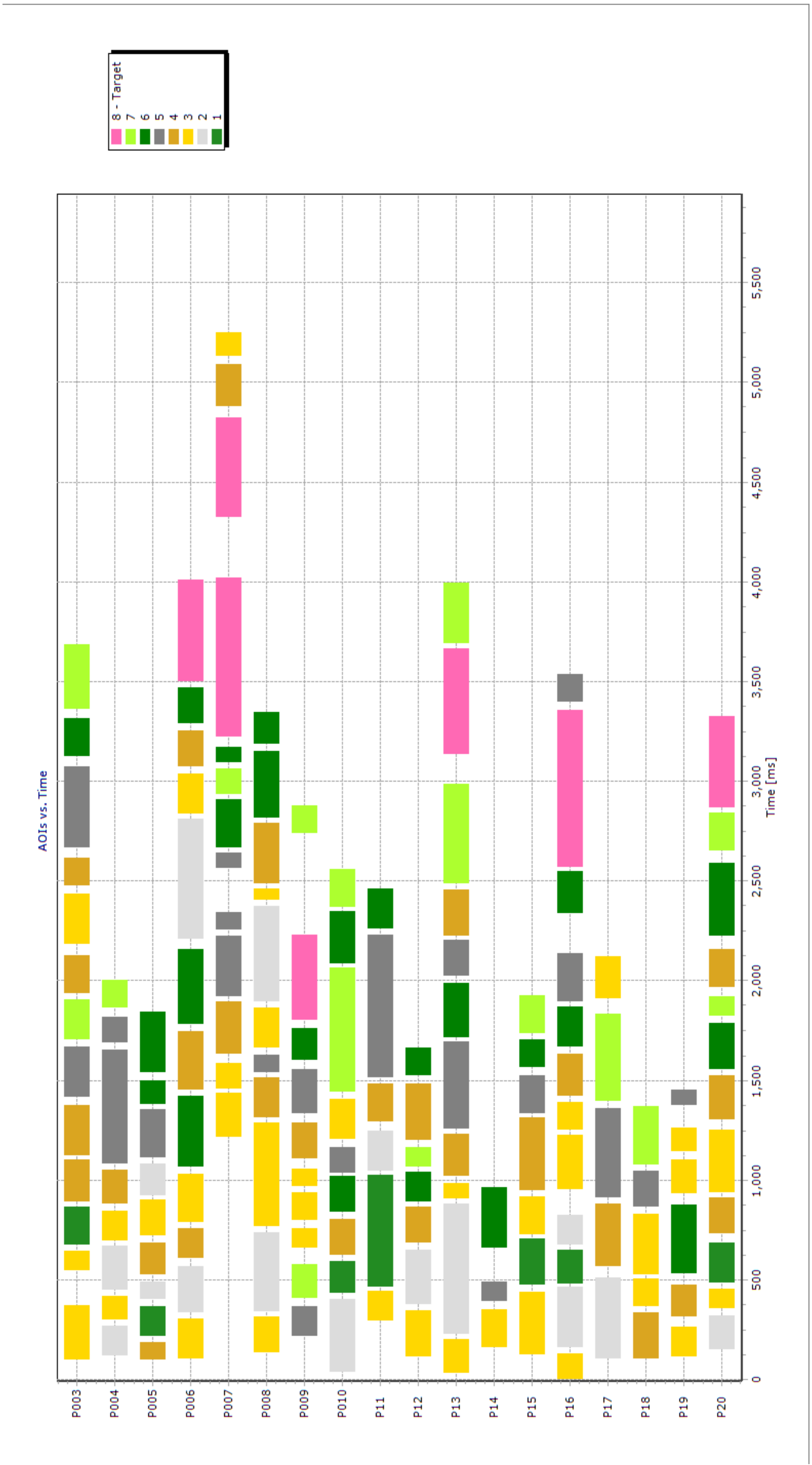

Figure C.4: AOI sequence chart for fixations E1

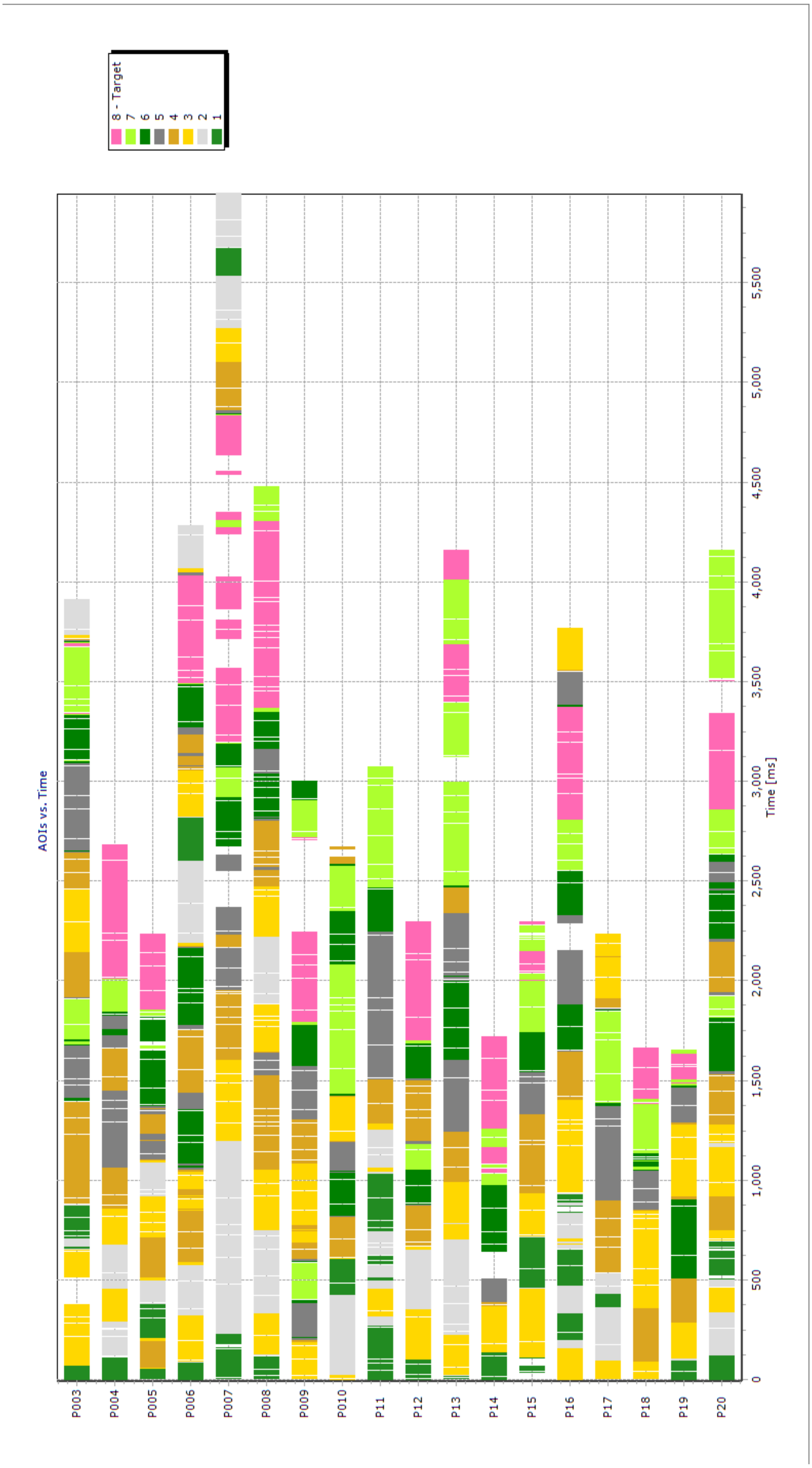

Figure C.5: AOI sequence chart for raw data E1

C.1.2 E1x

Figure C.6: Stimuli E1x

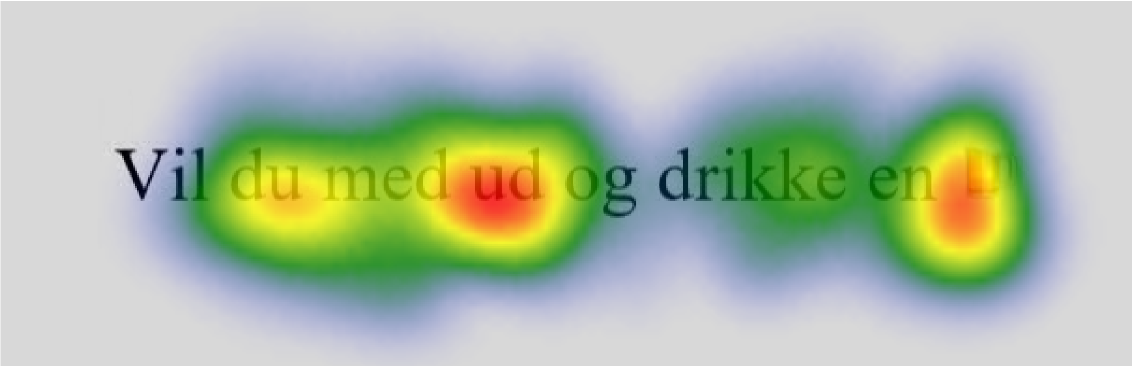

Figure C.7: Heat map E1x

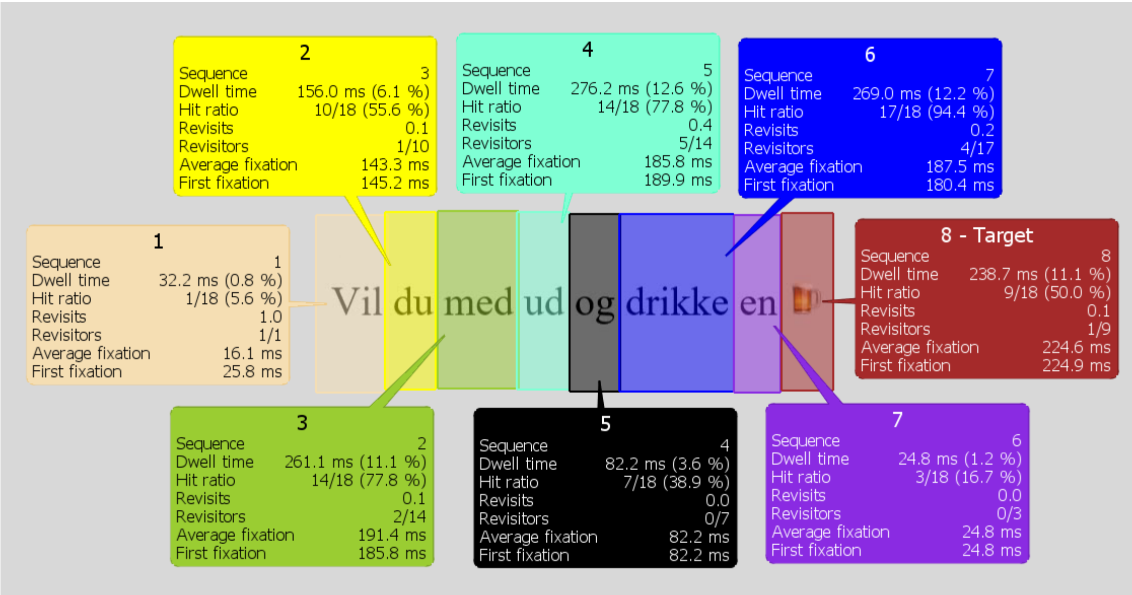

Figure C.8: Metrics E1x

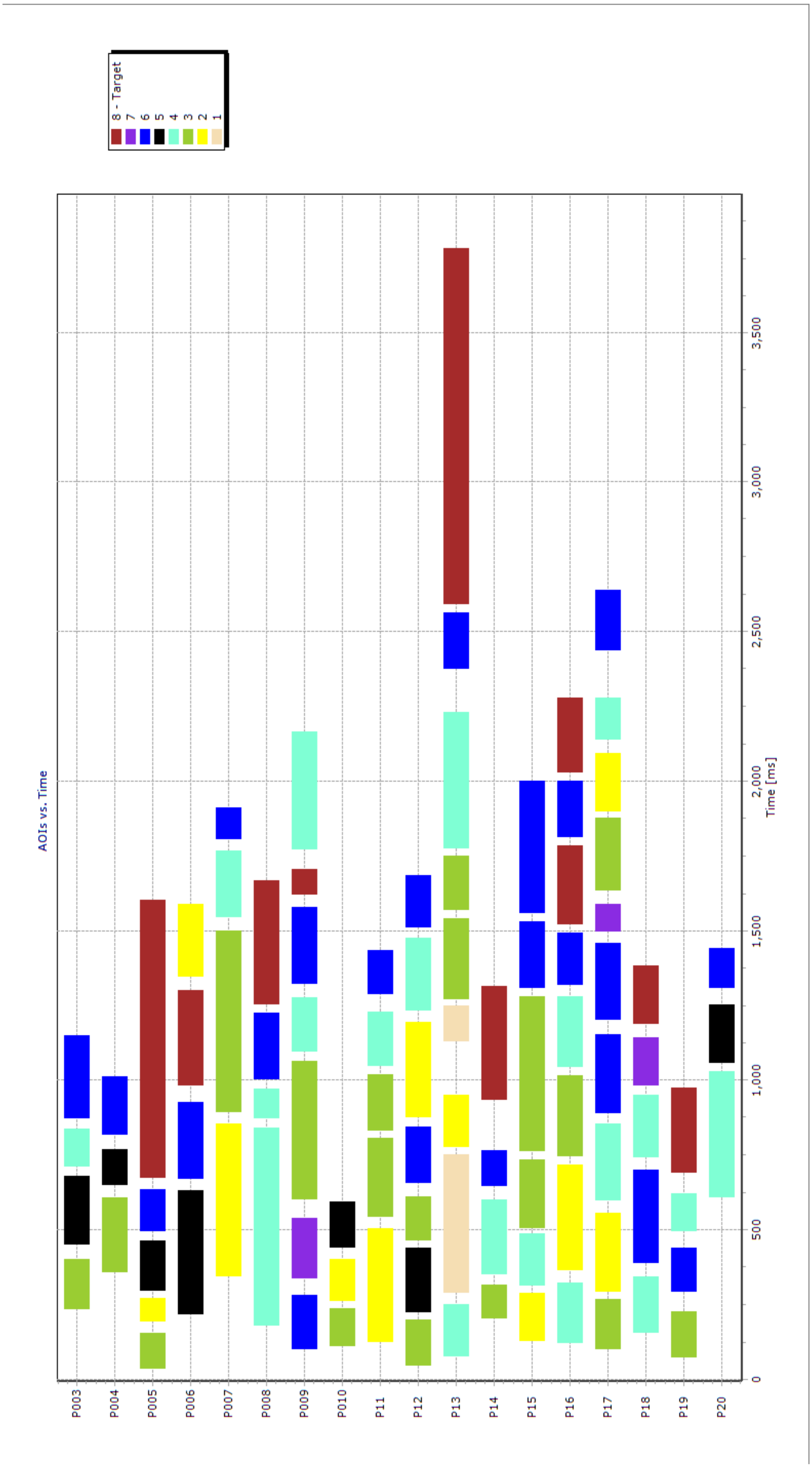

Figure C.9: AOI sequence chart for fixations E1x

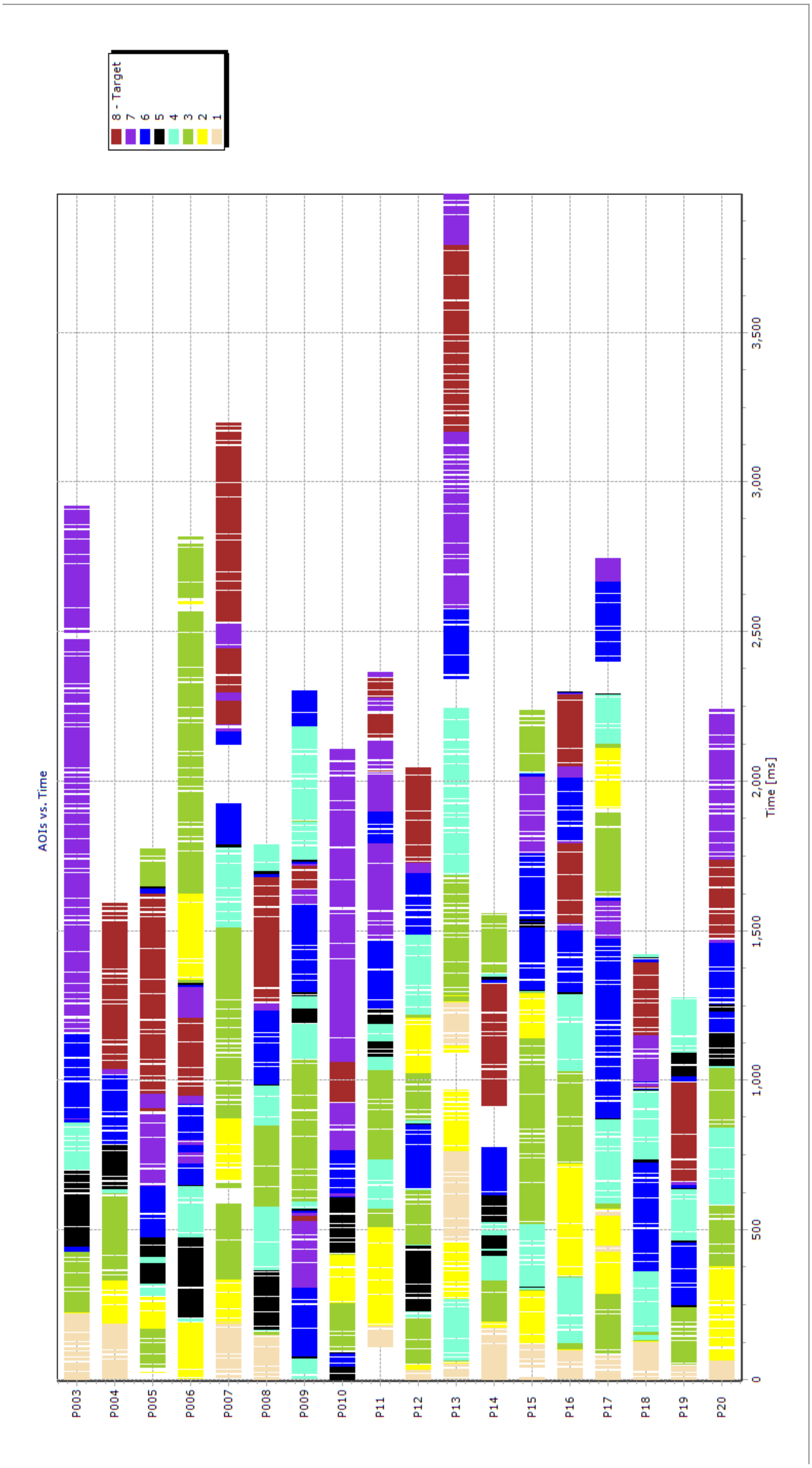

Figure C.10: AOI sequence chart for raw data E1x

C.1.3 E2

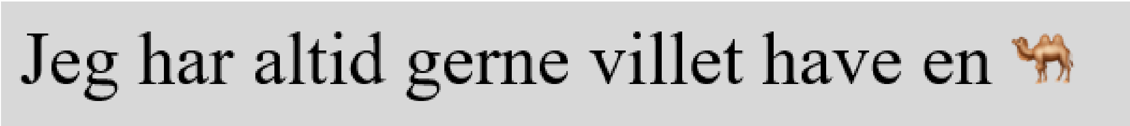

Figure C.11: Stimuli E2

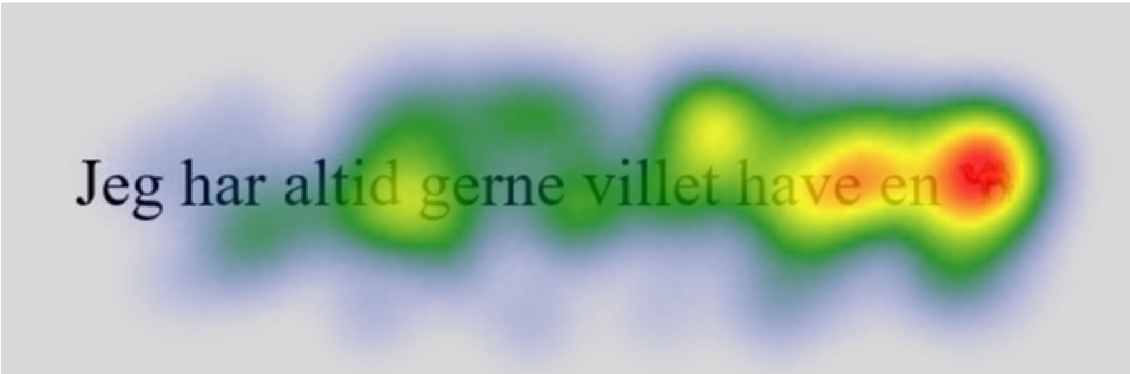

Figure C.12: Heat map E2

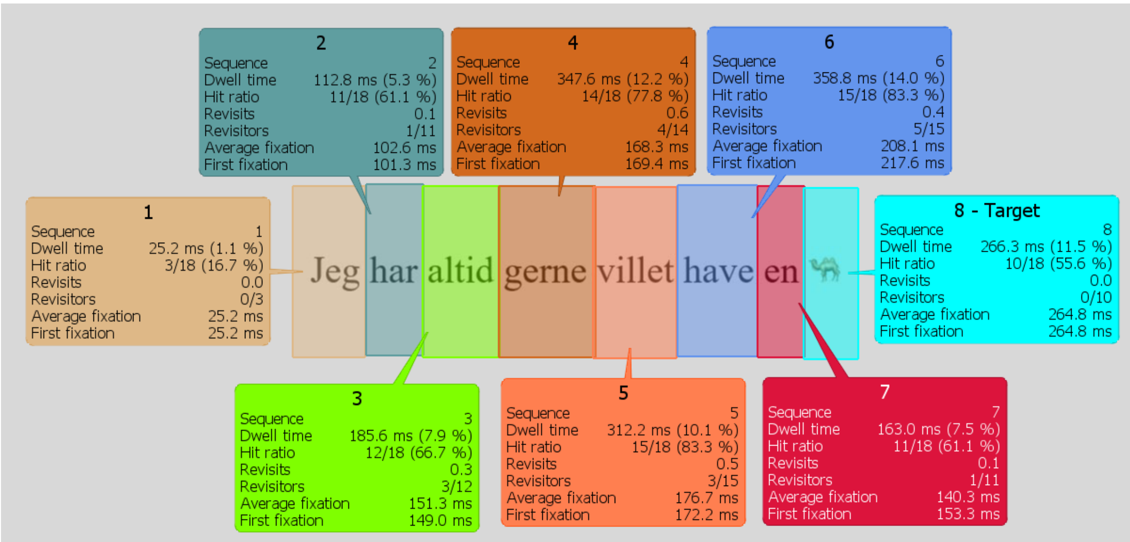

Figure C.13: Metrics E2

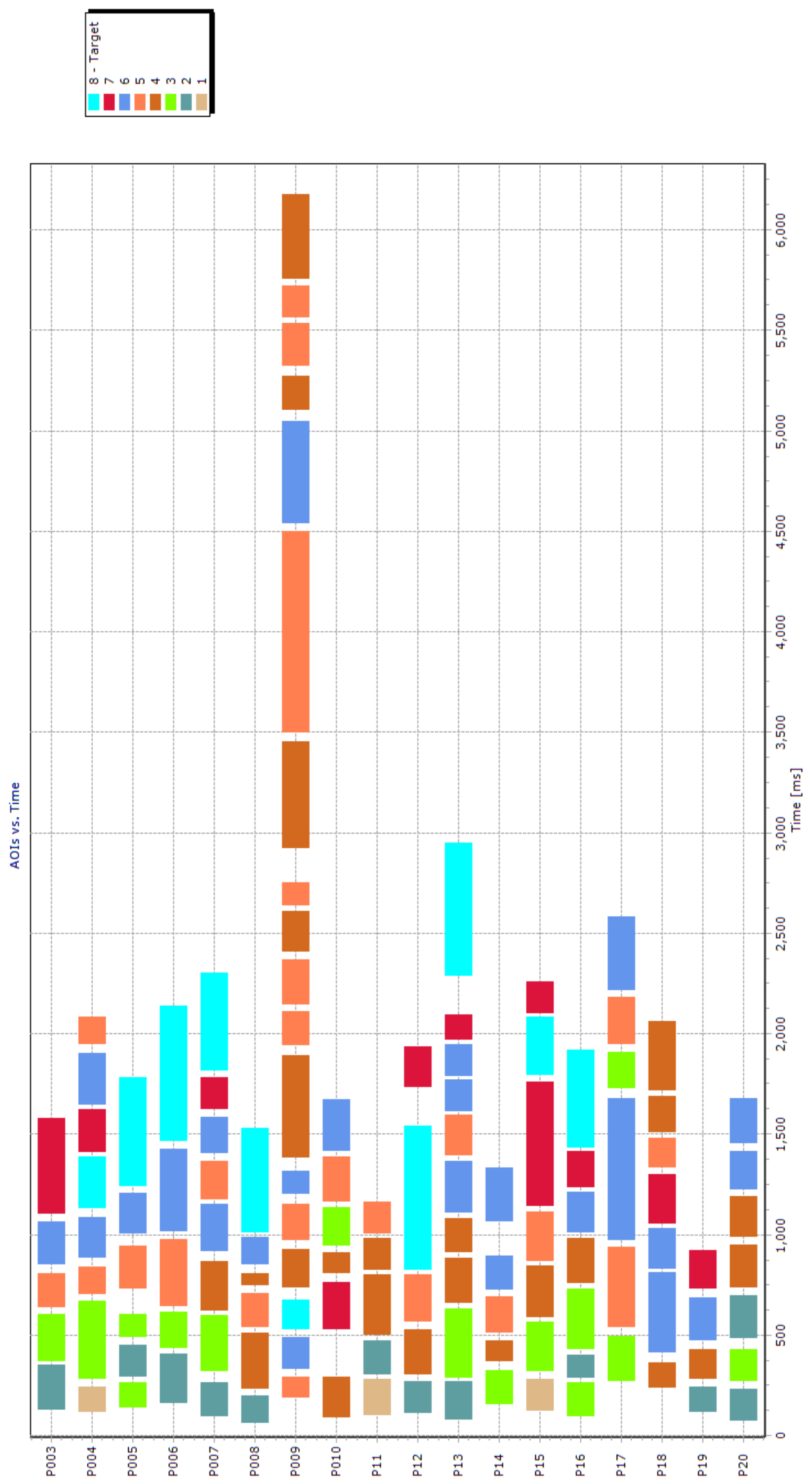

Figure C.14: AOI sequence chart for fixations E2

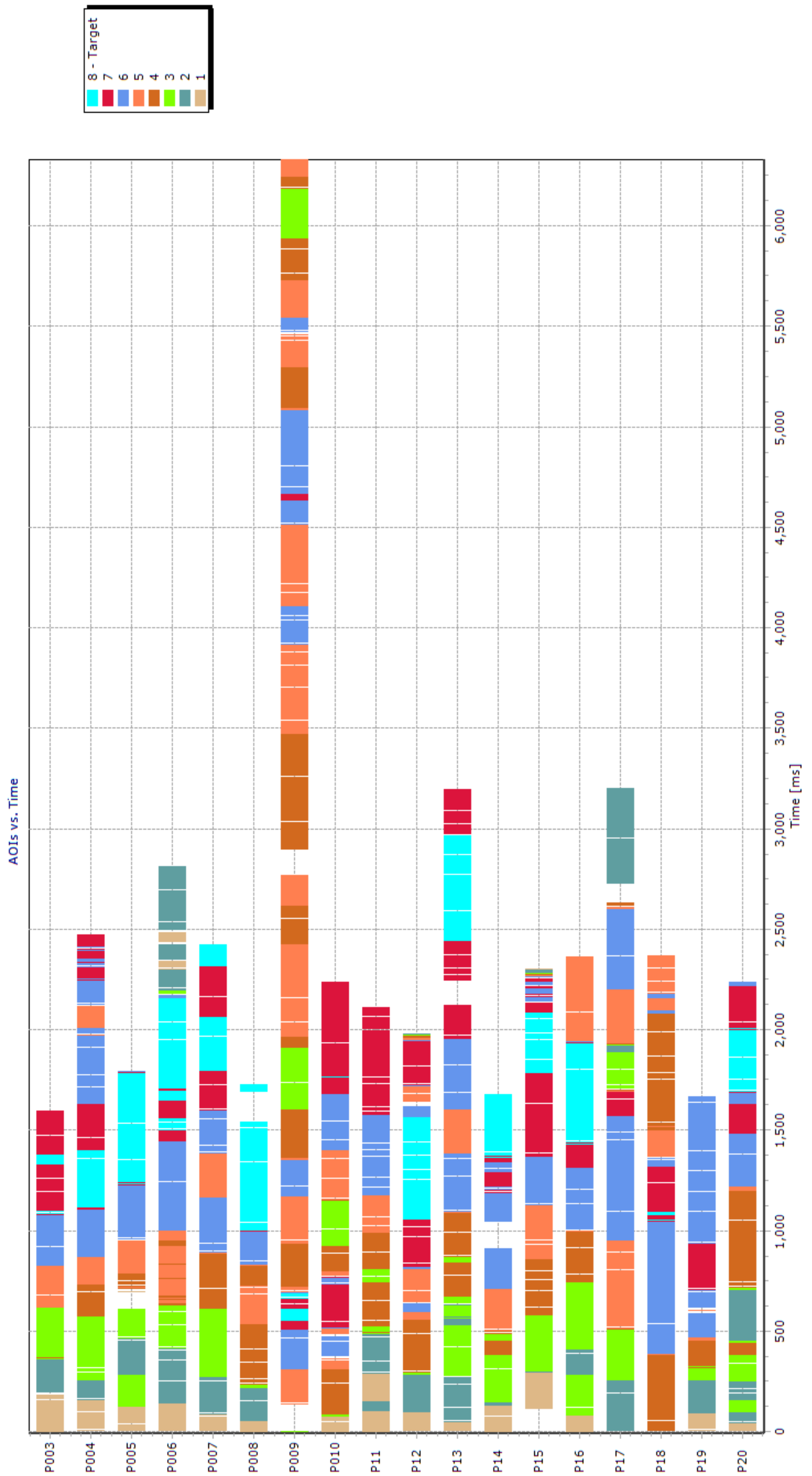

Figure C.15: AOI sequence chart for raw data E2

C.1.4 E2x

Figure C.16: Stimuli E2x

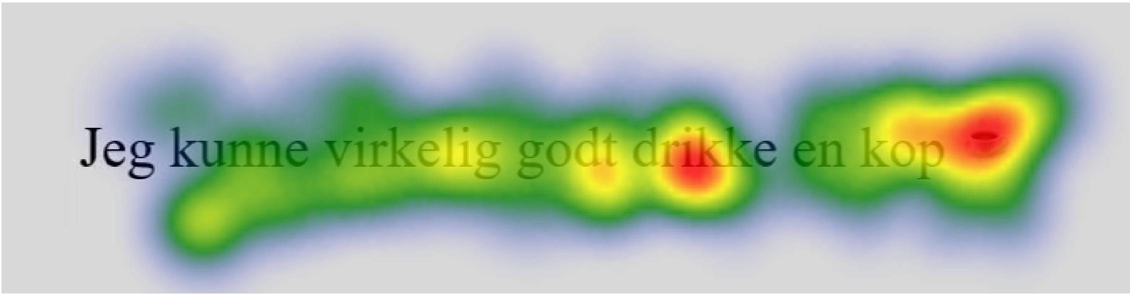

Figure C.17: Heat map E2x

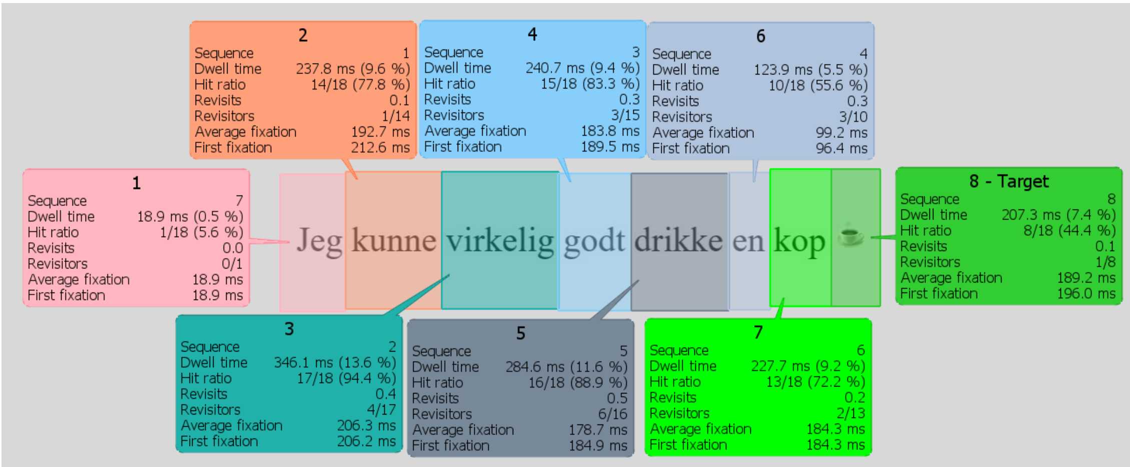

Figure C.18: Metrics E2x

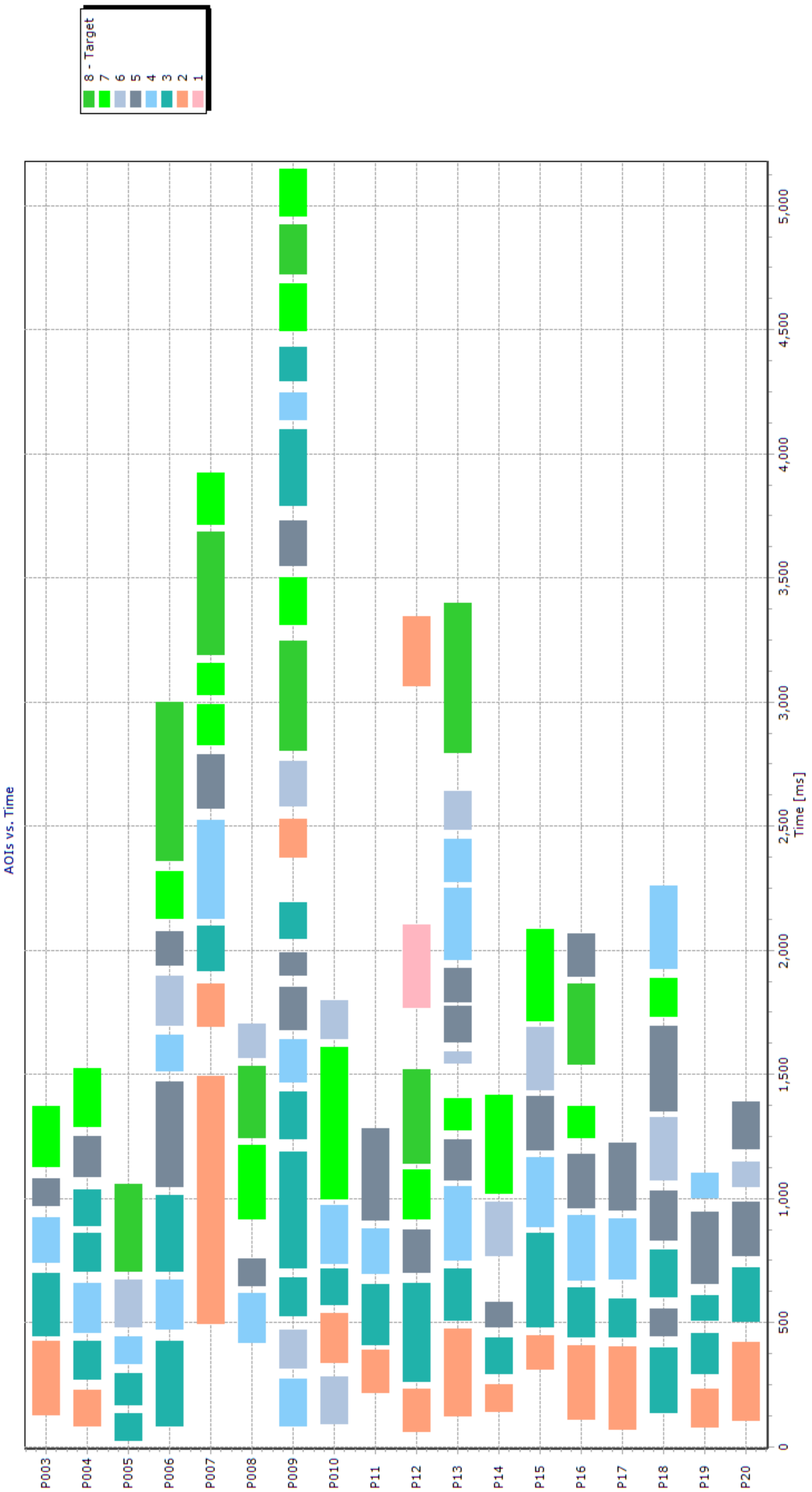

Figure C.19: AOI sequence chart for fixations E2x

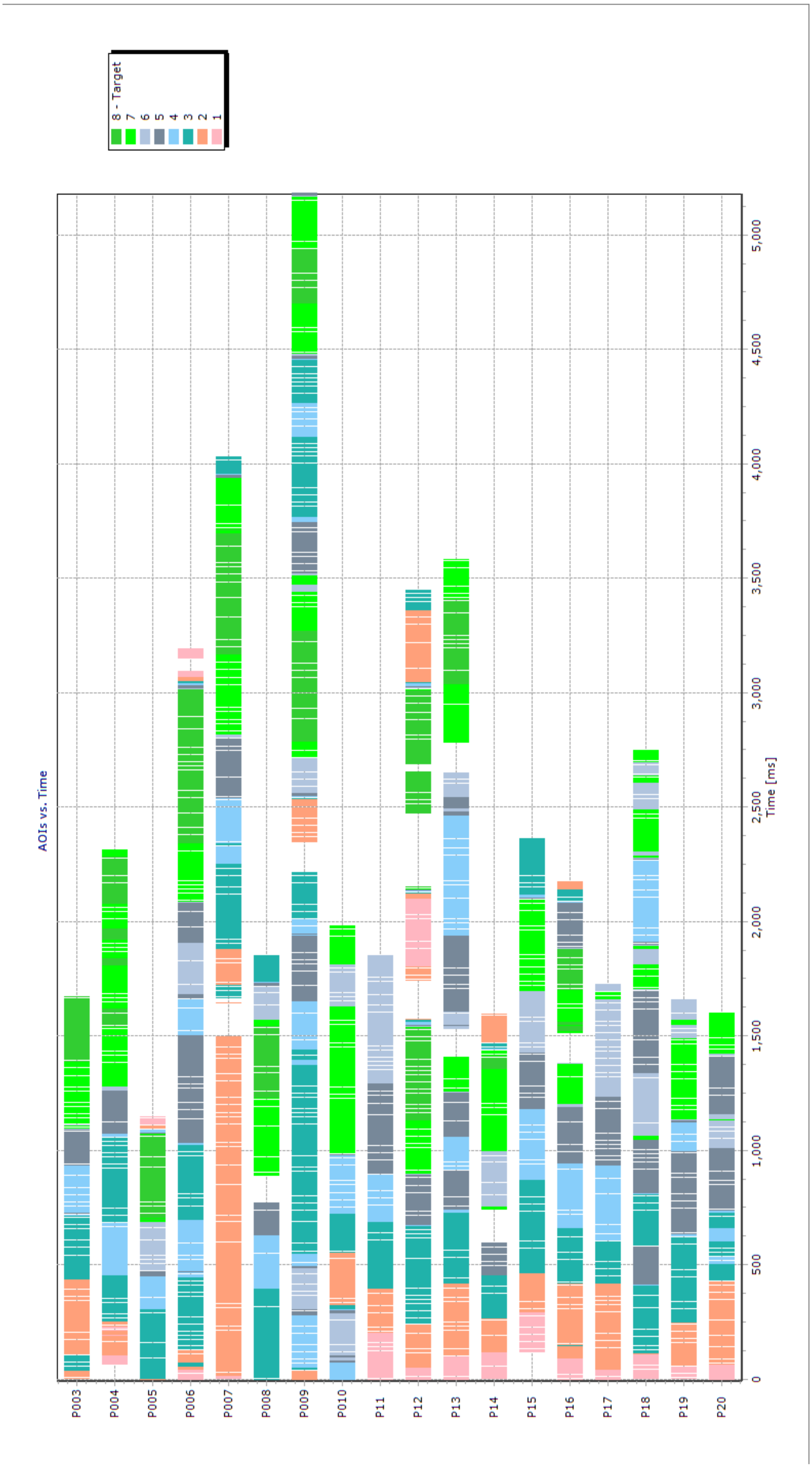

Figure C.20: AOI sequence chart for raw data E2x

C.1.5 E3

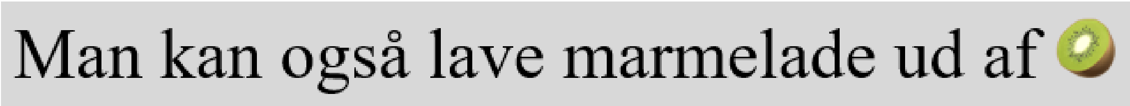

Figure C.21: Stimuli E3

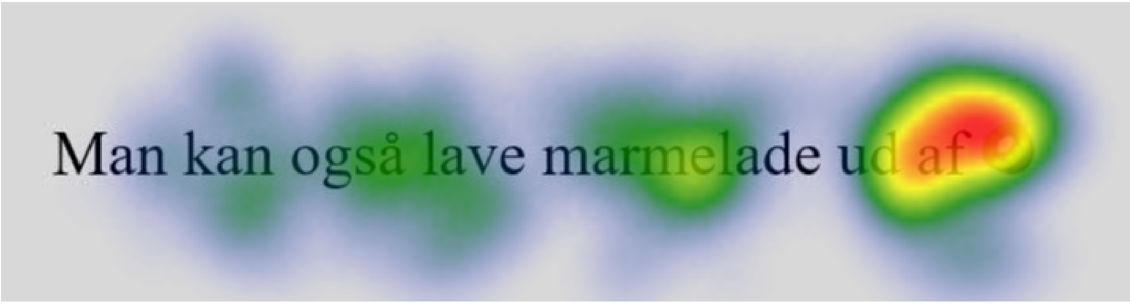

Figure C.22: Heat map E3

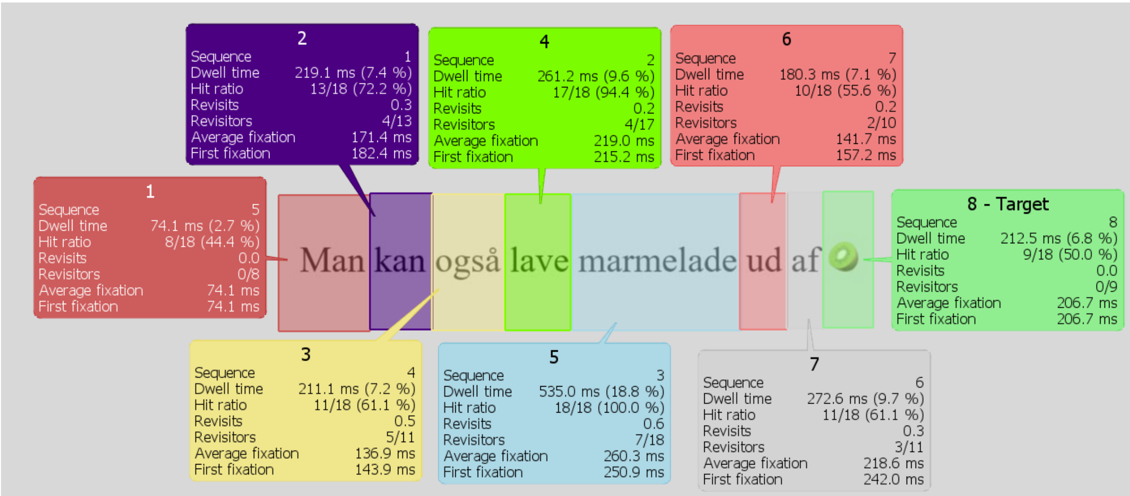

Figure C.23: Metrics E3

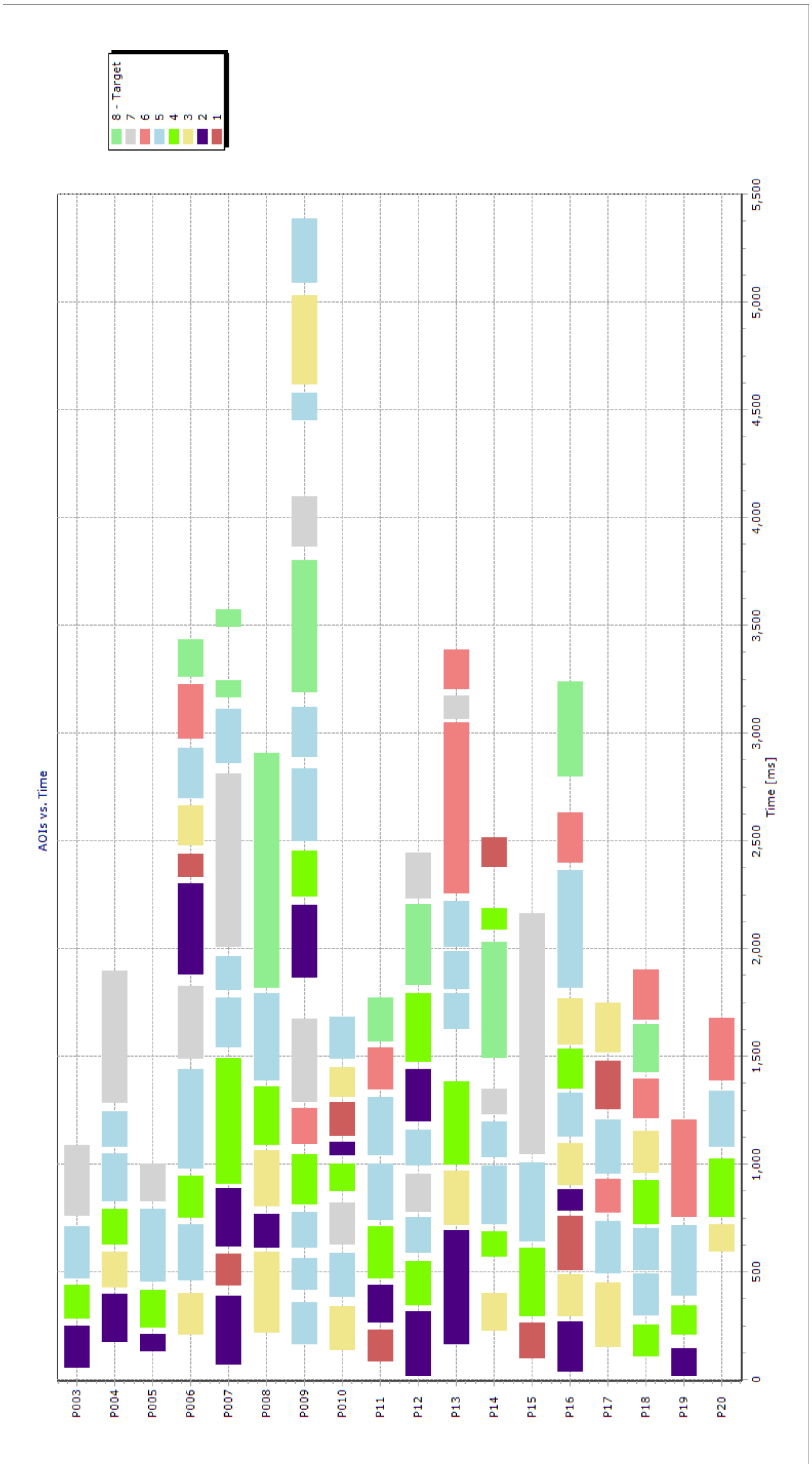

Figure C.24: AOI sequence chart for fixations E3

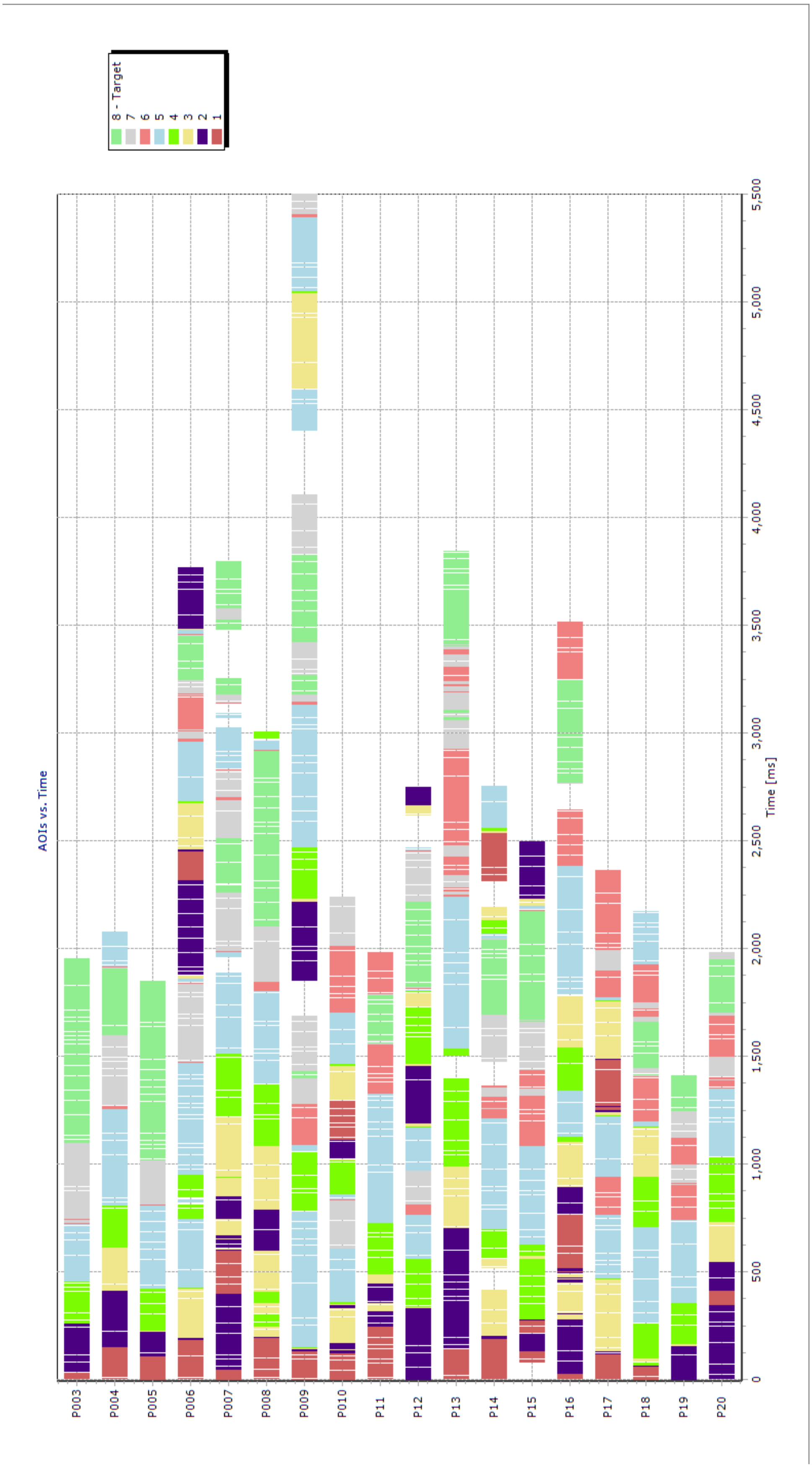

Figure C.25: AOI sequence chart for raw data E3

C.1.6 E3x

Figure C.26: Stimuli E3x

Figure C.27: Heat map E3x

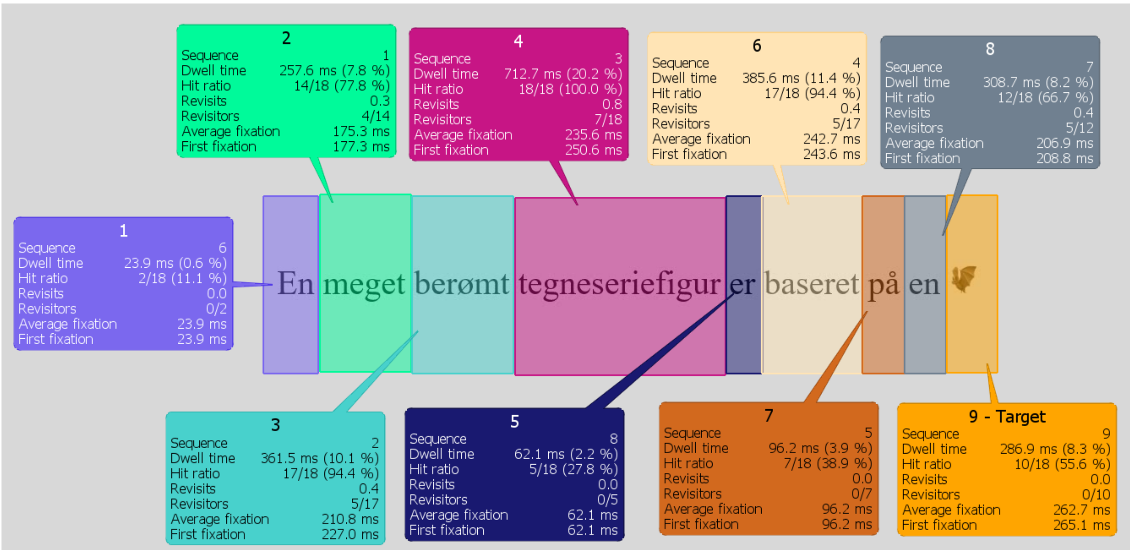

Figure C.28: Metrics E3x

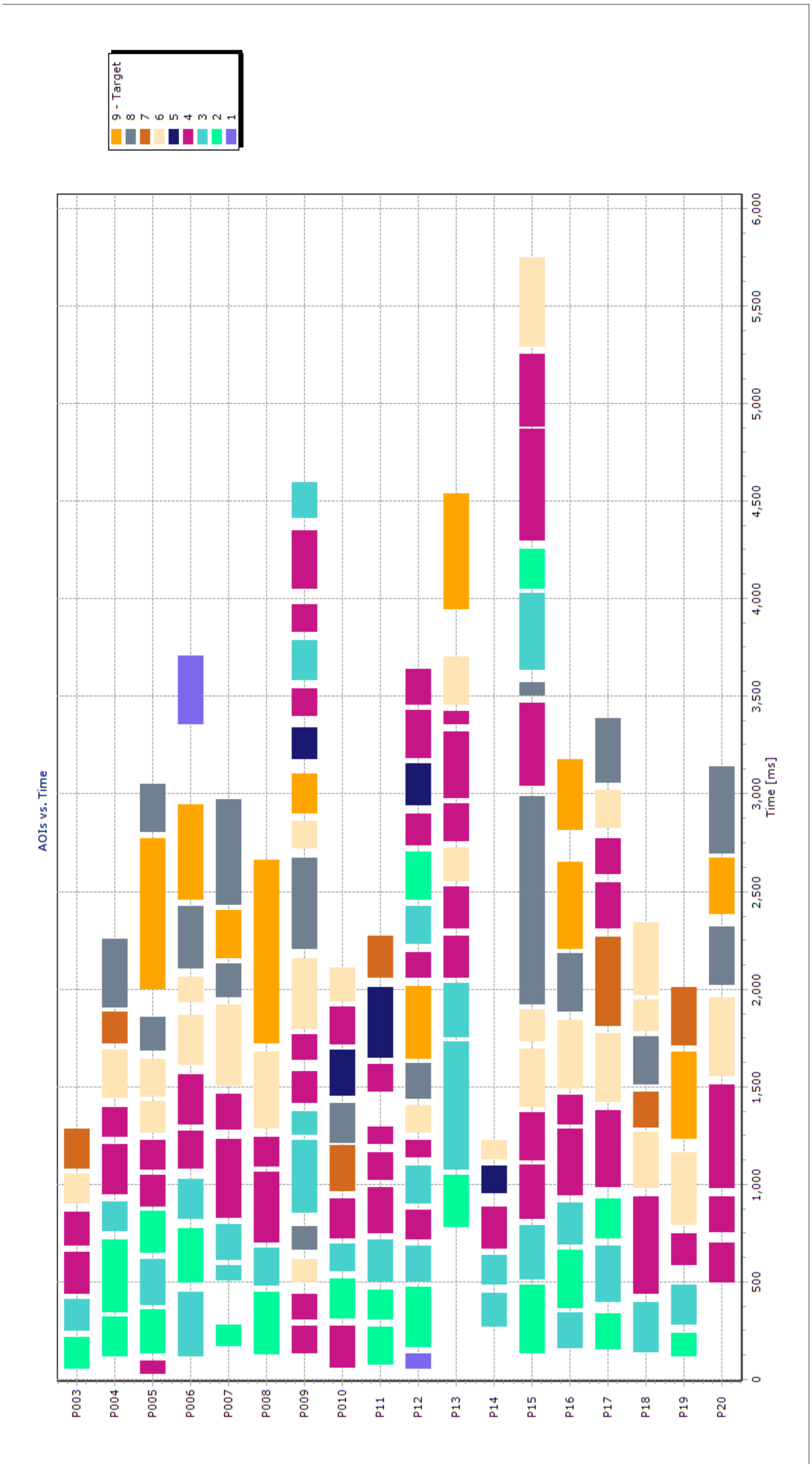

Figure C.29: AOI sequence chart for fixations E3x

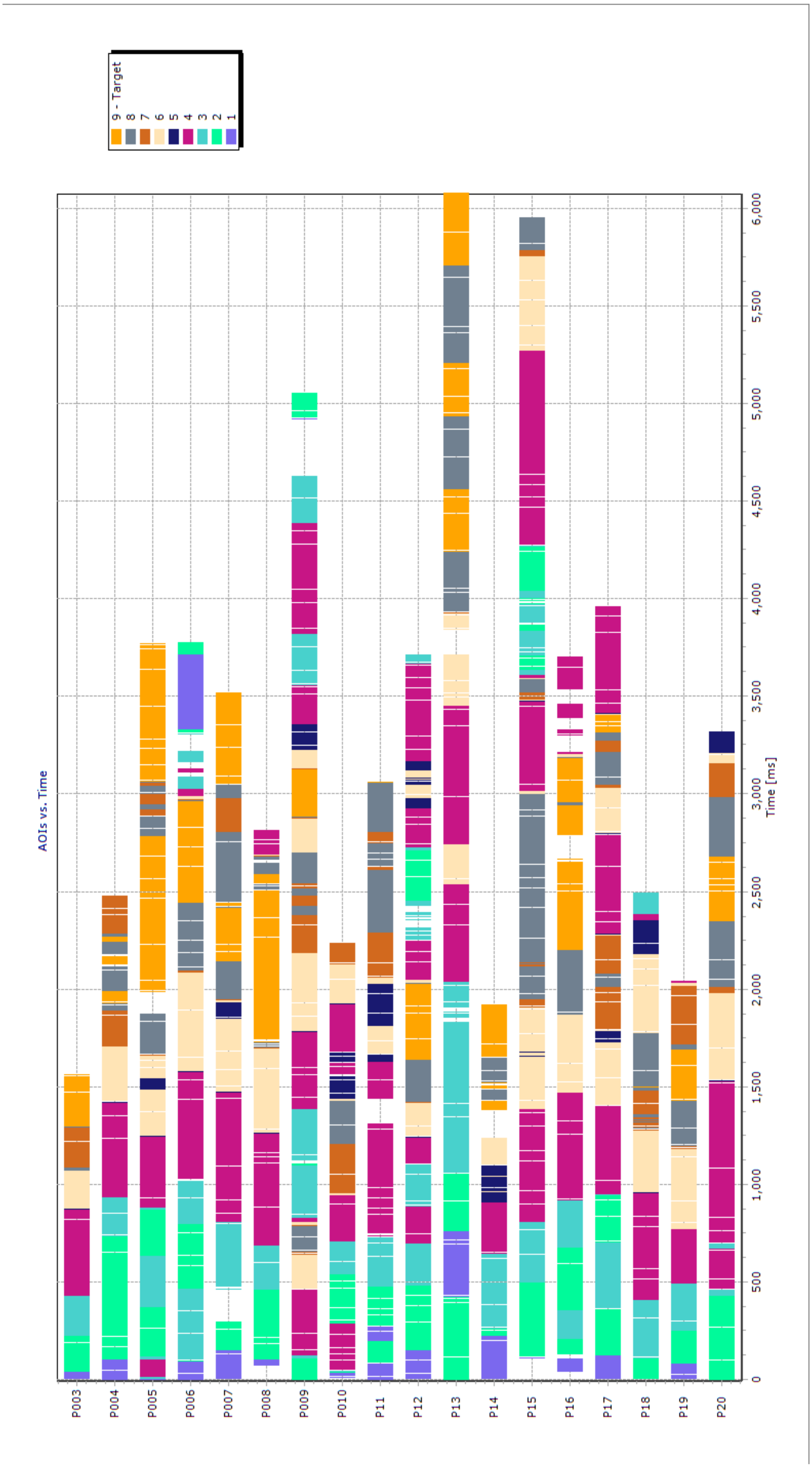

Figure C.30: AOI sequence chart for raw data E3x

C.1.7 E4

Figure C.31: Stimuli E4

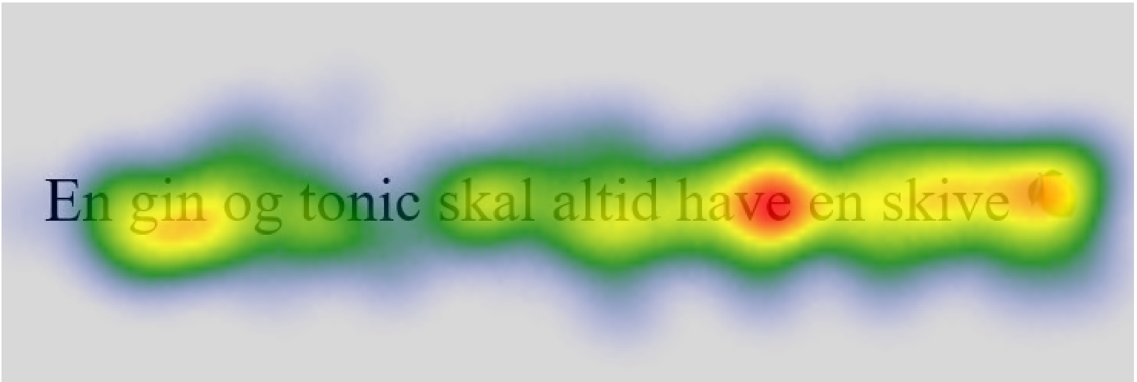

Figure C.32: Heat map E4

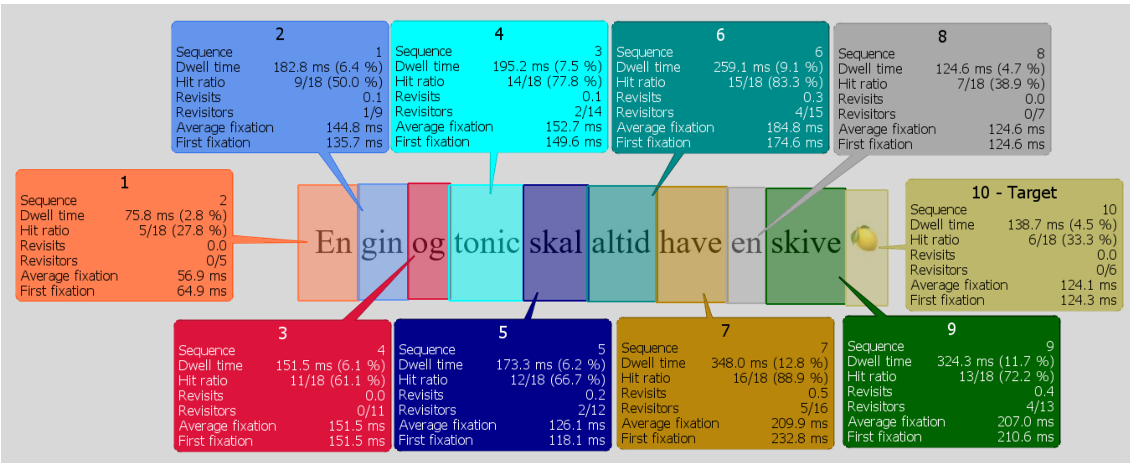

Figure C.33: Metrics E4

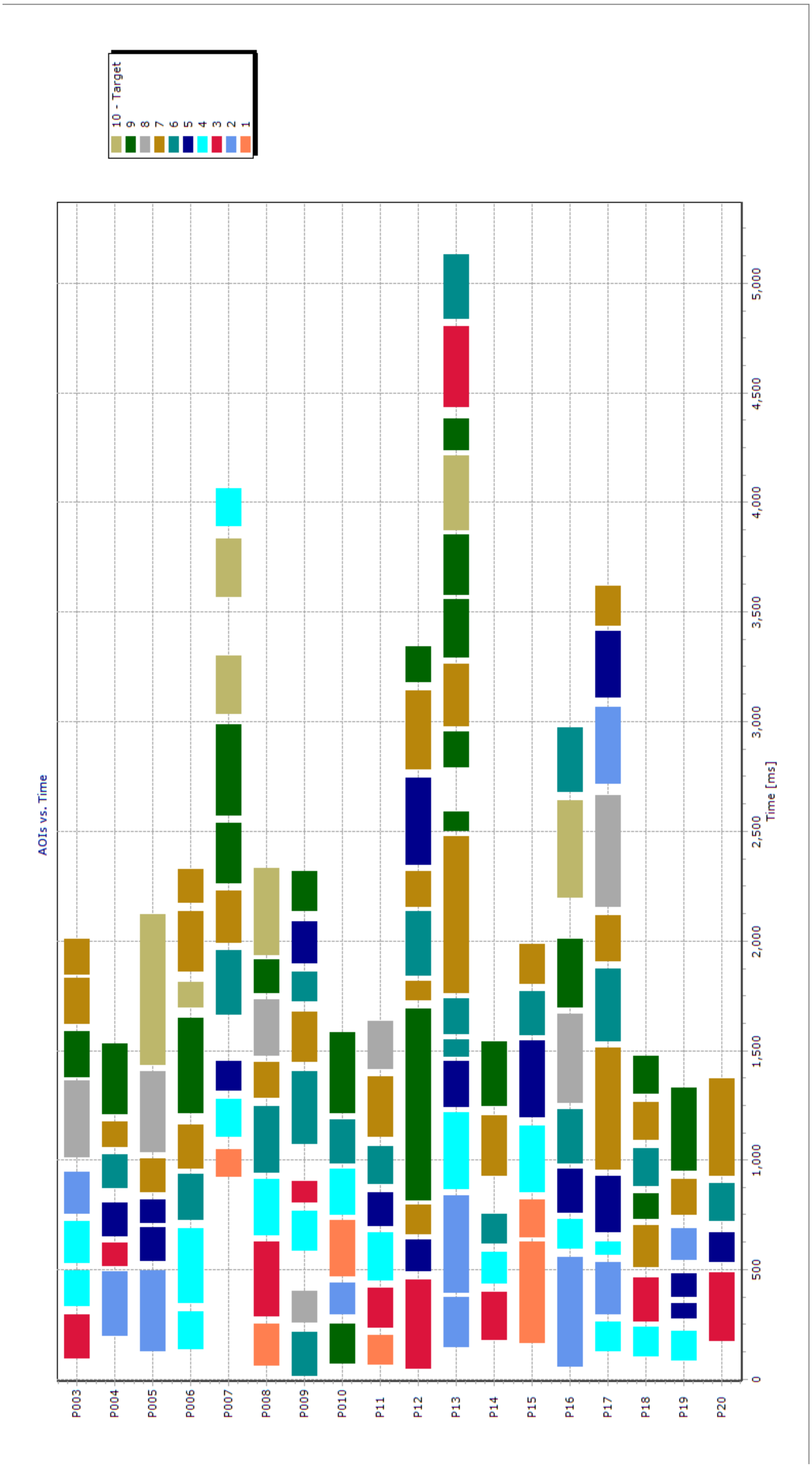

Figure C.34: AOI sequence chart for fixations E4

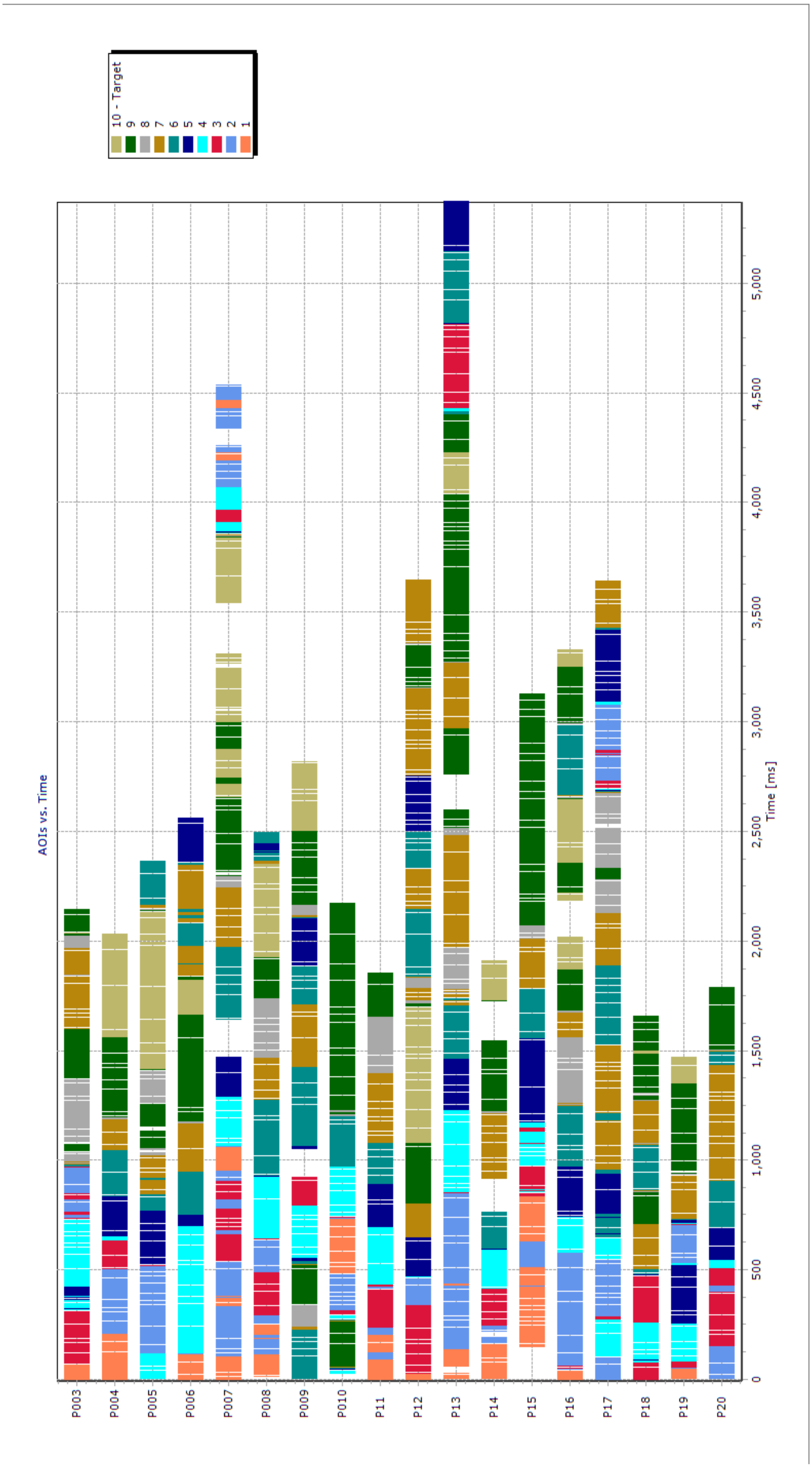

Figure C.35: AOI sequence chart for raw data E4

C.1.8 E5

Figure C.36: Stimuli E5

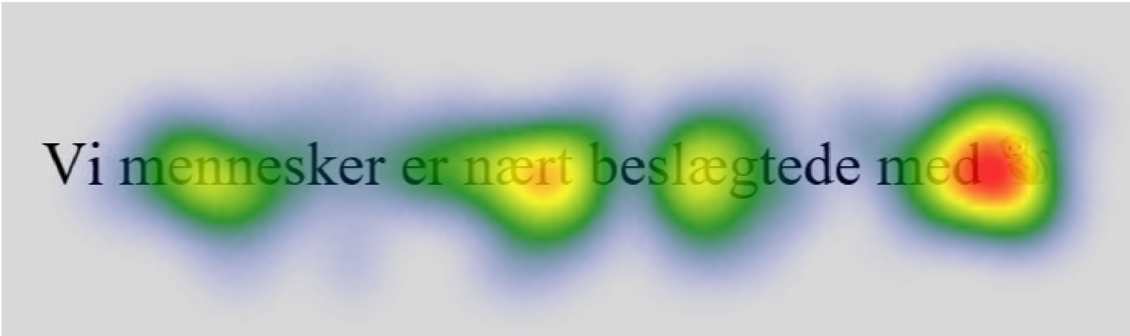

Figure C.37: Heat map E5

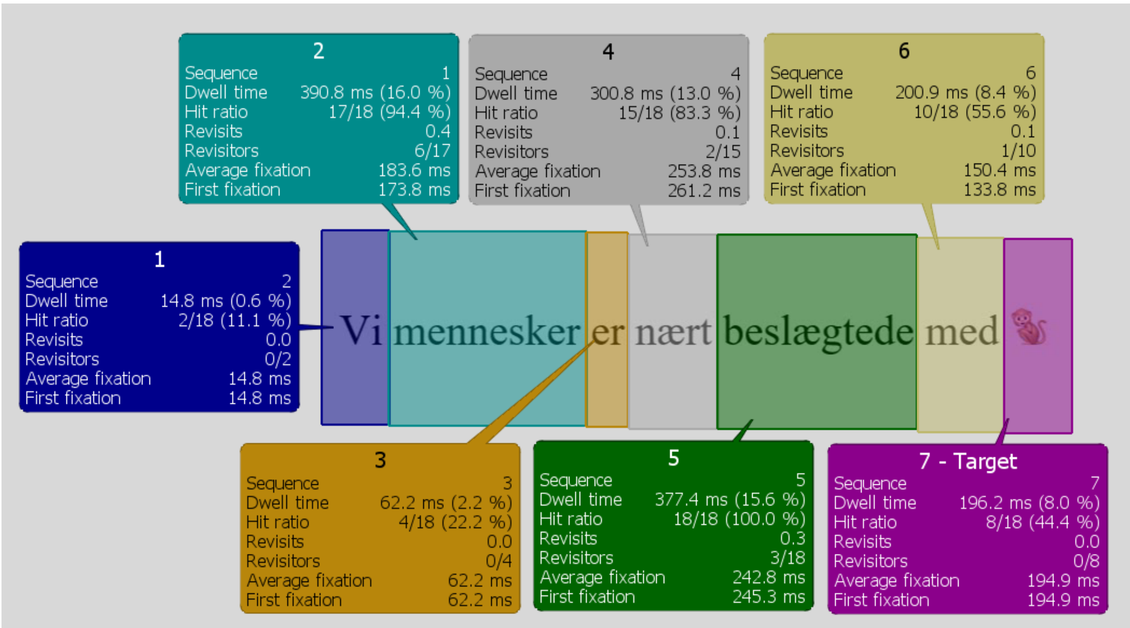

Figure C.38: Metrics E5

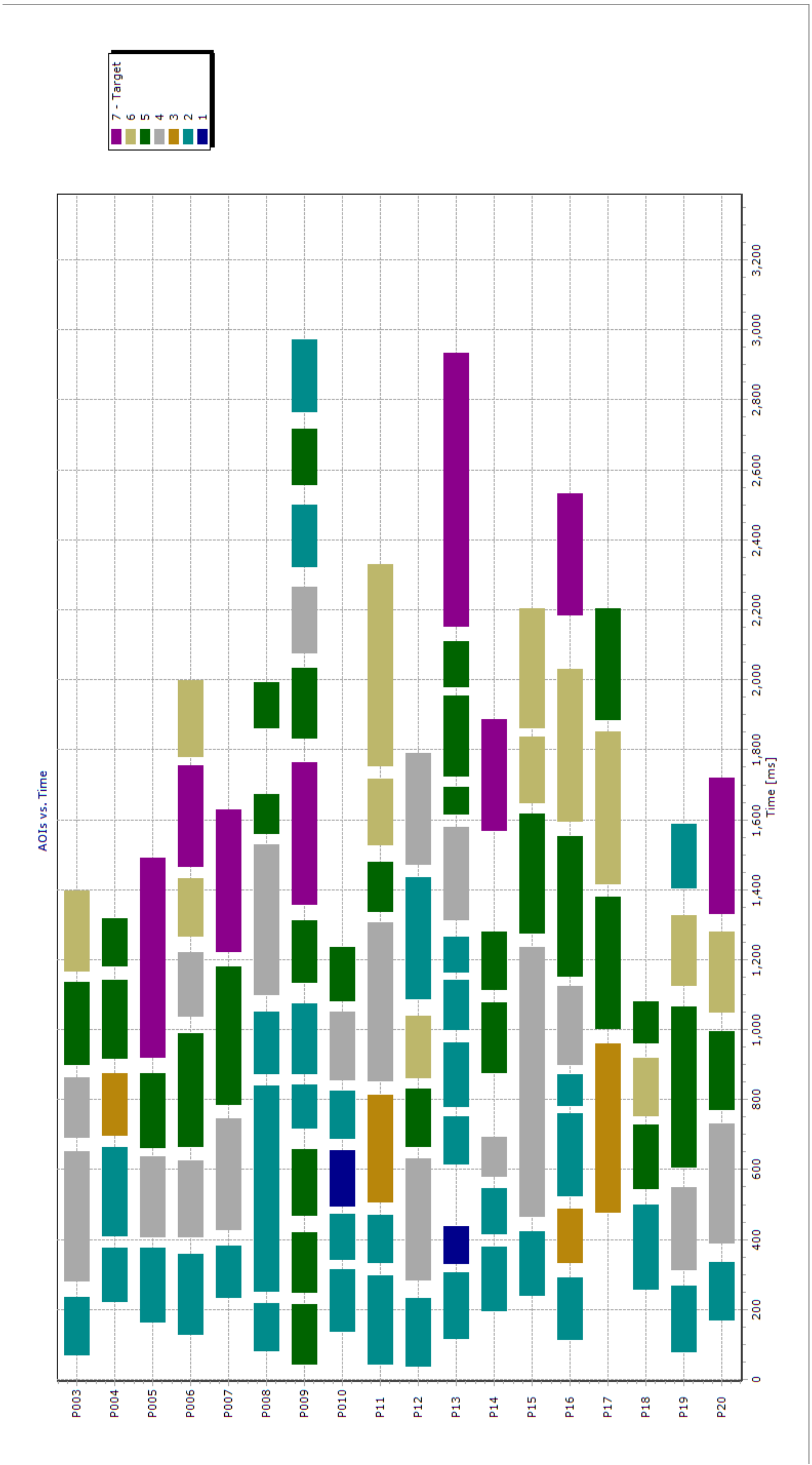

Figure C.39: AOI sequence chart for fixations E5

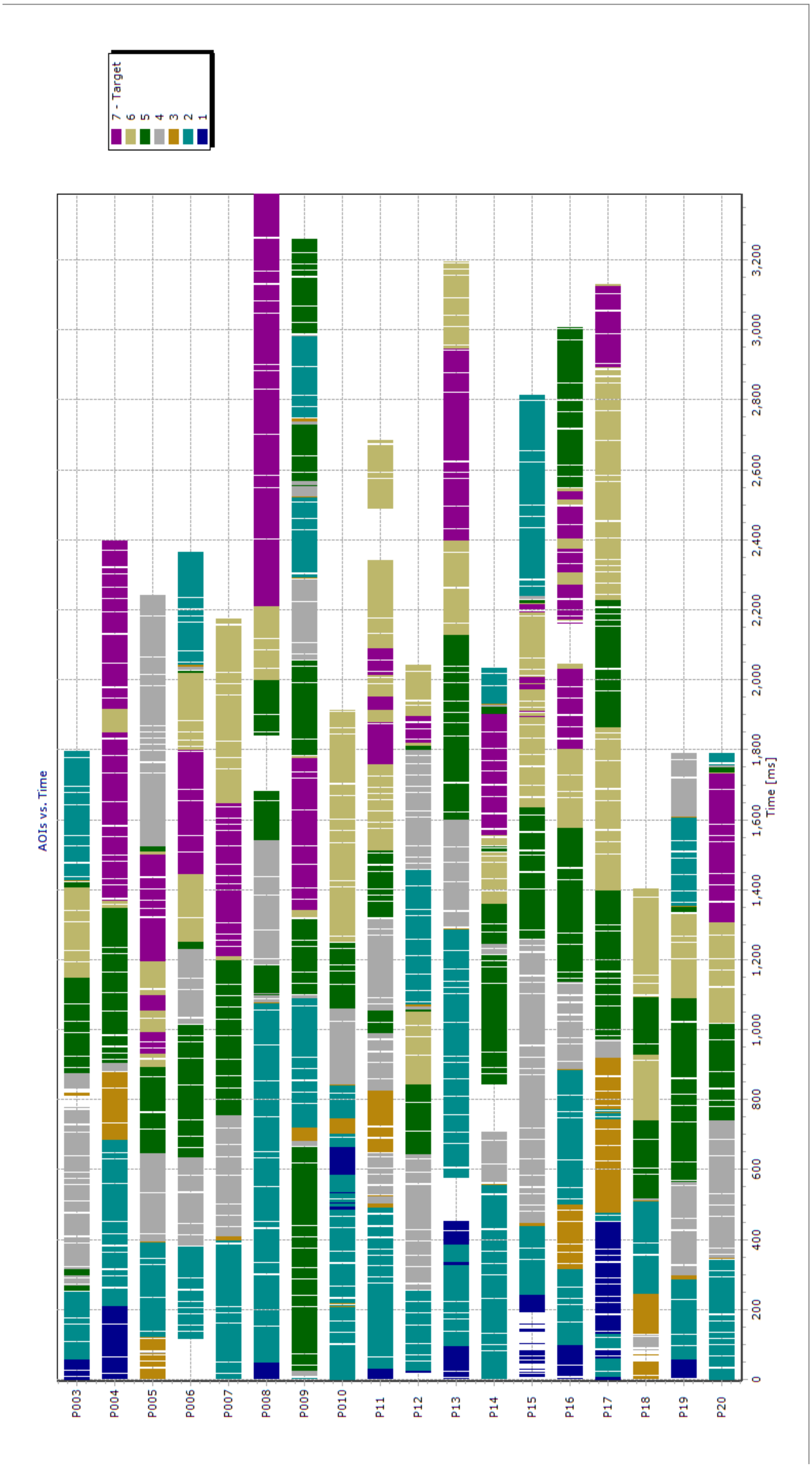

Figure C.40: AOI sequence chart for raw data E5

C.1.9 E6

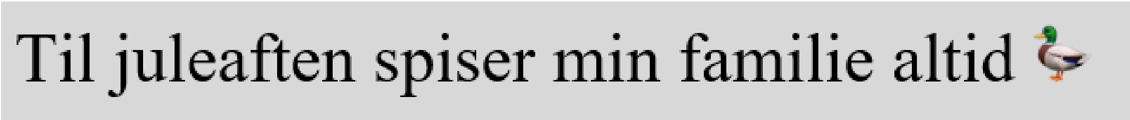

Figure C.41: Stimuli E6

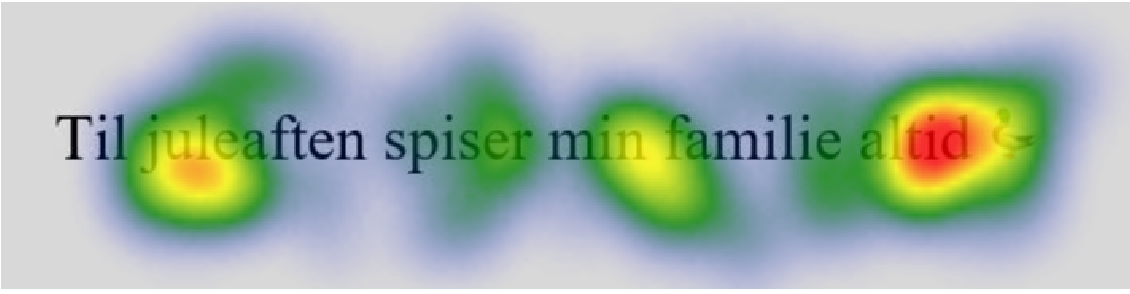

Figure C.42: Heat map E6

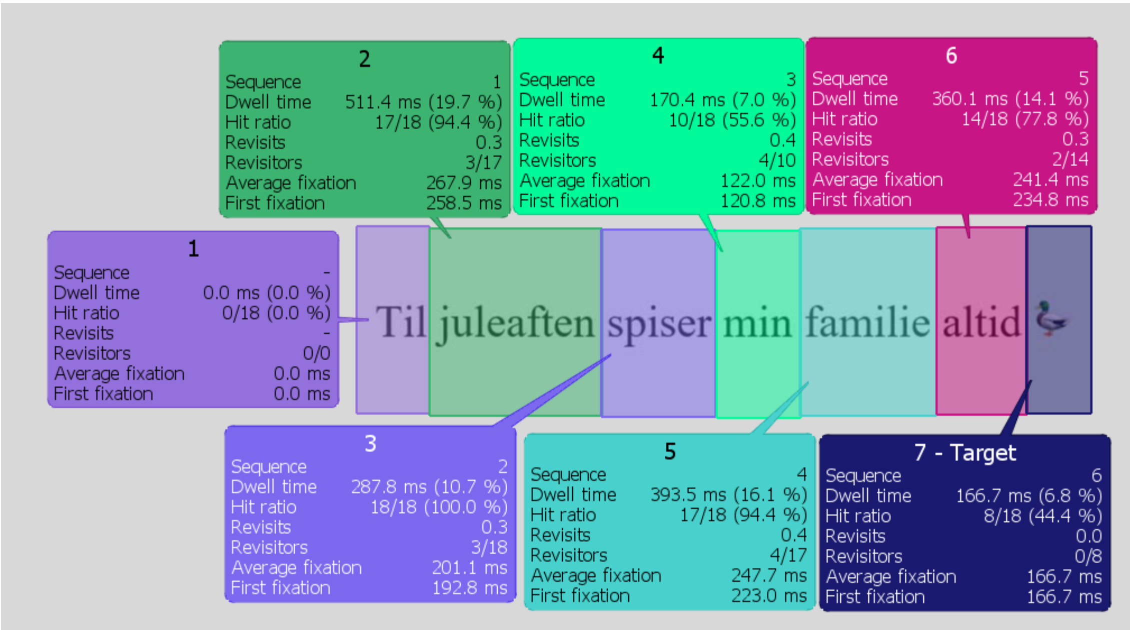

Figure C.43: Metrics E6

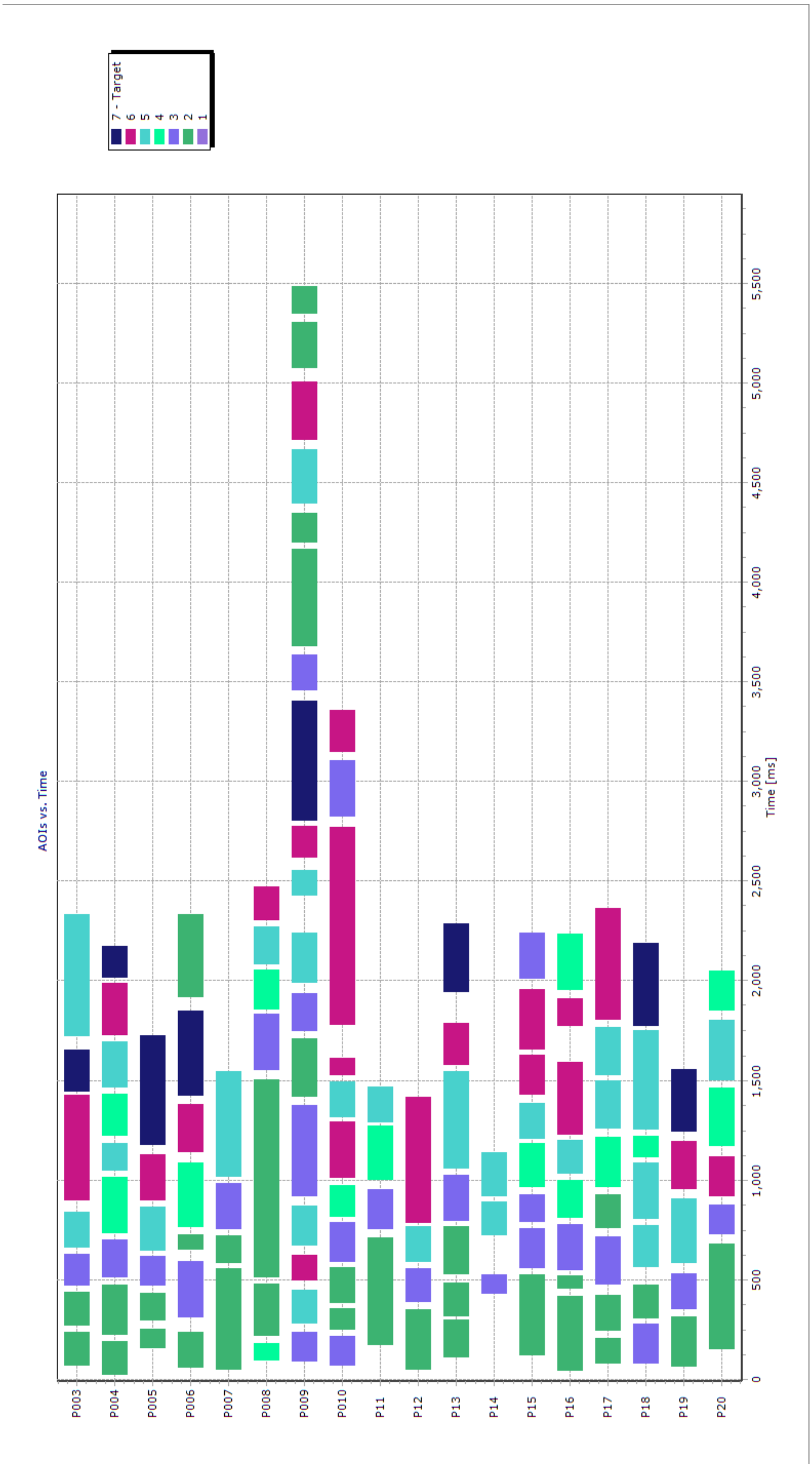

Figure C.44: AOI sequence chart for fixations E6

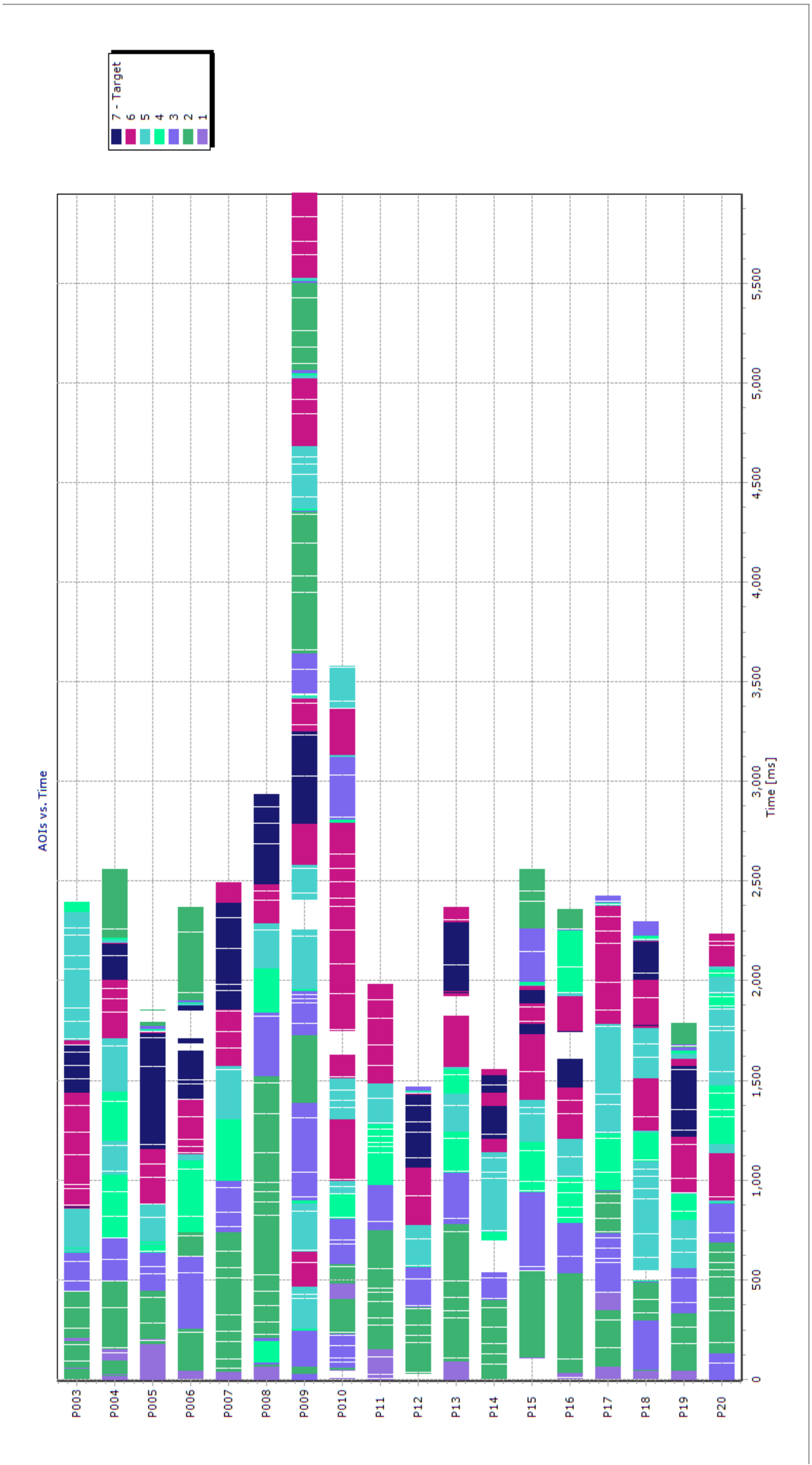

Figure C.45: AOI sequence chart for raw data E6

C.2 Front targets(F)

C.2.1 F1

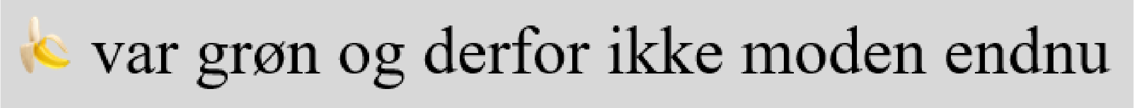

Figure C.46: Stimuli F1

Figure C.47: Heat map F1

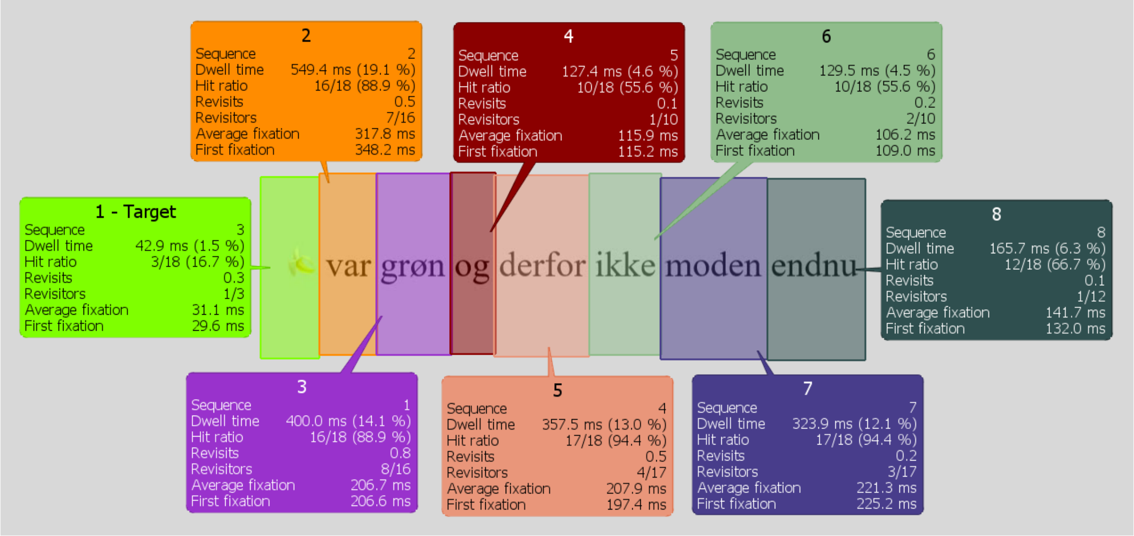

Figure C.48: Metrics F1

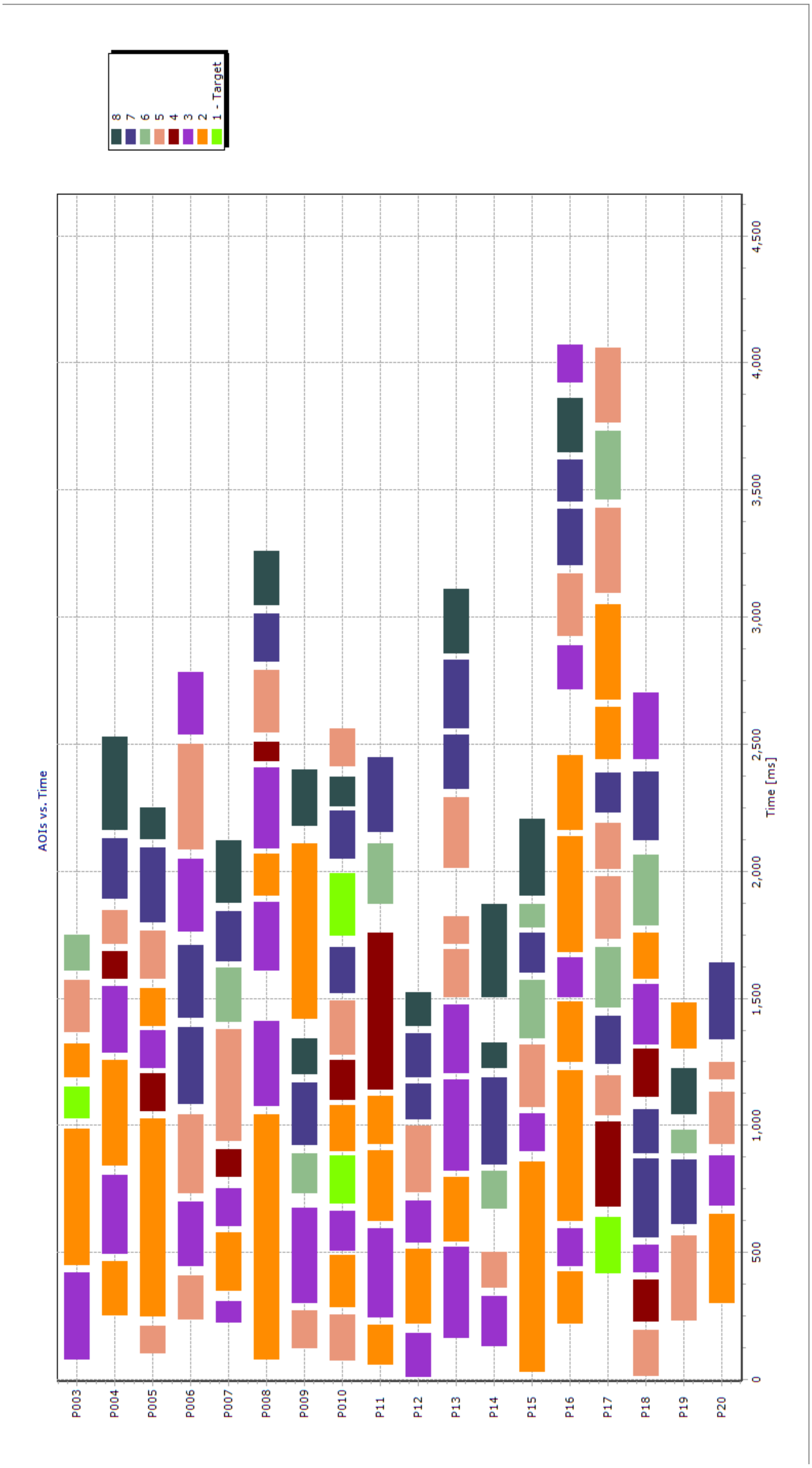

Figure C.49: AOI sequence chart for fixations F1

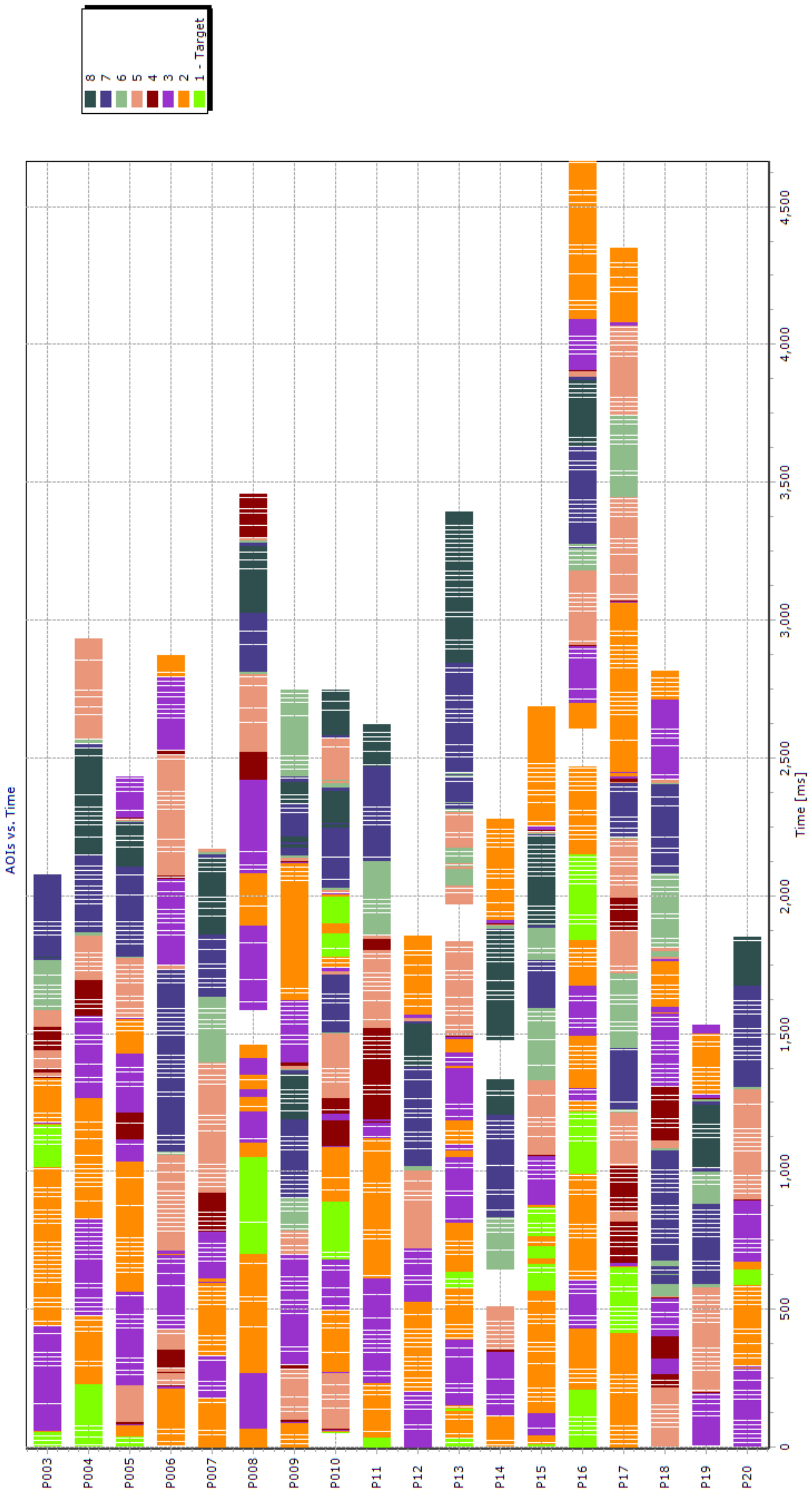

Figure C.50: AOI sequence chart for raw data F1

C.2.2 F1x

Figure C.51: Stimuli F1x

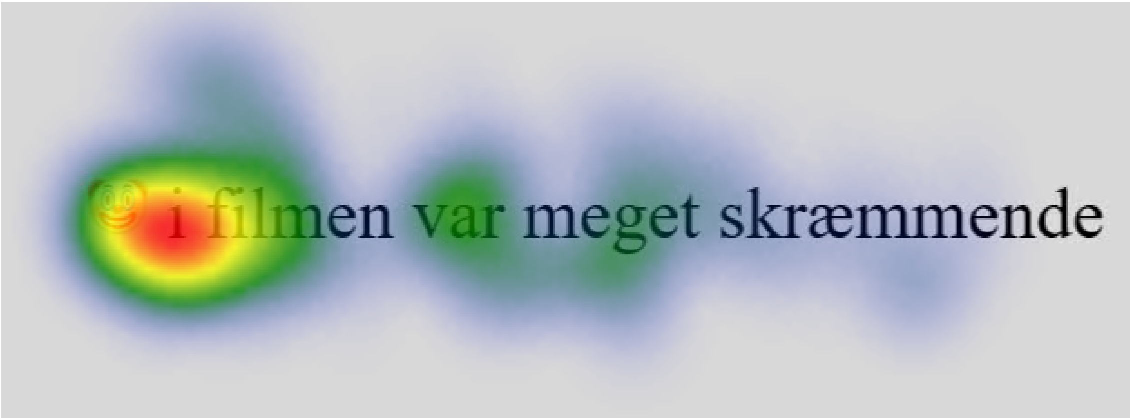

Figure C.52: Heat map F1x

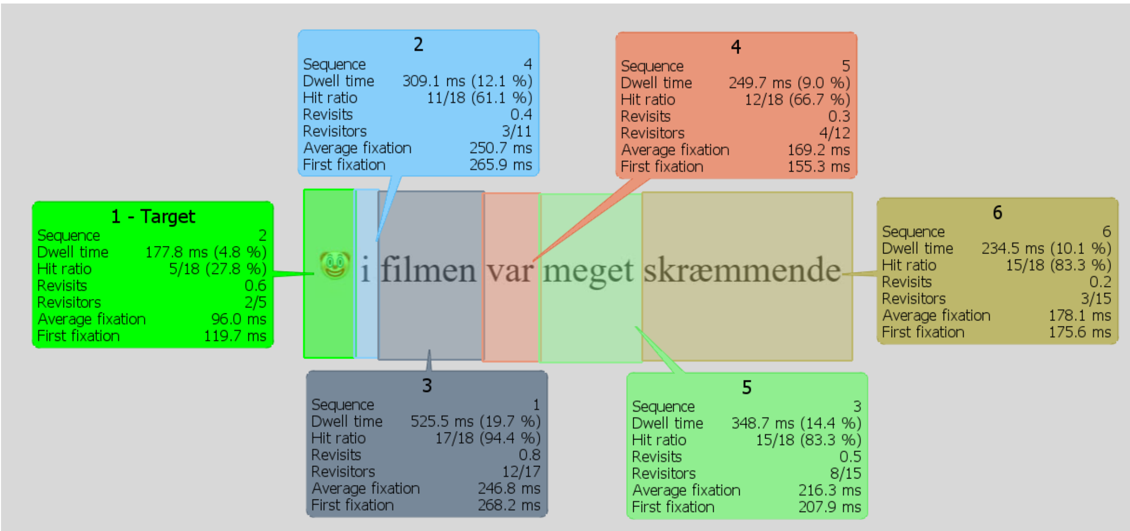

Figure C.53: Metrics F1x

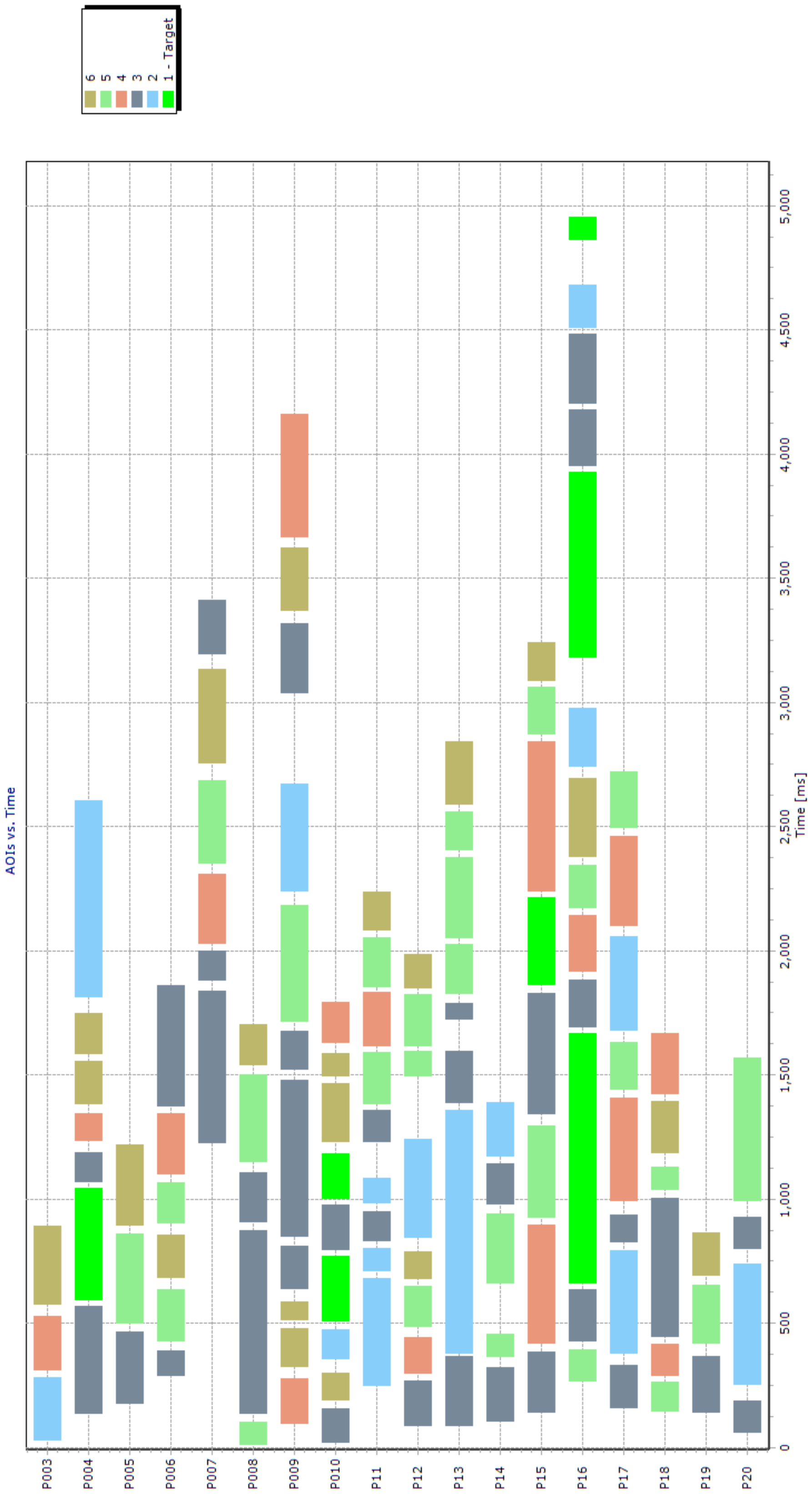

Figure C.54: AOI sequence chart for fixations F1x

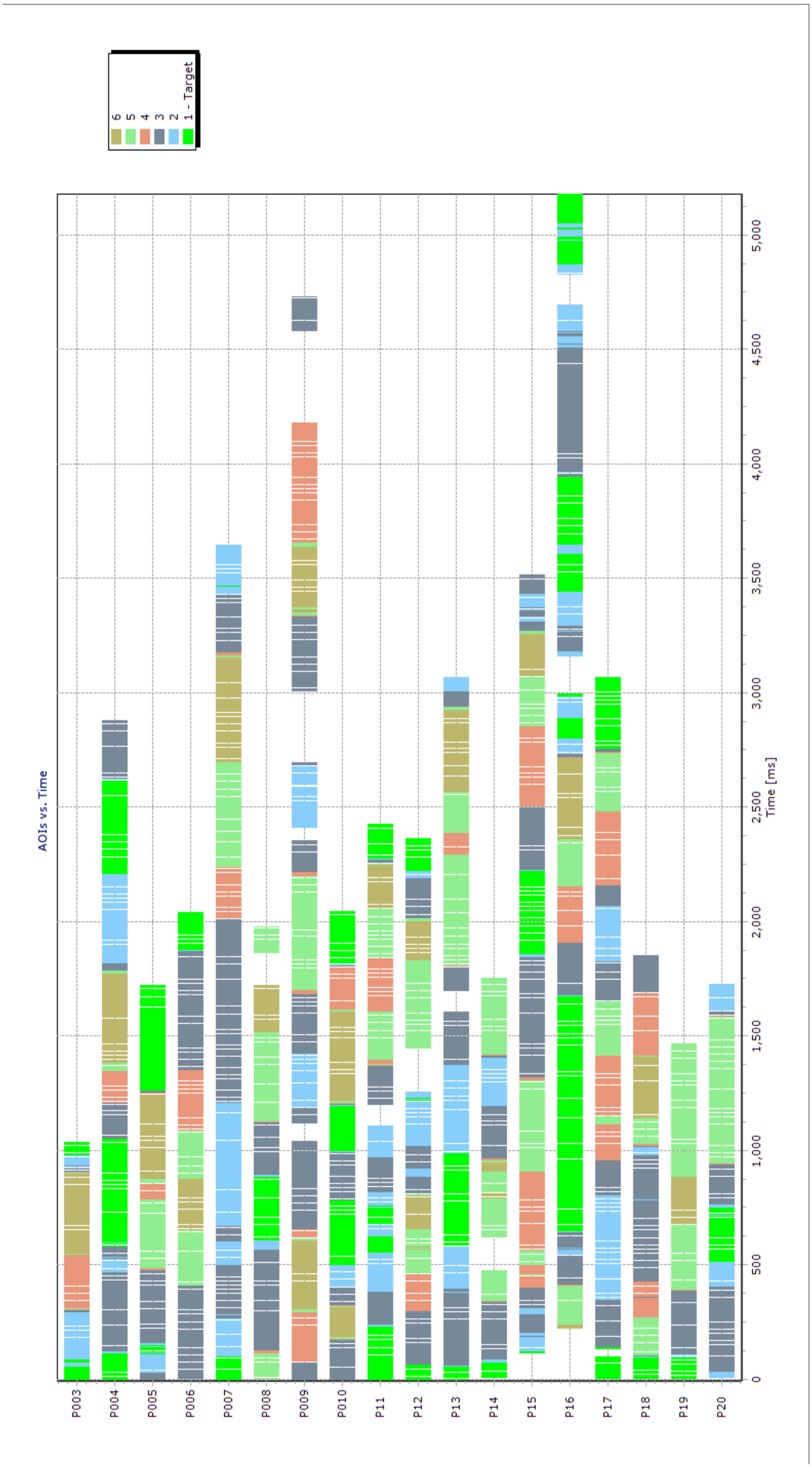

Figure C.55: AOI sequence chart for raw data F1x

C.2.3 F2

Figure C.56: Stimuli F2

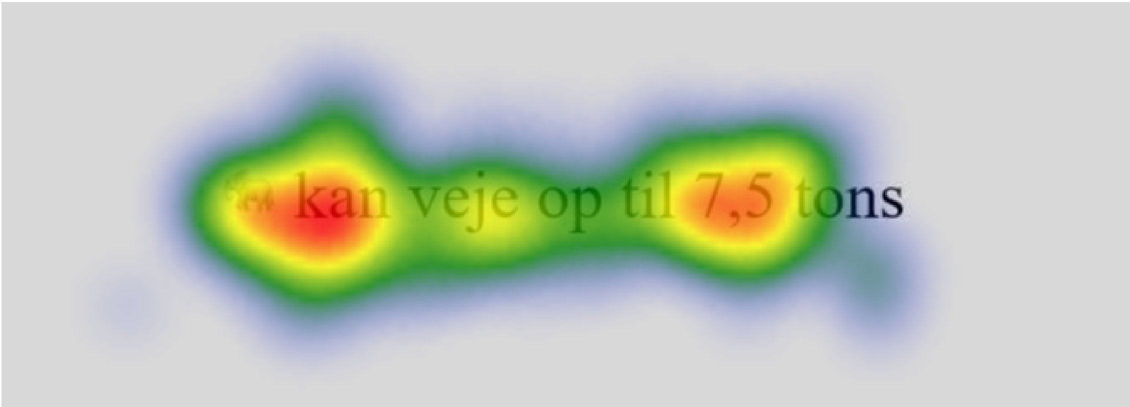

Figure C.57: Heat map F2

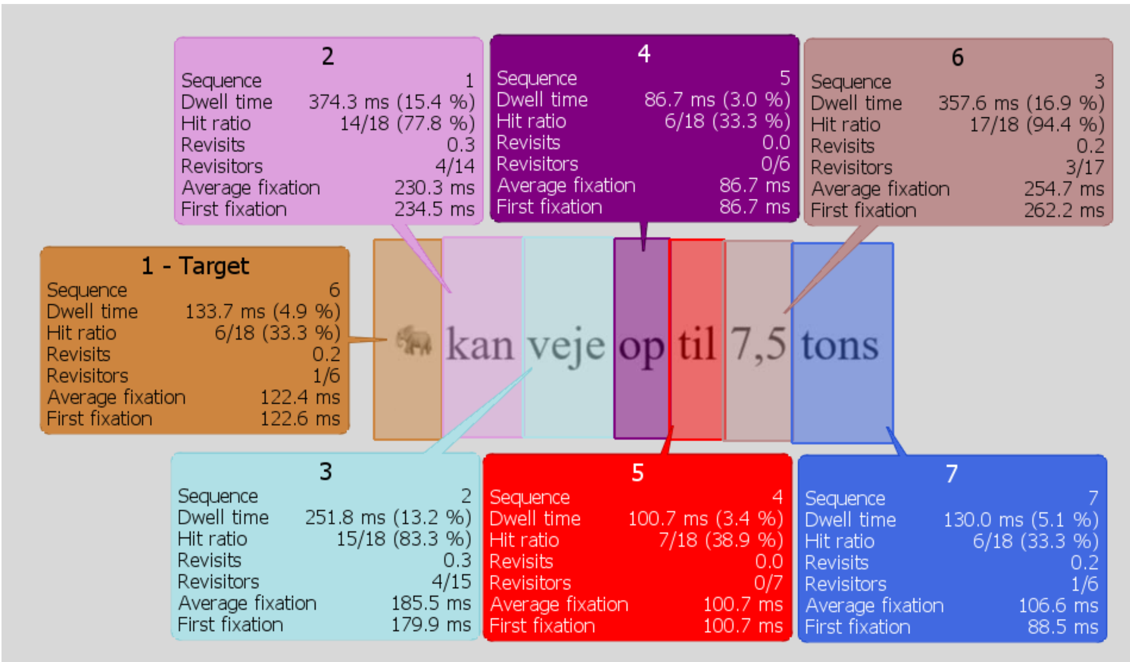

Figure C.58: Metrics F2

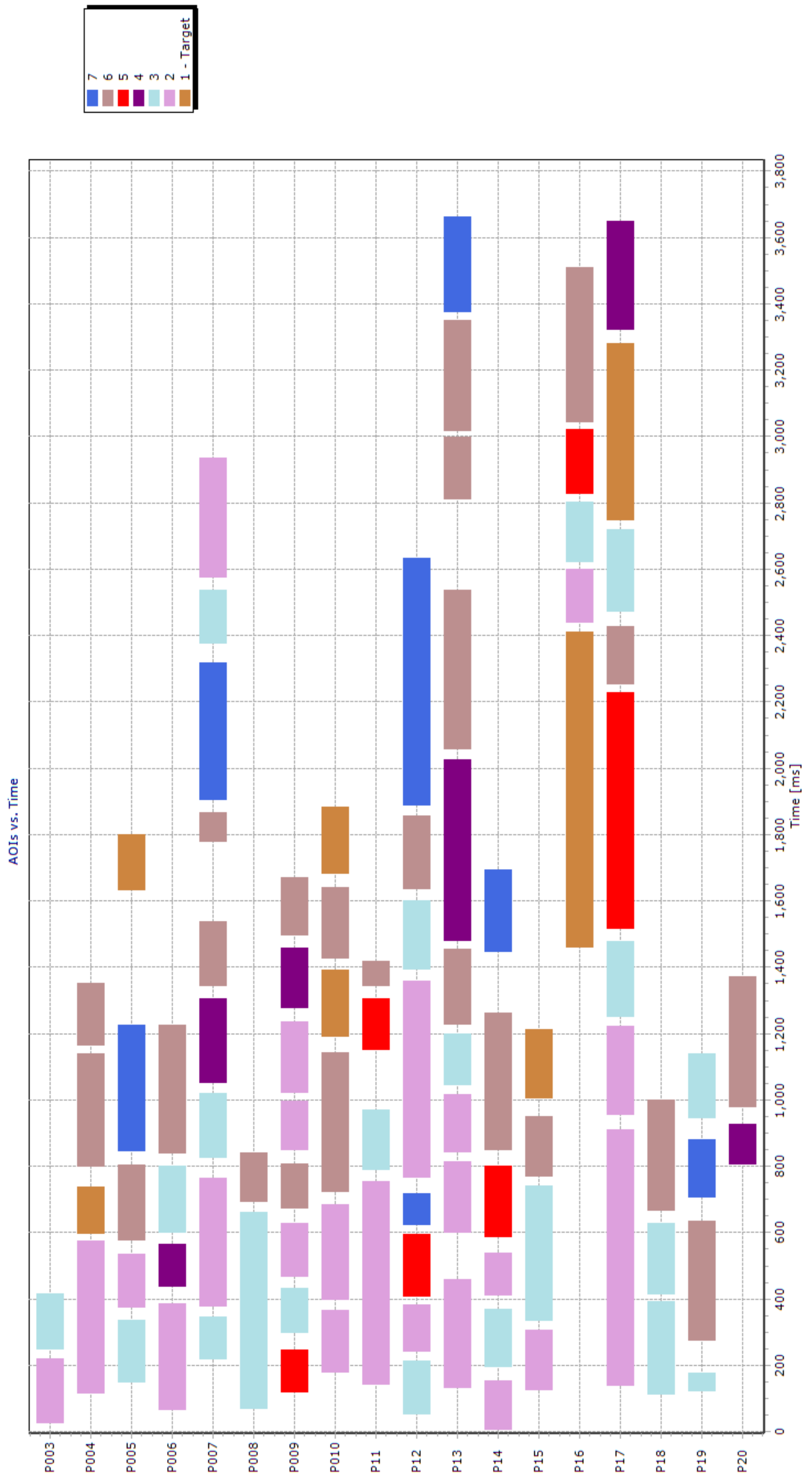

Figure C.59: AOI sequence chart for fixations F2

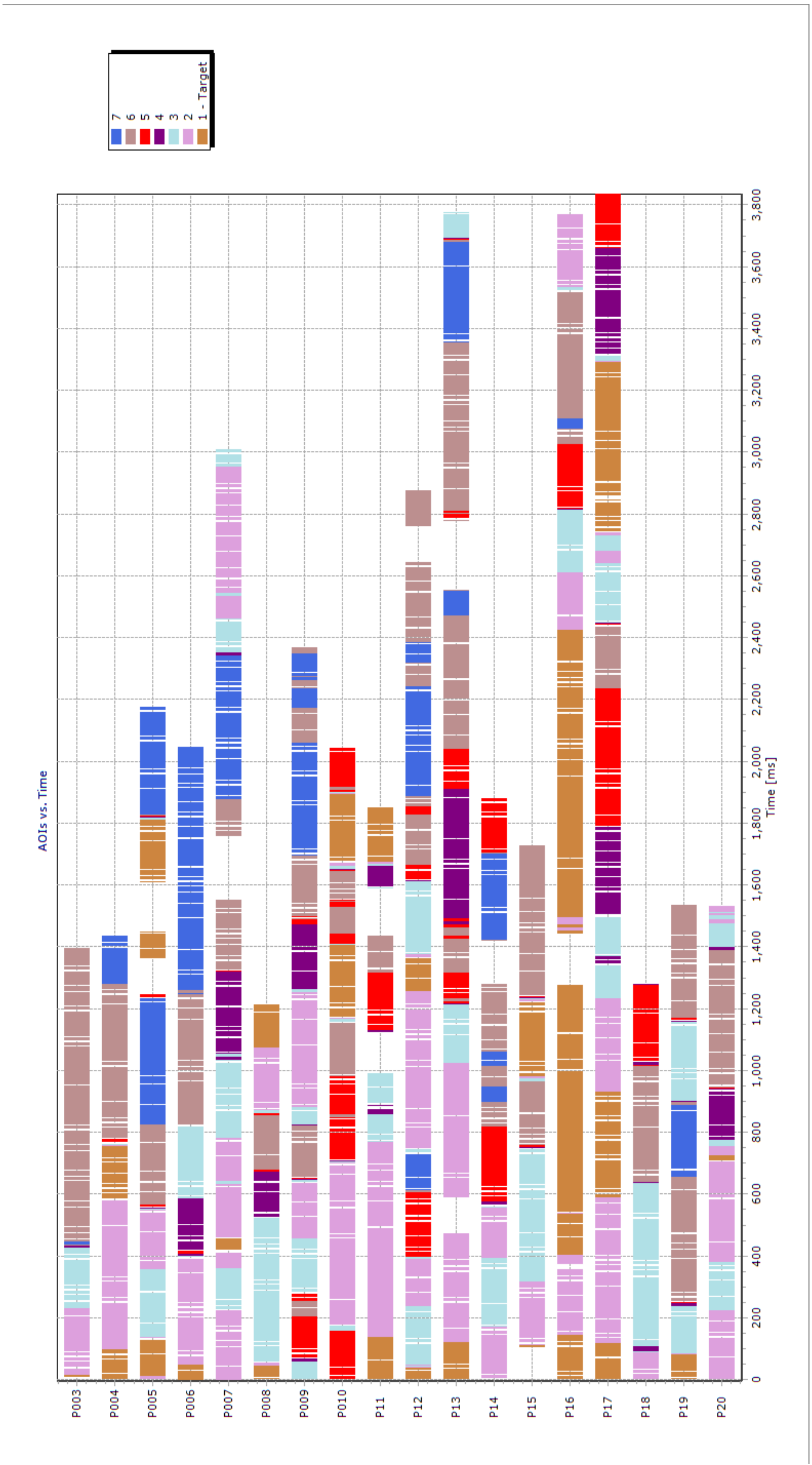

Figure C.60: AOI sequence chart for raw data F2

C.2.4 F2x

Figure C.61: Stimuli F2x

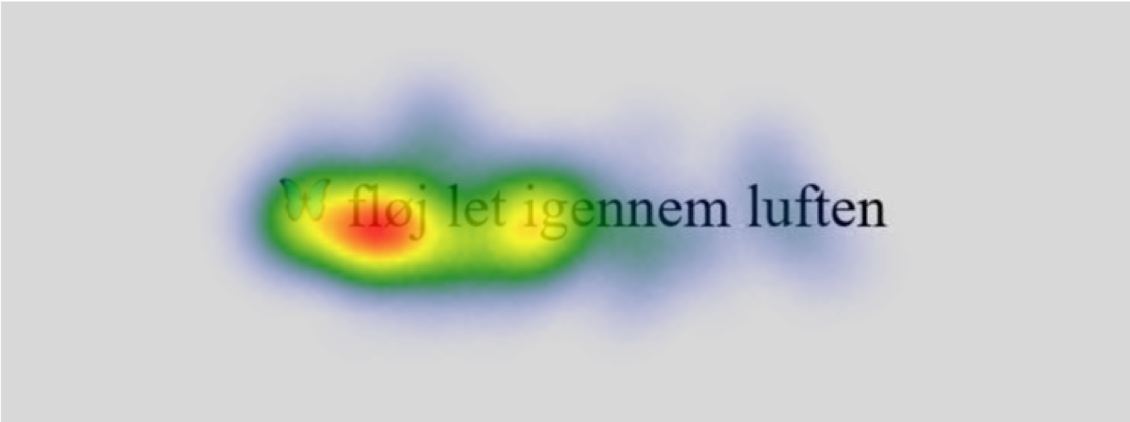

Figure C.62: Heat map F2x

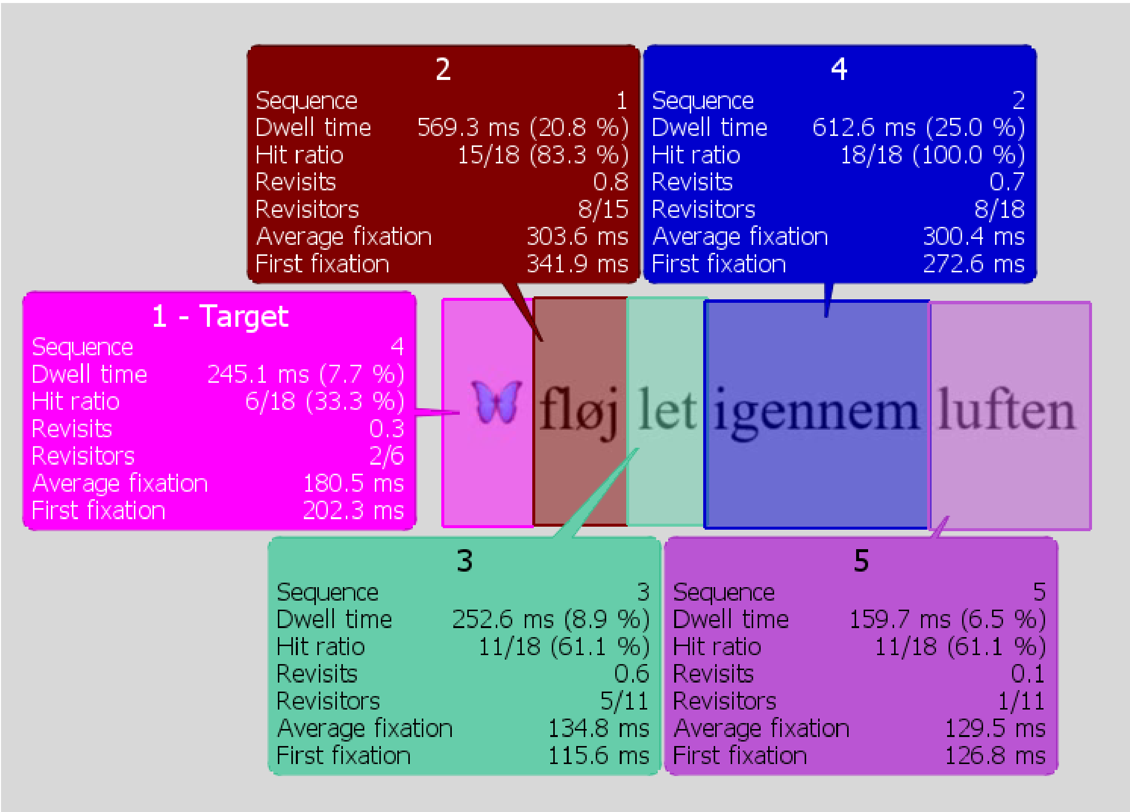

Figure C.63: Metrics F2x

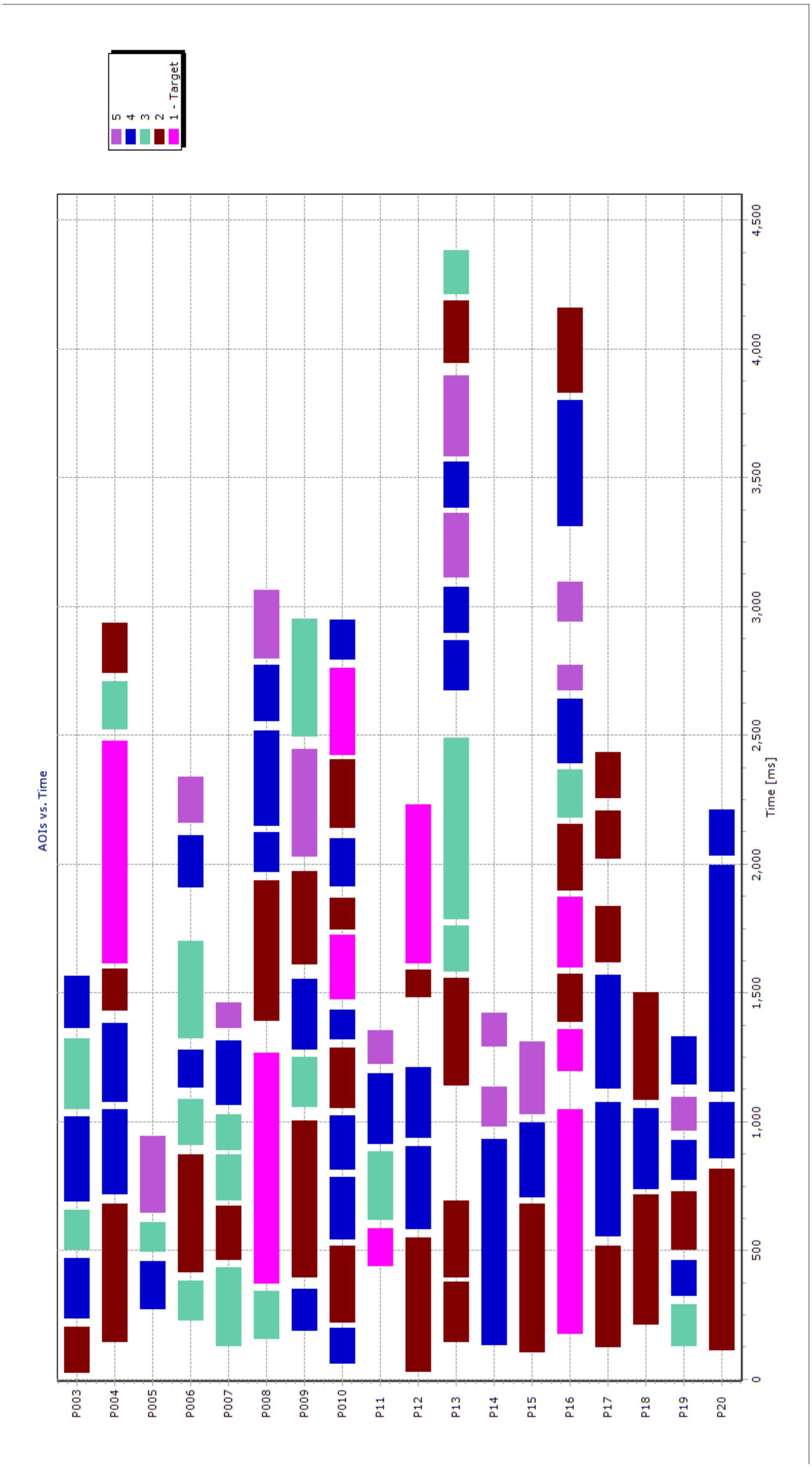

Figure C.64: AOI sequence chart for fixations F2x

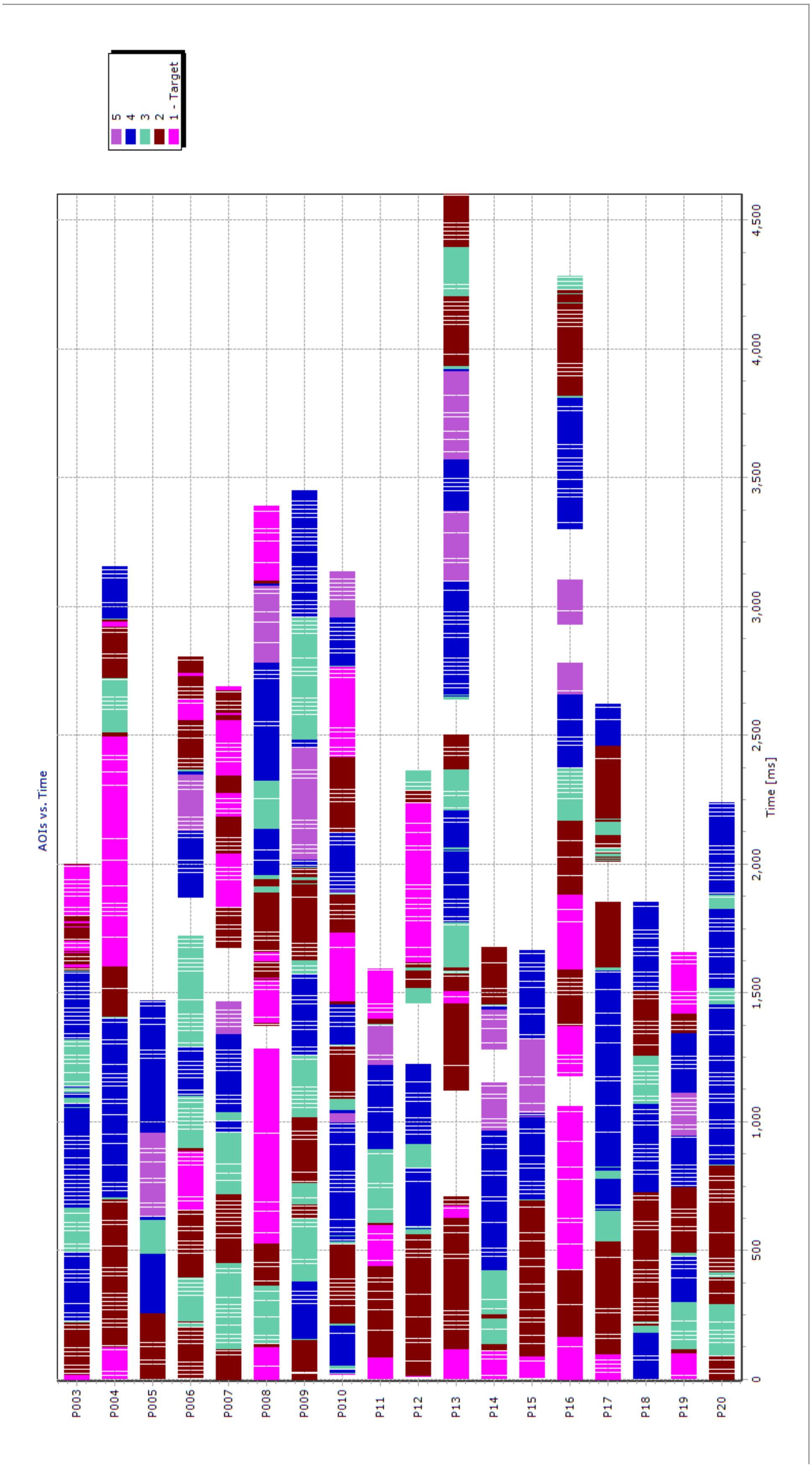

Figure C.65: AOI sequence chart for raw data F2x

C.2.5 F3

Figure C.66: Stimuli F3

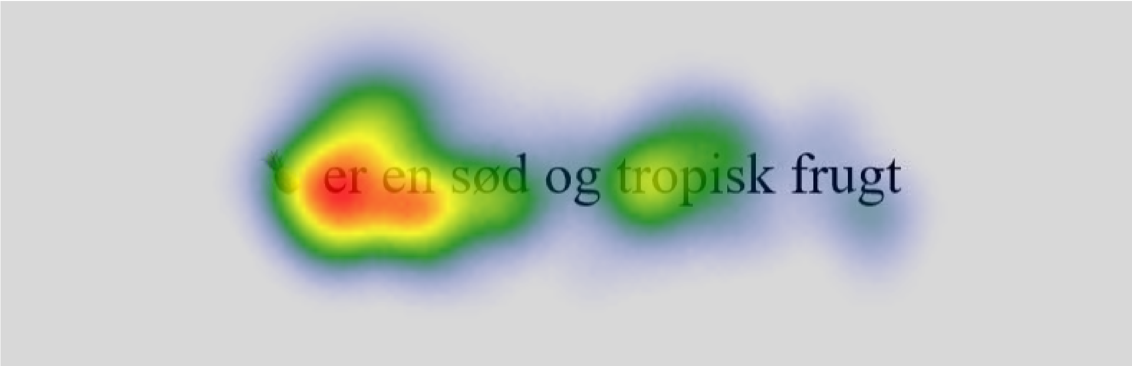

Figure C.67: Heat map F3

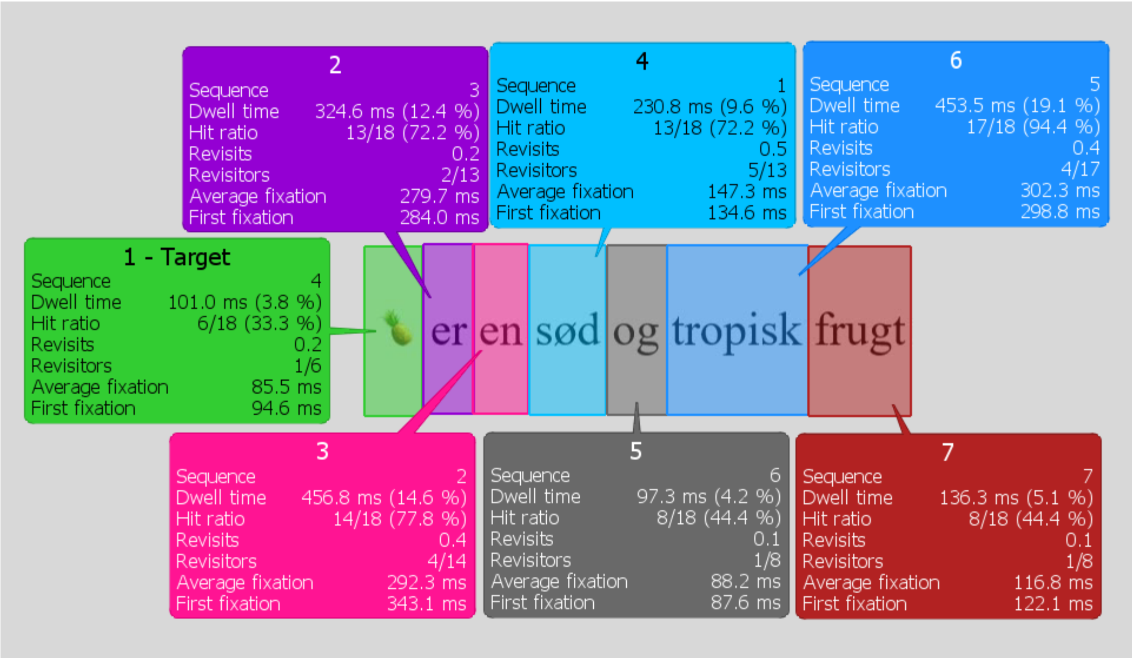

Figure C.68: Metrics F3

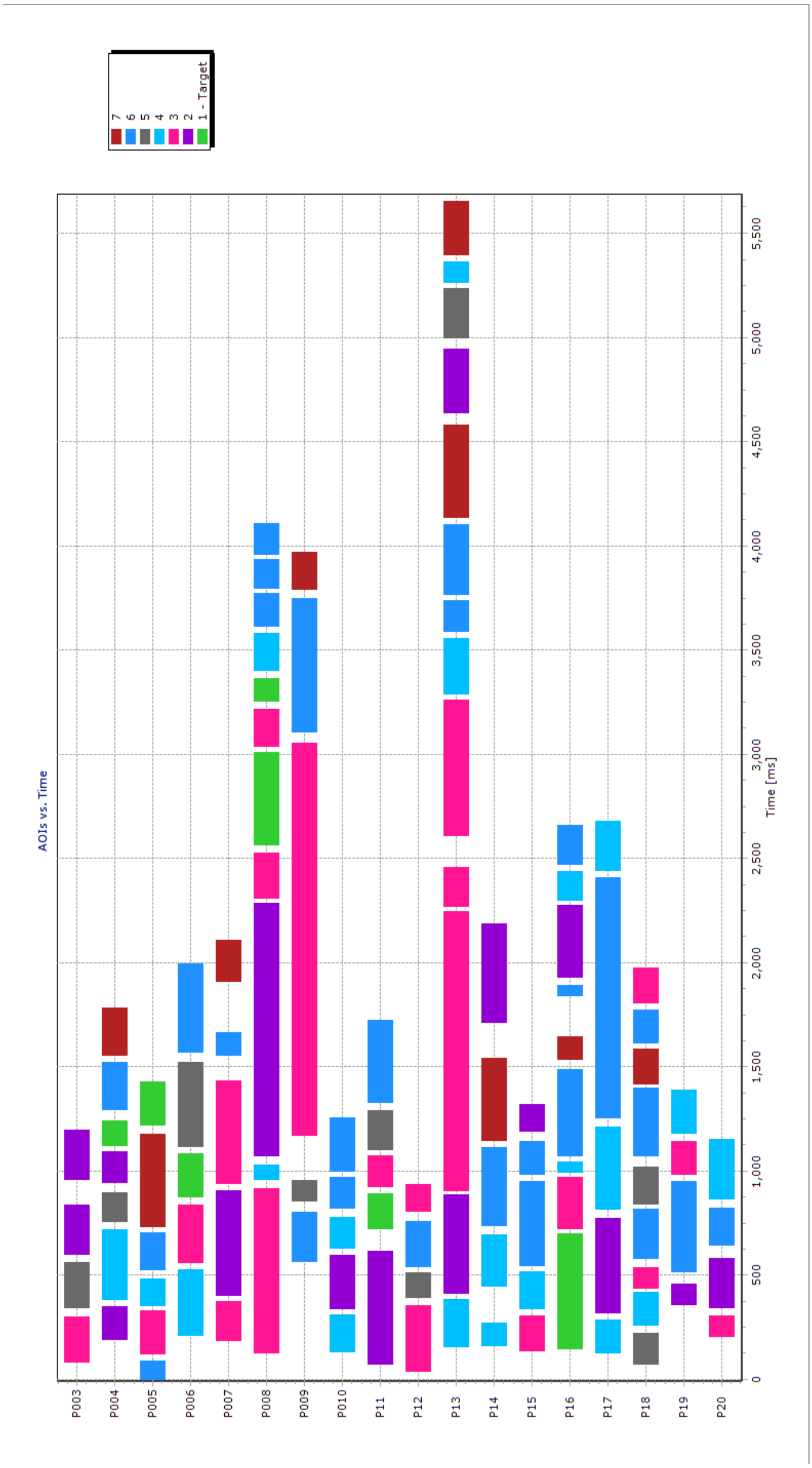

Figure C.69: AOI sequence chart for fixations F3

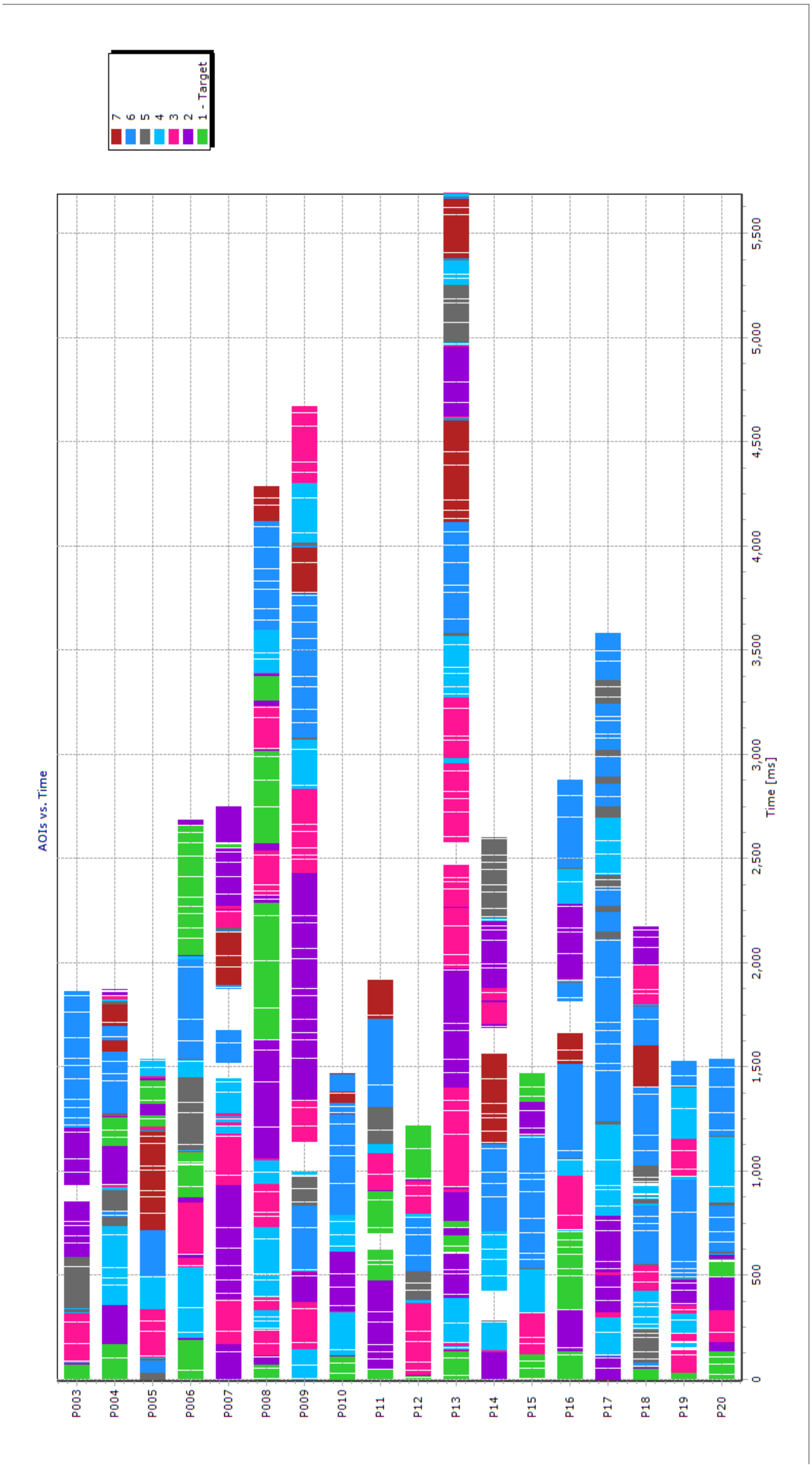

Figure C.70: AOI sequence chart for raw data F3

C.2.6 F3x

Figure C.71: Stimuli F3x

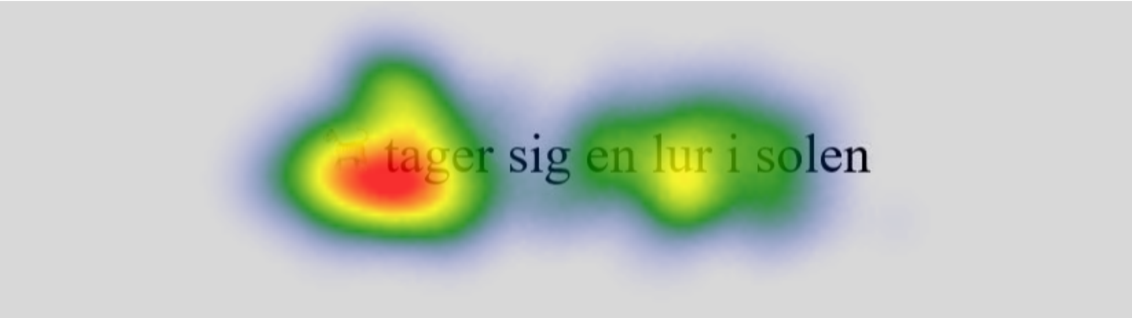

Figure C.72: Heat map F3x

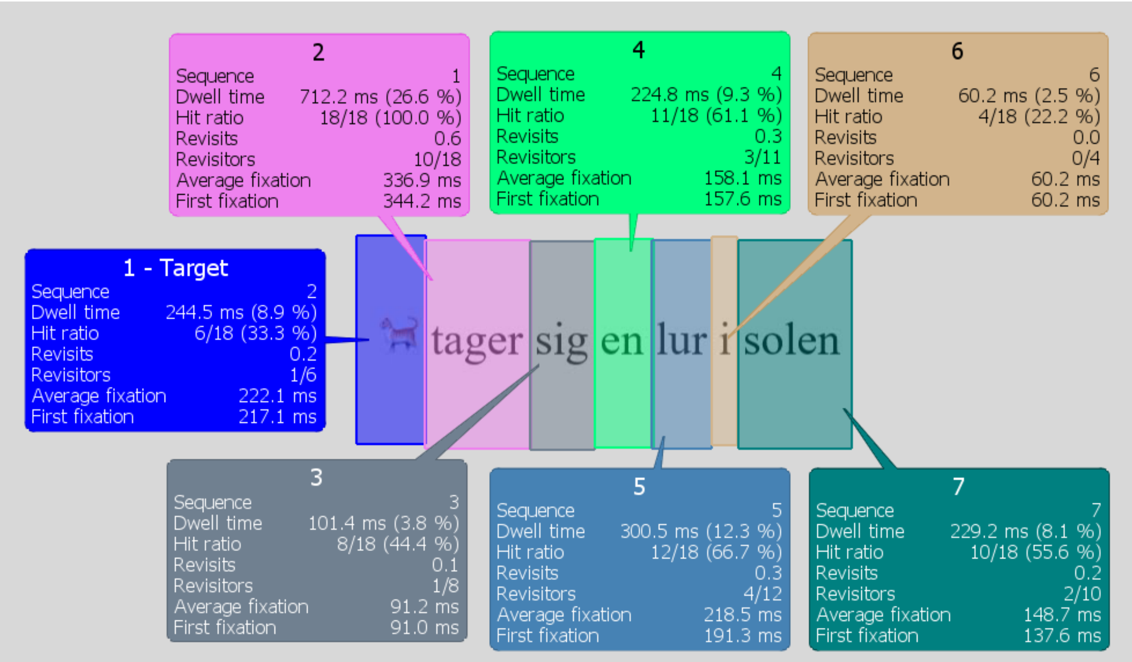

Figure C.73: Metrics F3x

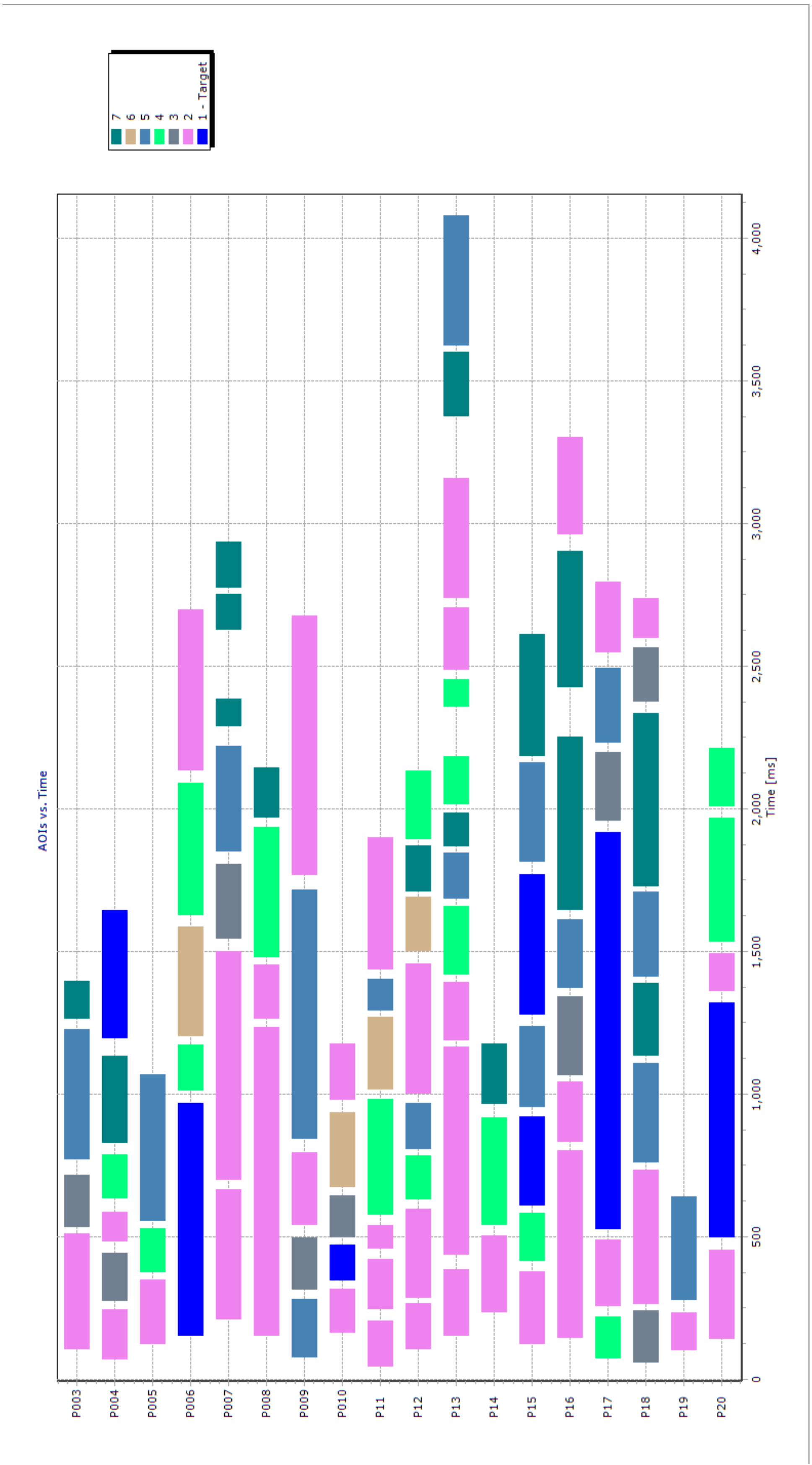

Figure C.74: AOI sequence chart for fixations F3x

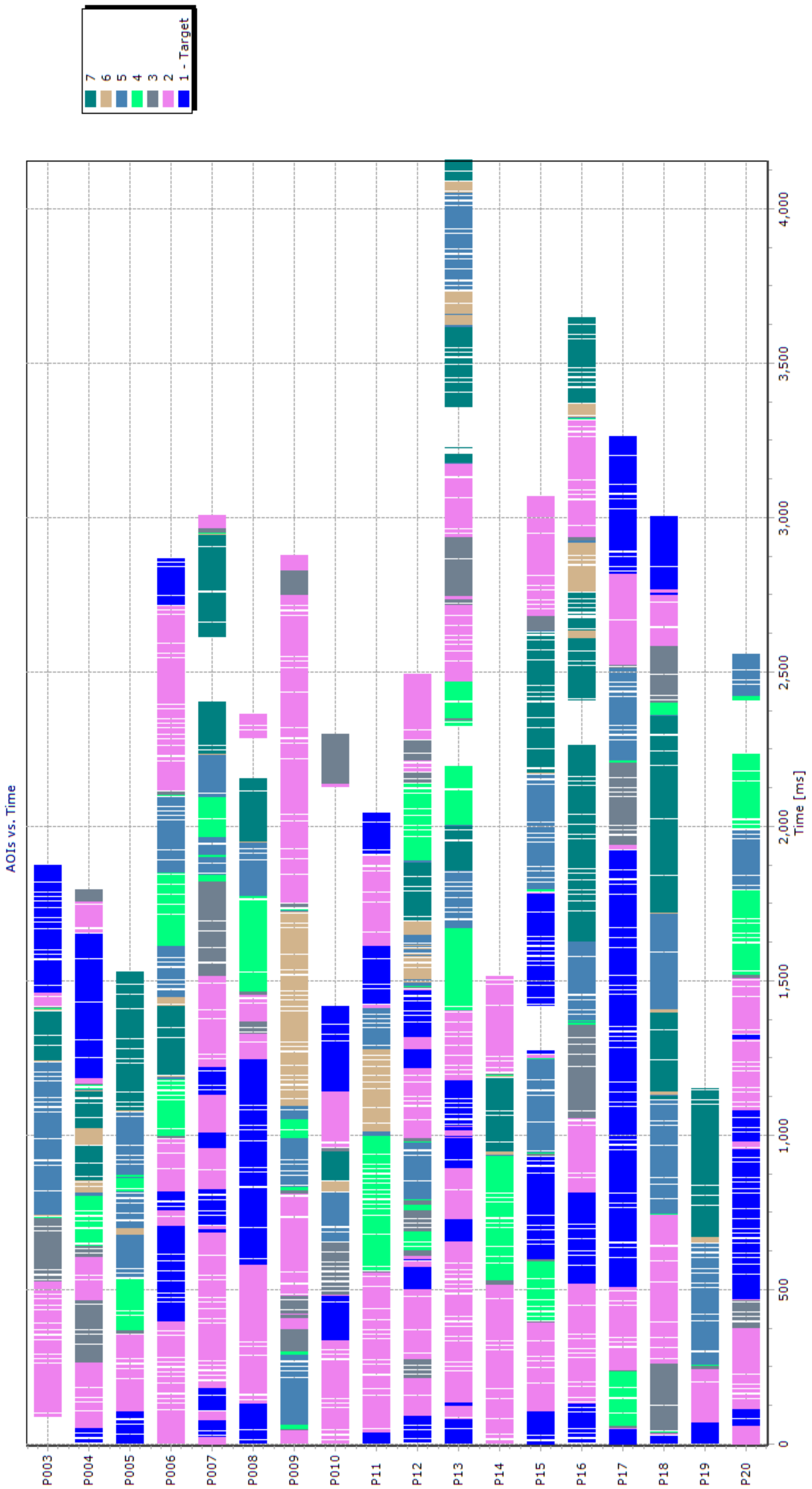

Figure C.75: AOI sequence chart for raw data F3x

C.2.7 F4

Figure C.76: Stimuli F4

Figure C.77: Heat map F4

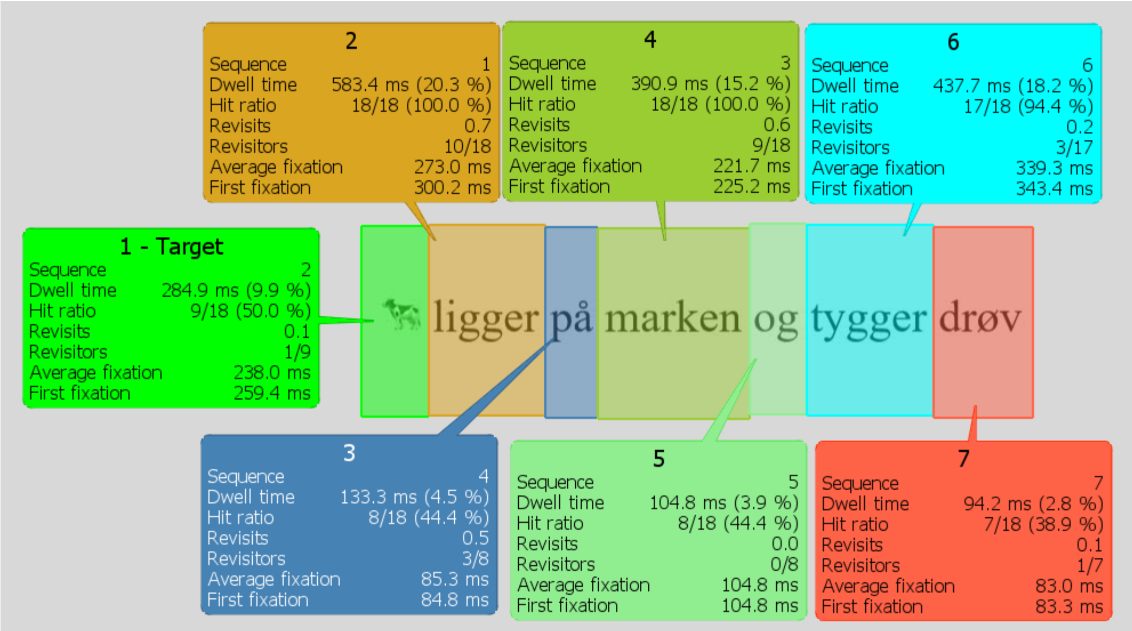

Figure C.78: Metrics F4

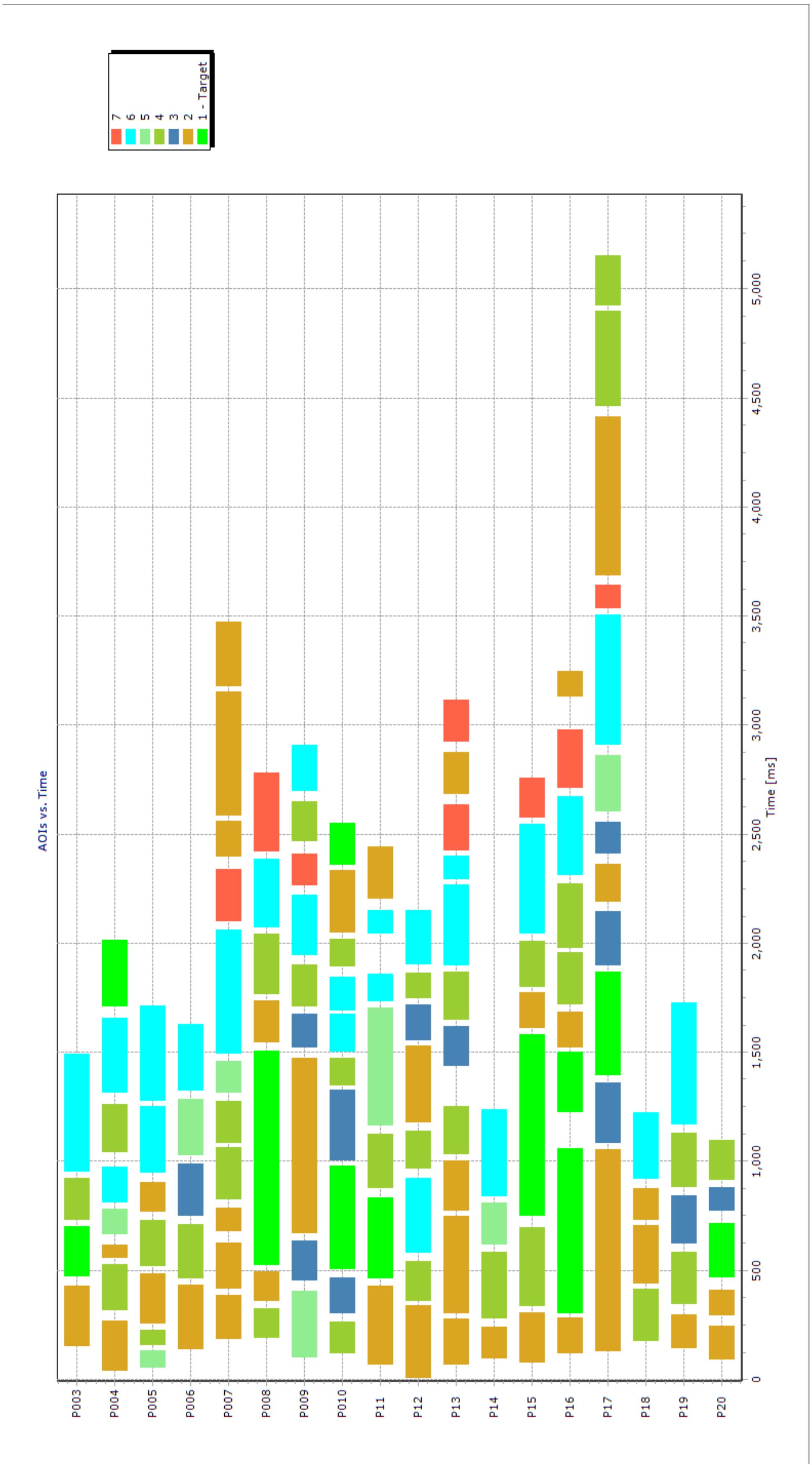

Figure C.79: AOI sequence chart for fixations F4

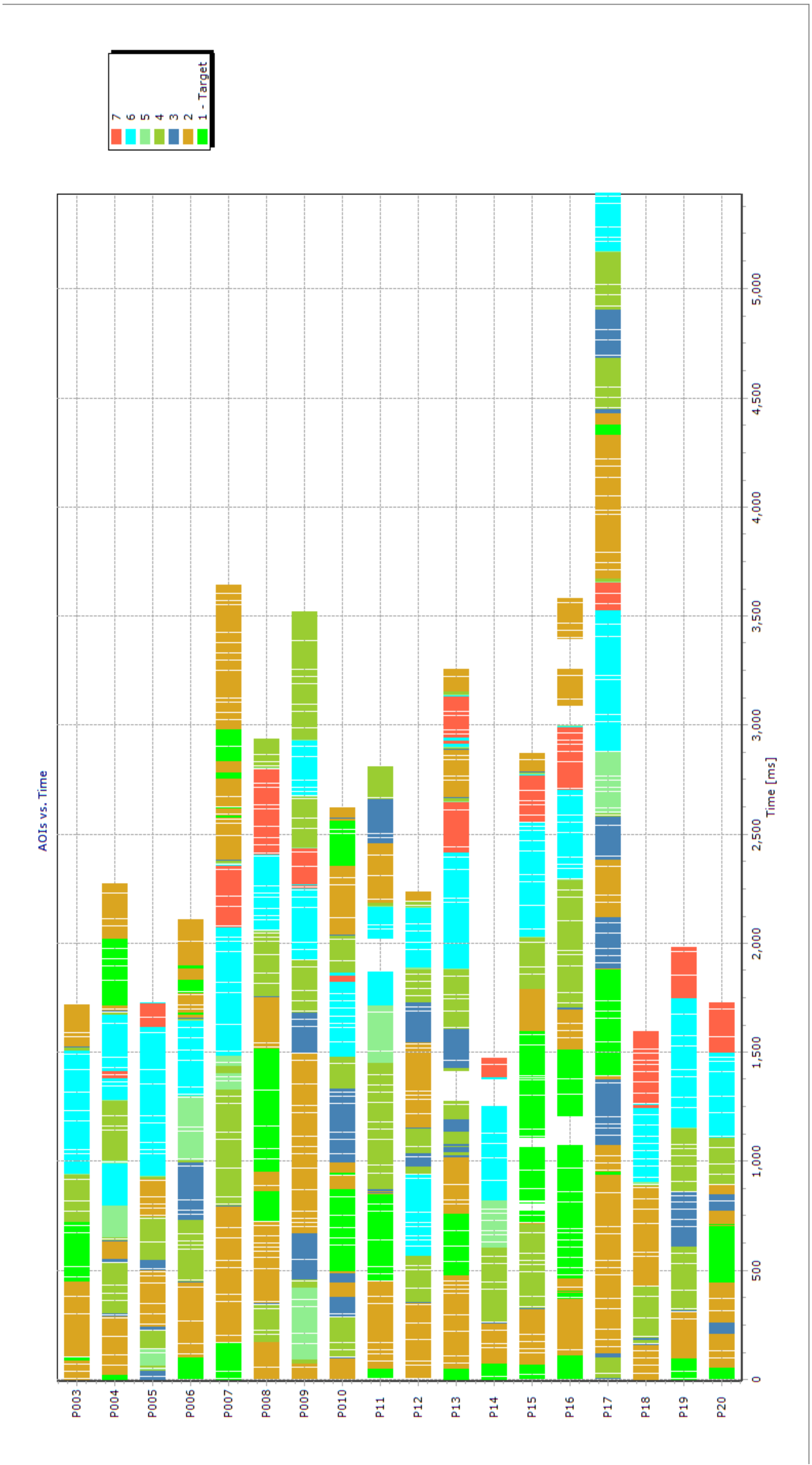

Figure C.80: AOI sequence chart for raw data F4

C.2.8 F5

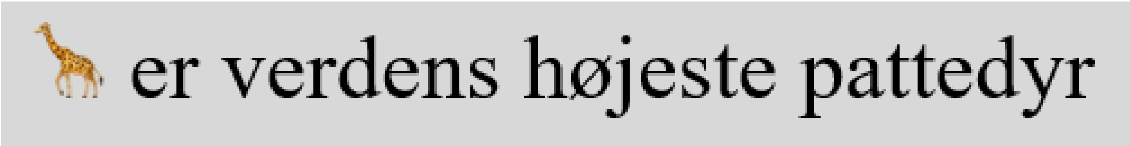

Figure C.81: Stimuli F5

Figure C.82: Heat map F5

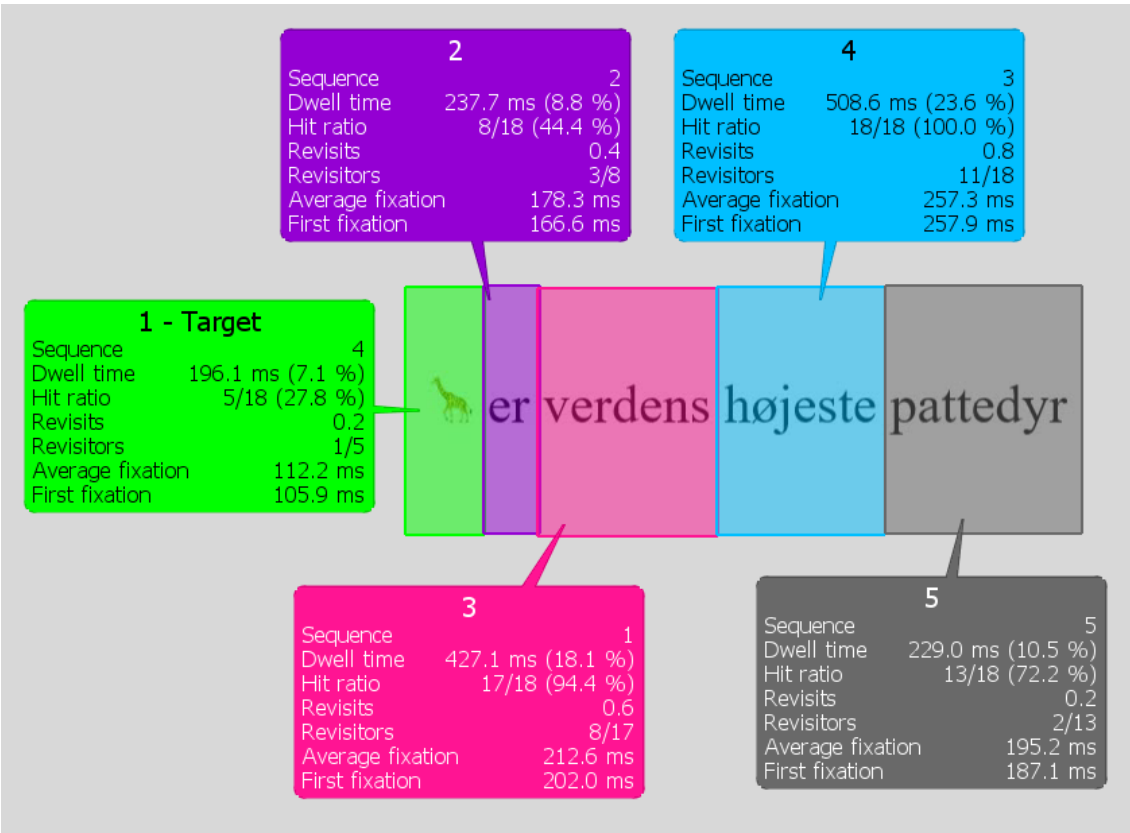

Figure C.83: Metrics F5

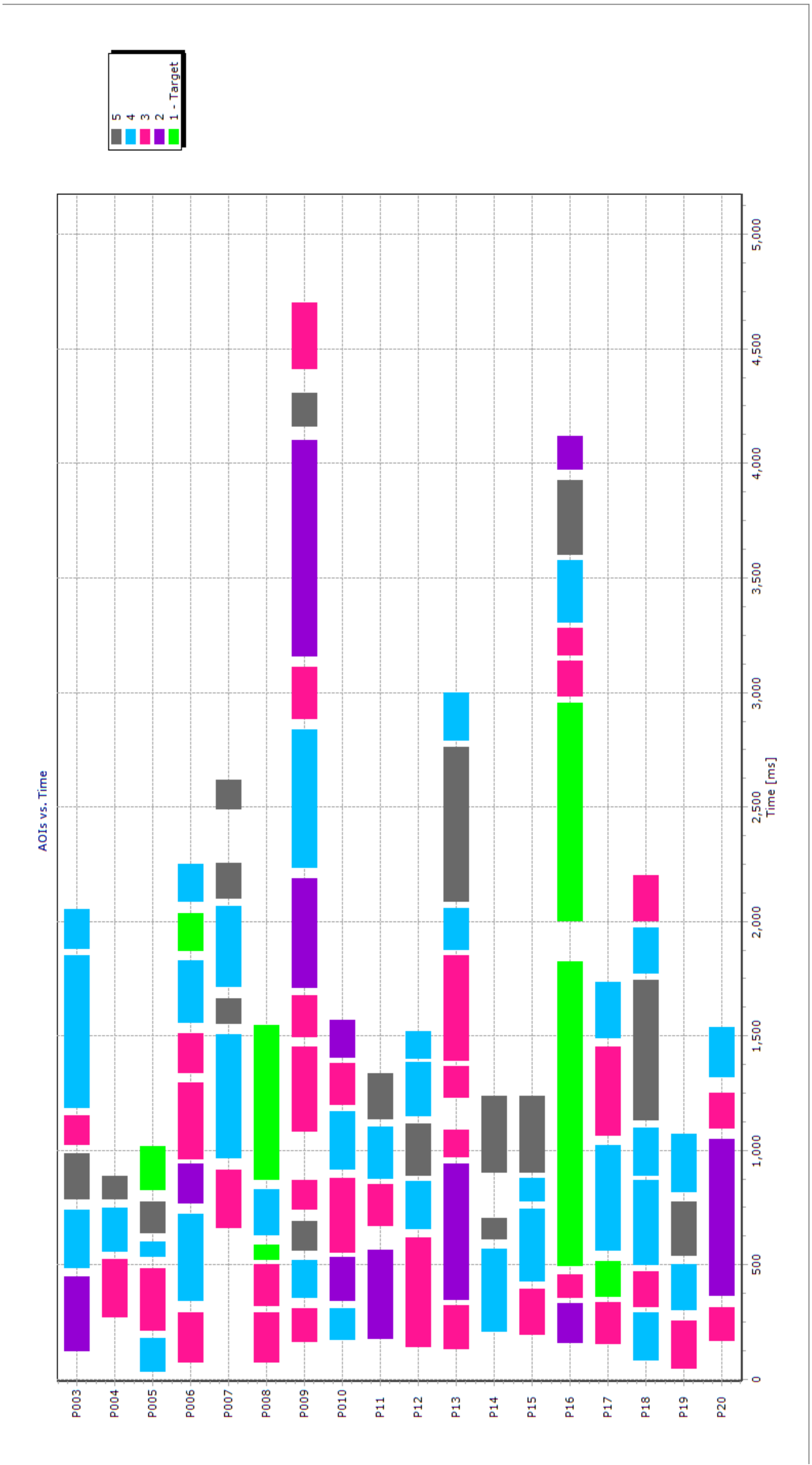

Figure C.84: AOI sequence chart for fixations F5

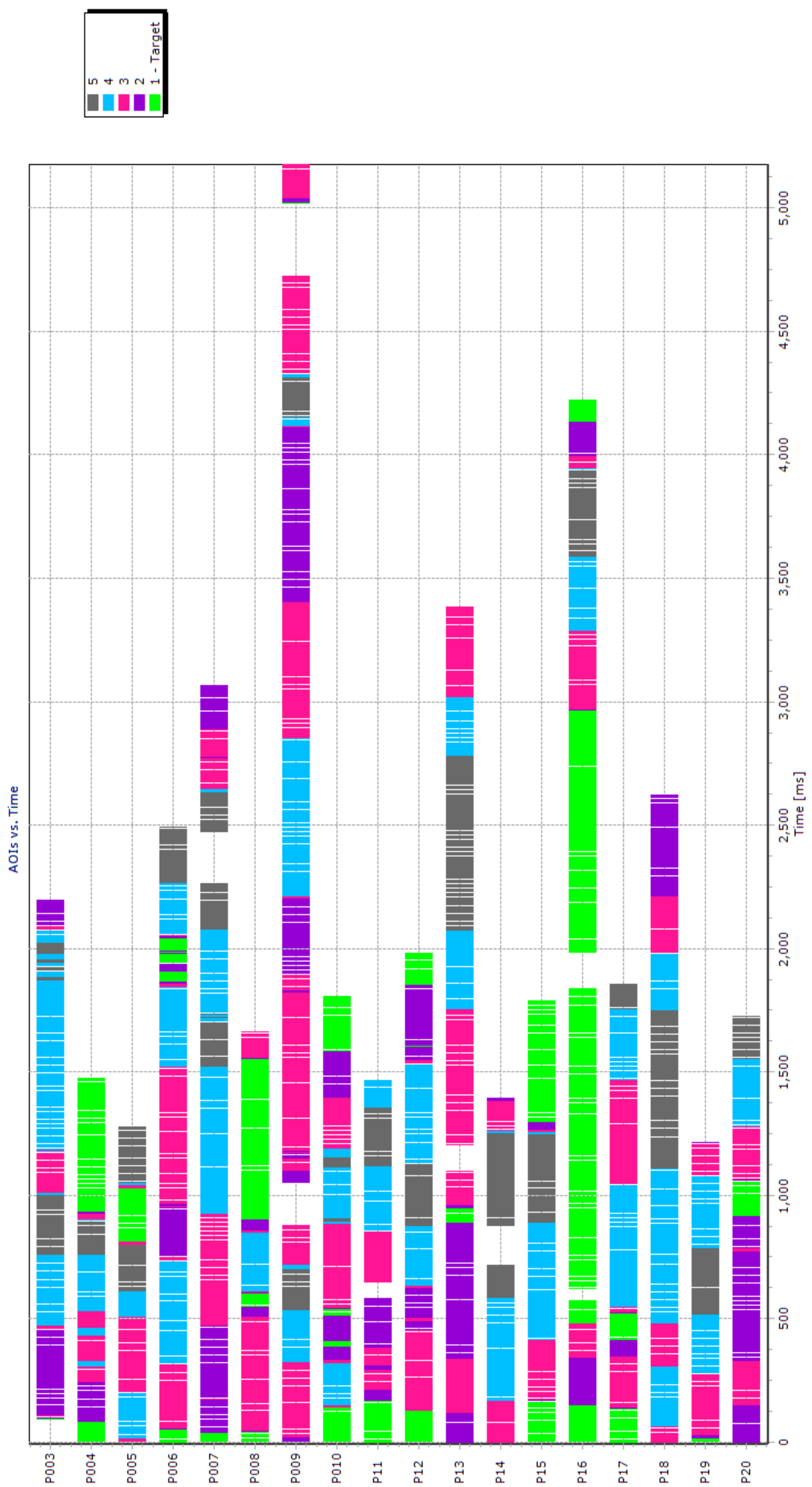

Figure C.85: AOI sequence chart for raw data F5

C.2.9 F6

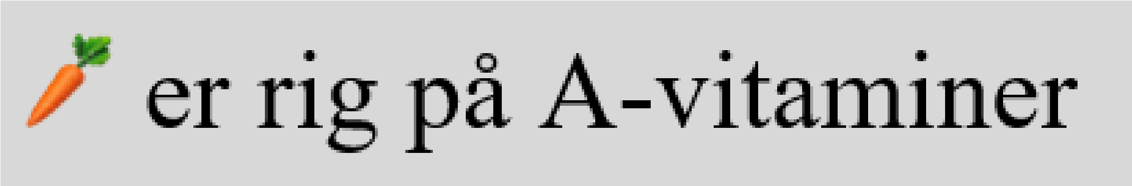

Figure C.86: Stimuli F6

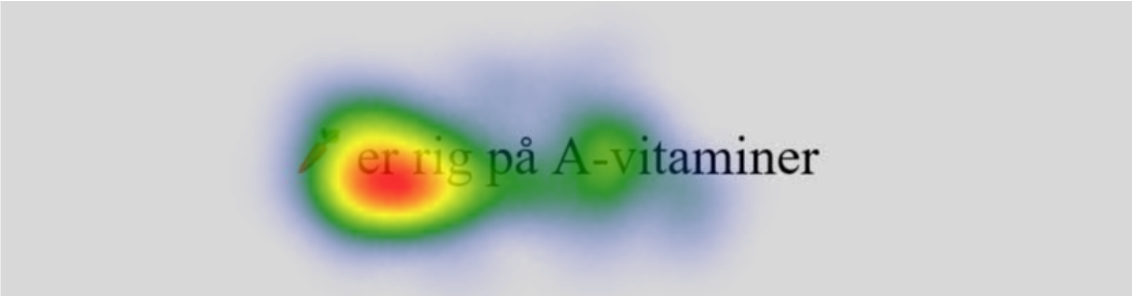

Figure C.87: Heat map F6

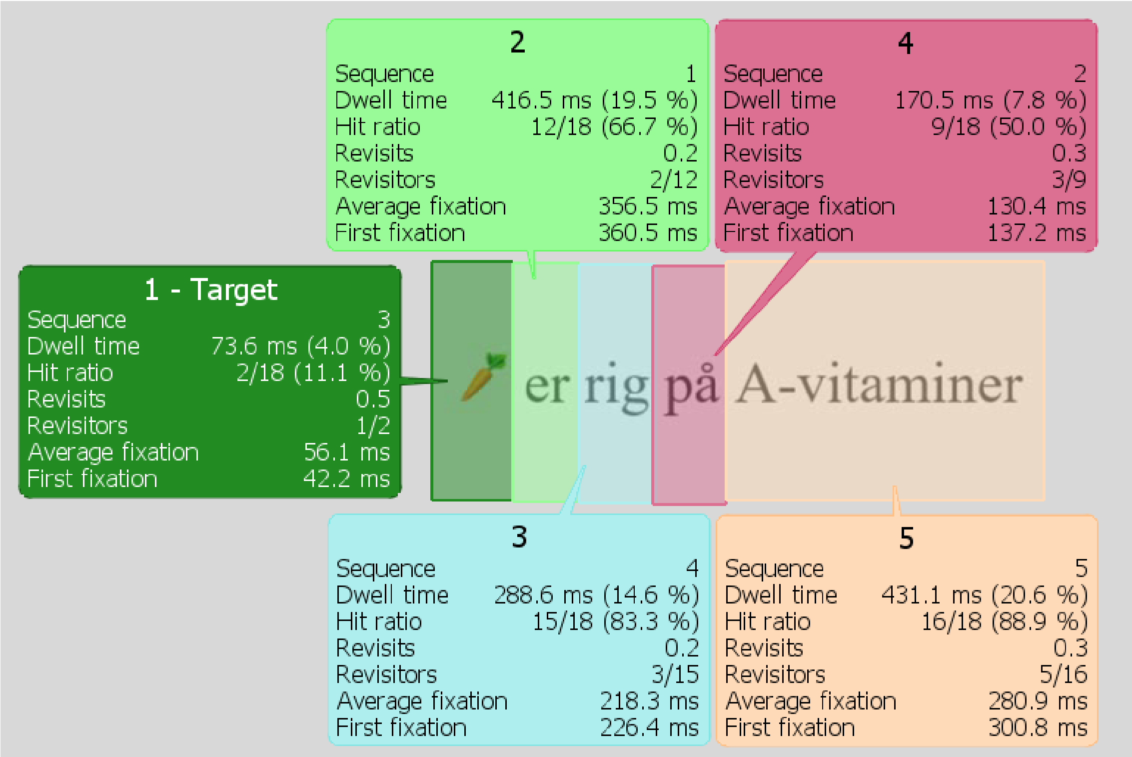

Figure C.88: Metrics F6

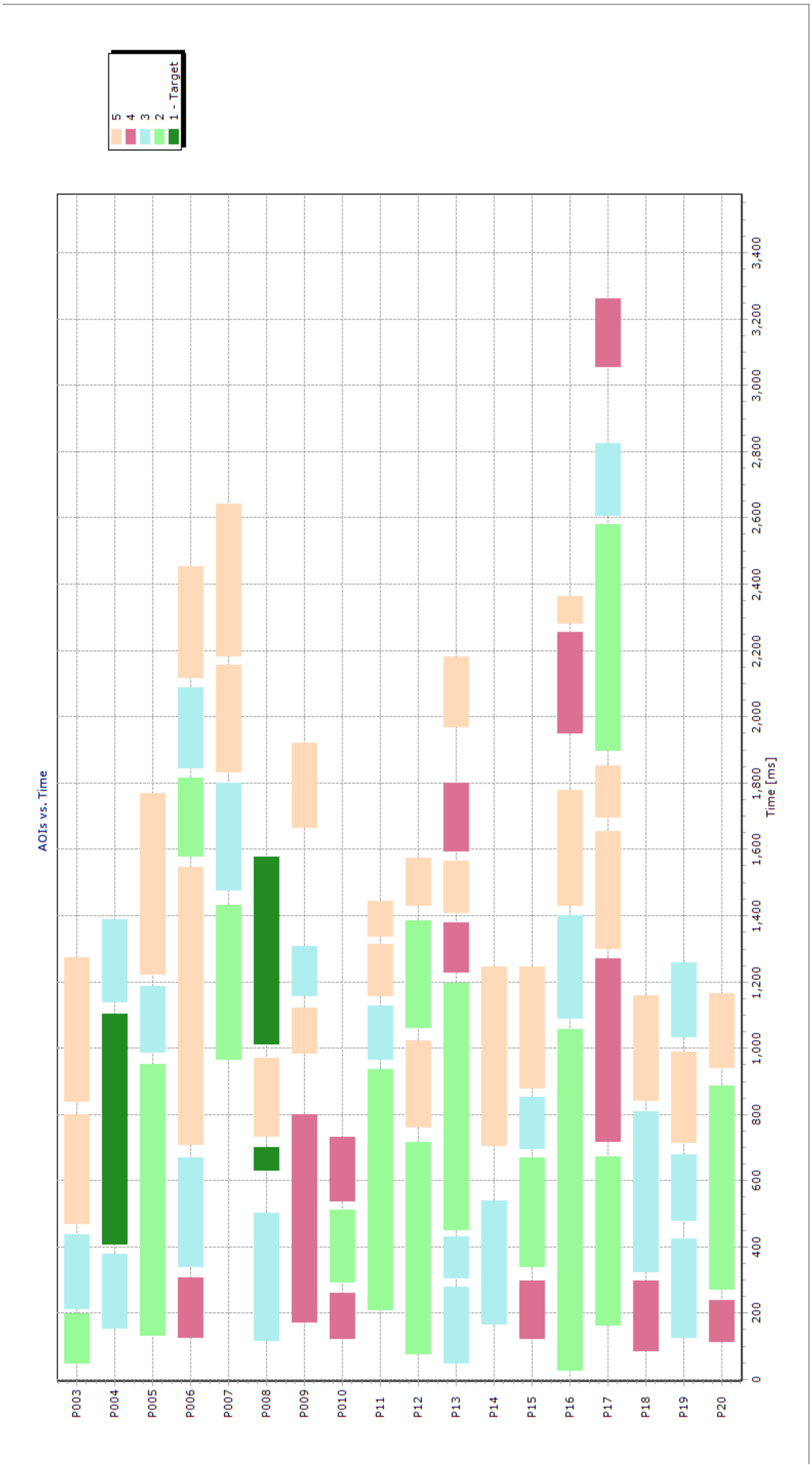

Figure C.89: AOI sequence chart for fixations F6

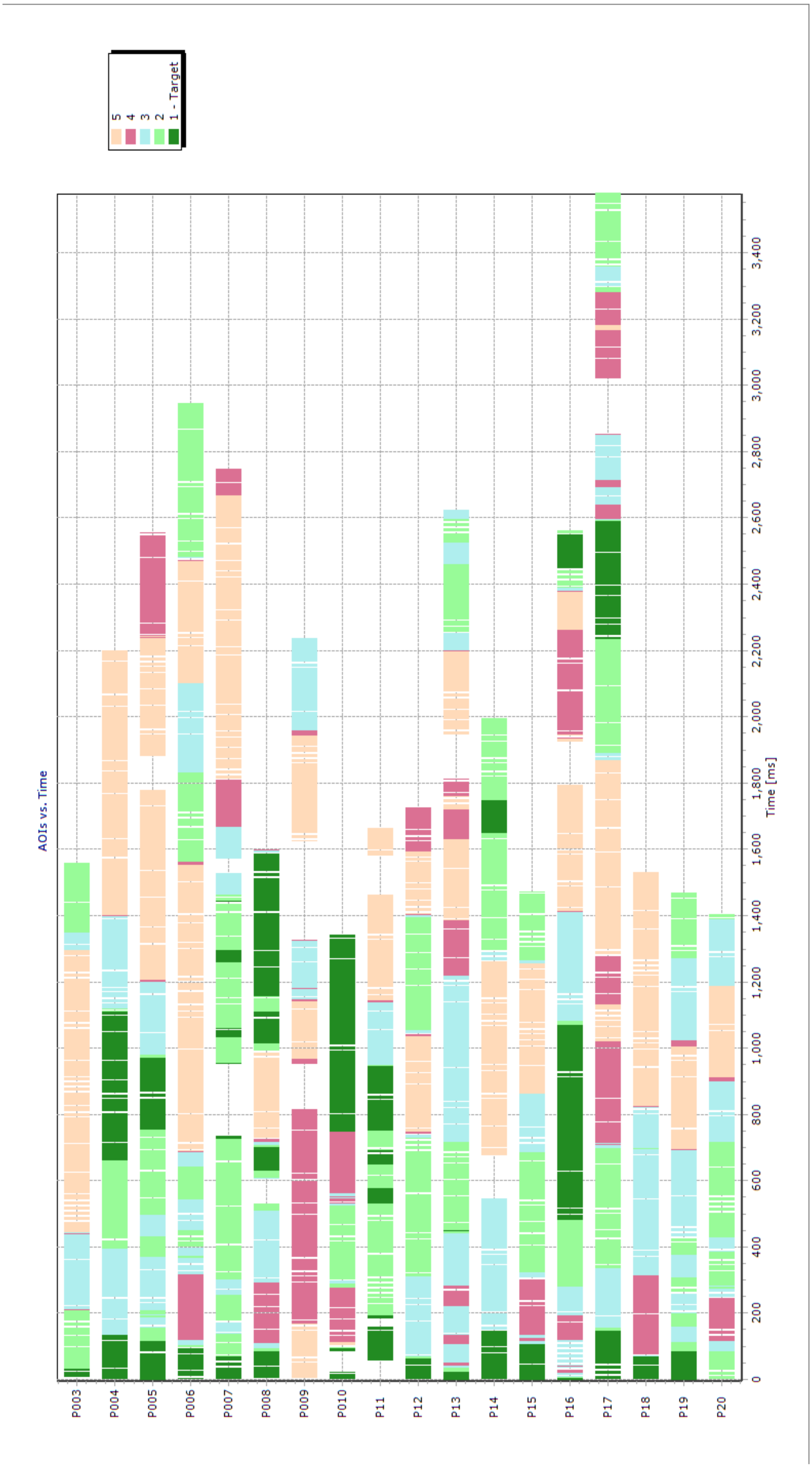

Figure C.90: AOI sequence chart for raw data F6

C.3 Middle targets(M)

C.3.1 M1

Figure C.91: Stimuli M1

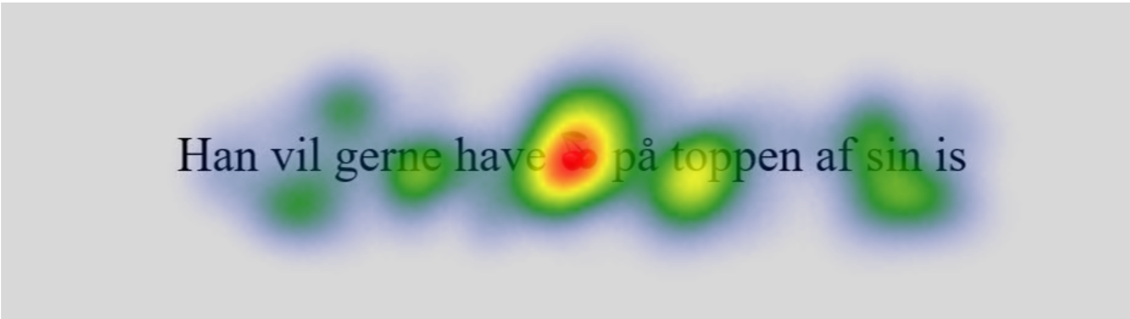

Figure C.92: Heat map M1

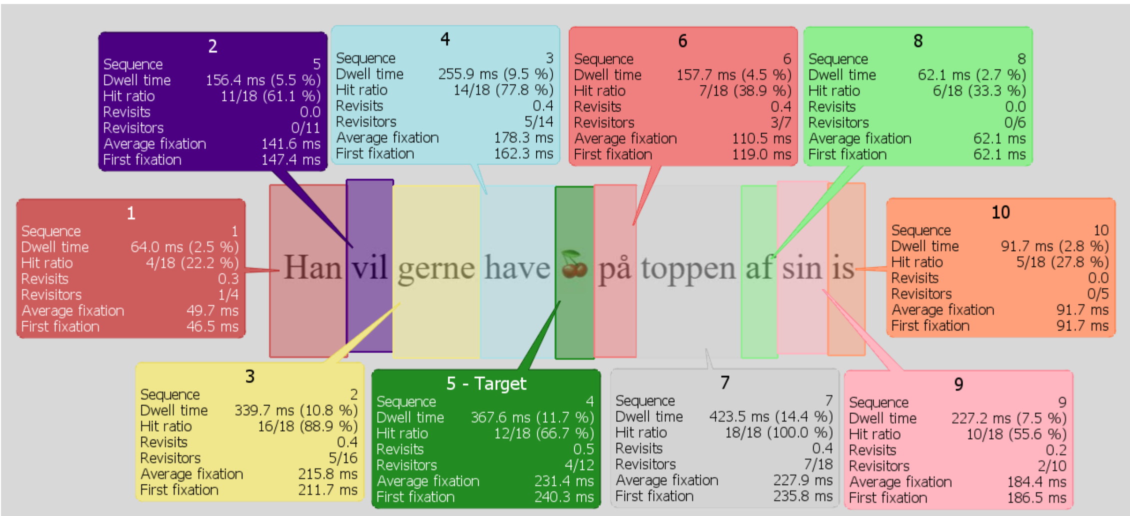

Figure C.93: Metrics M1

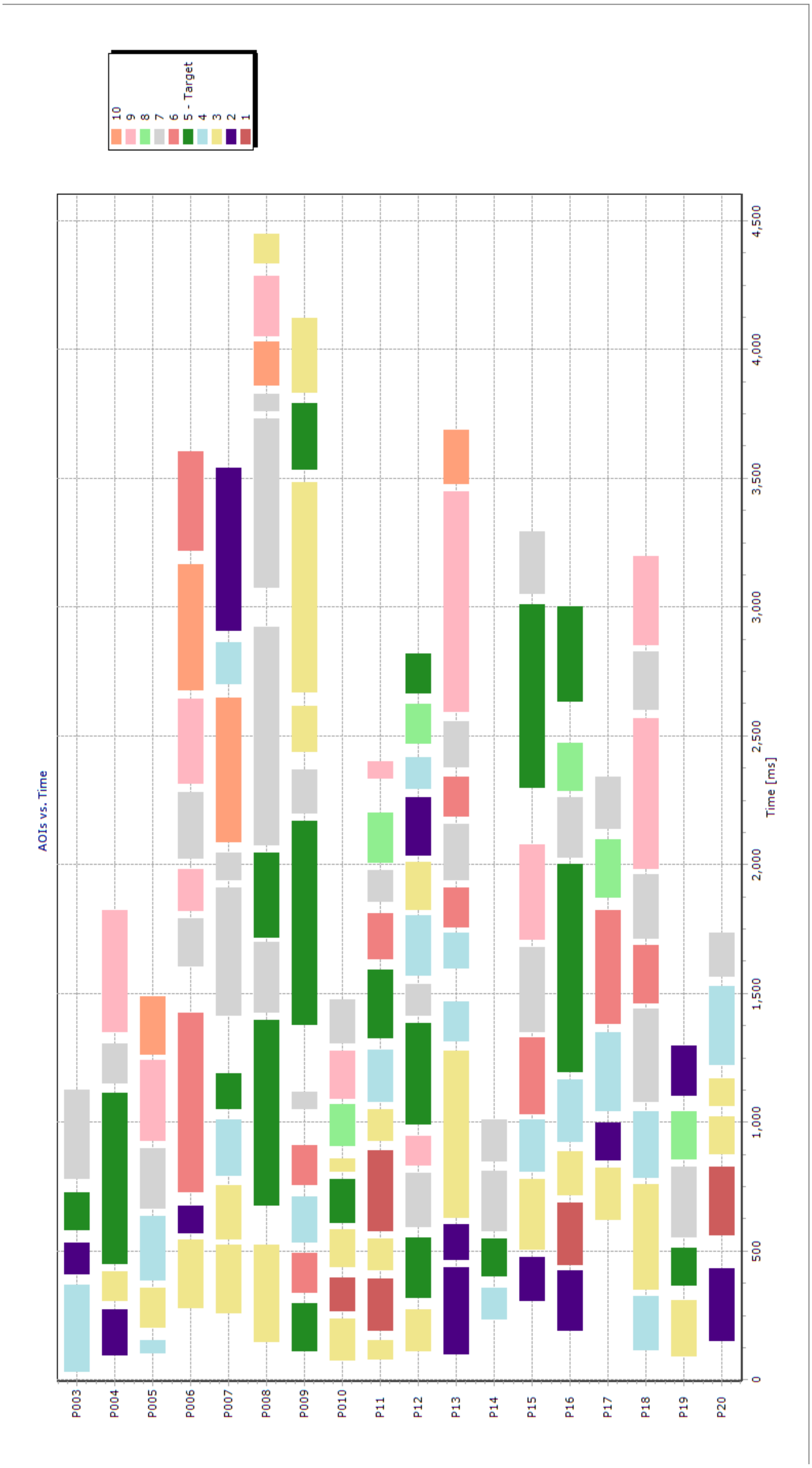

Figure C.94: AOI sequence chart for fixations M1

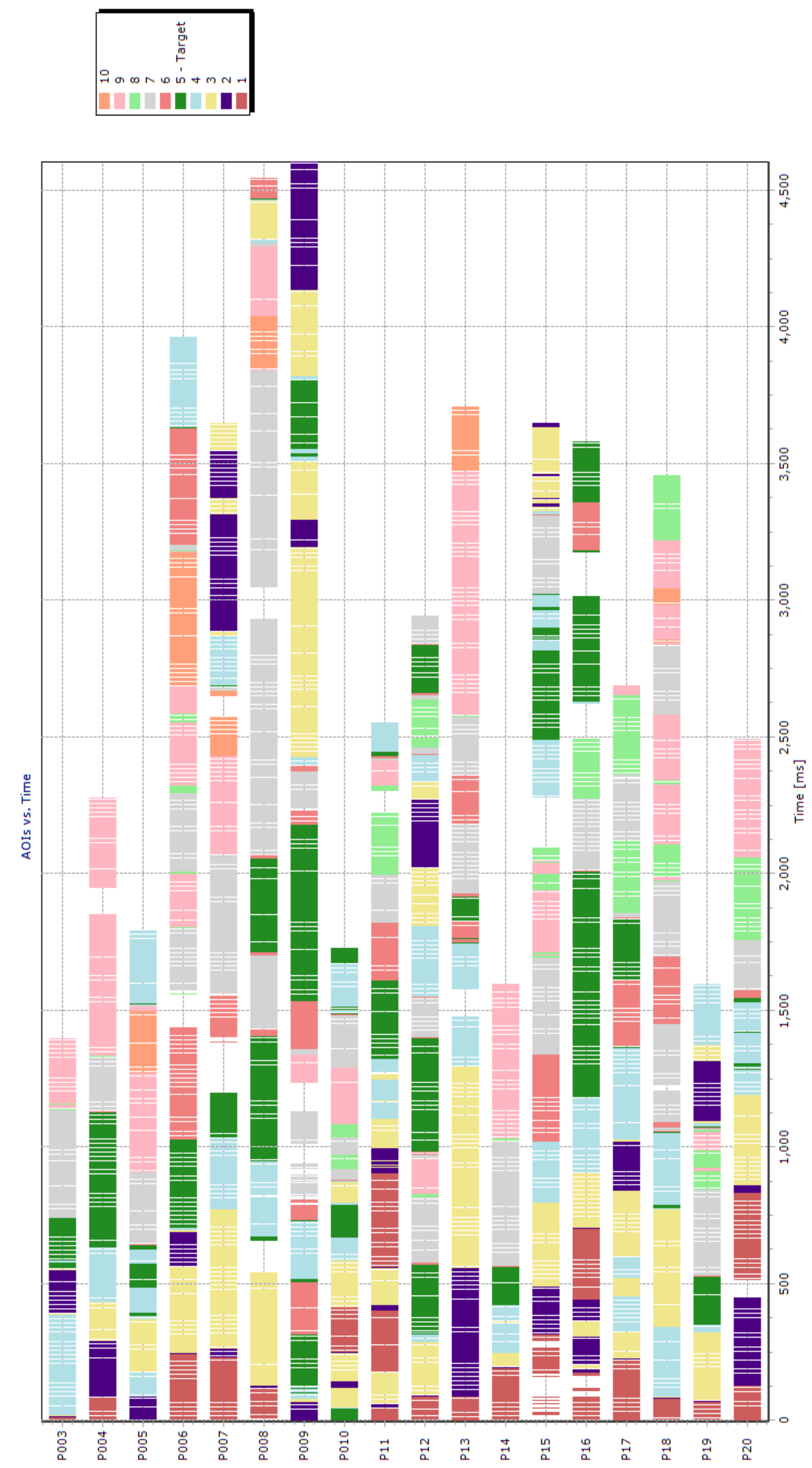

Figure C.95: AOI sequence chart for raw data M1

C.3.2 M1x

Figure C.96: Stimuli M1x

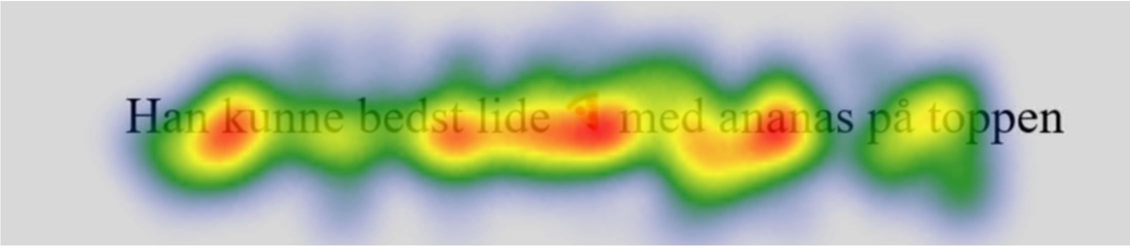

Figure C.97: Heat map M1x

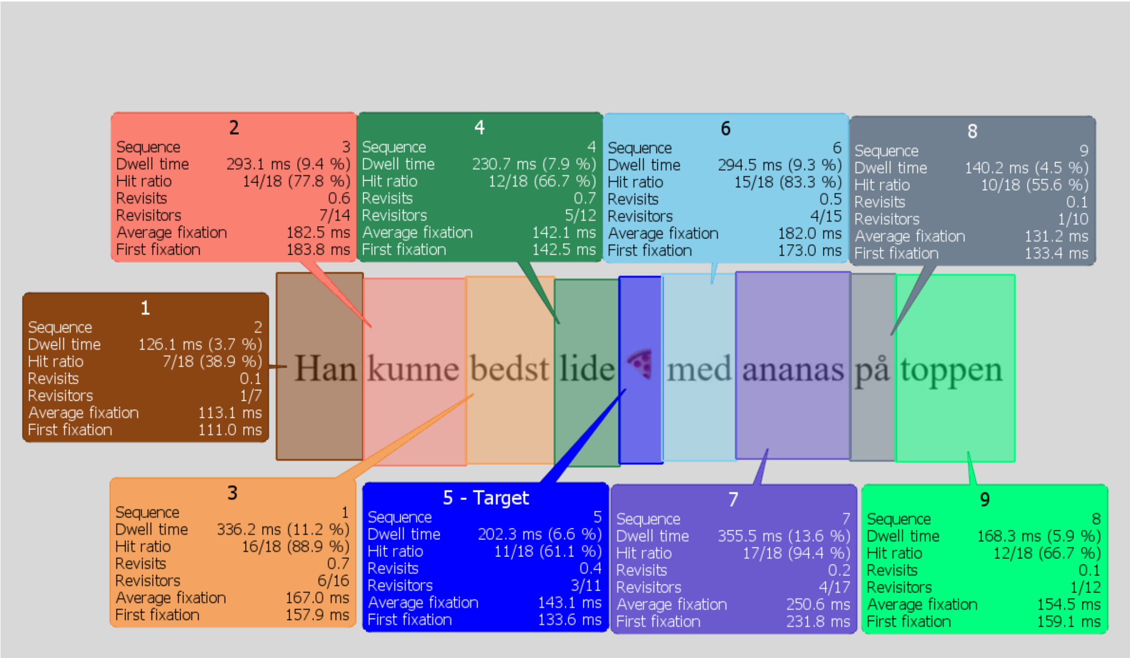

Figure C.98: Metrics M1x

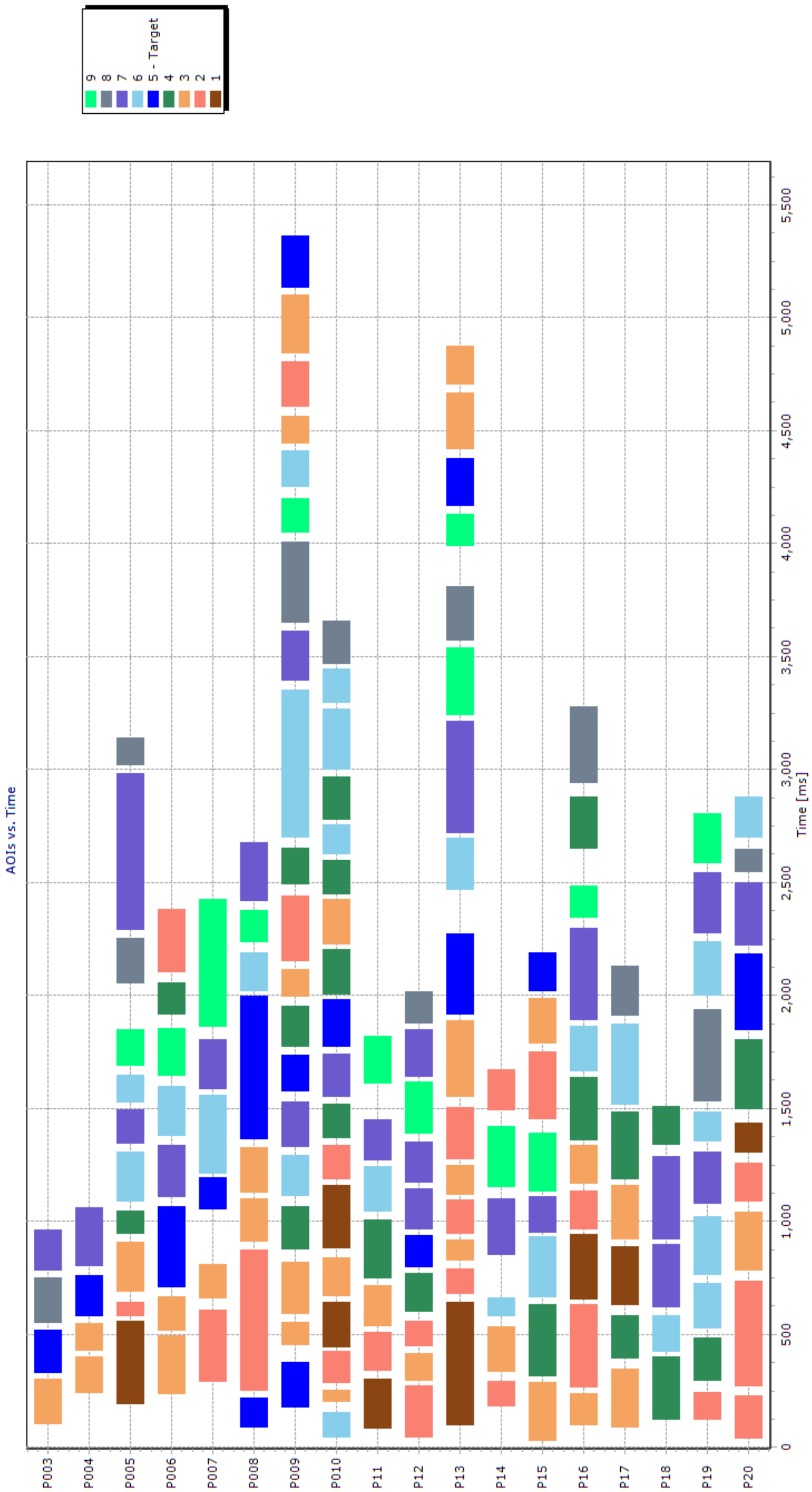

Figure C.99: AOI sequence chart for fixations M1x

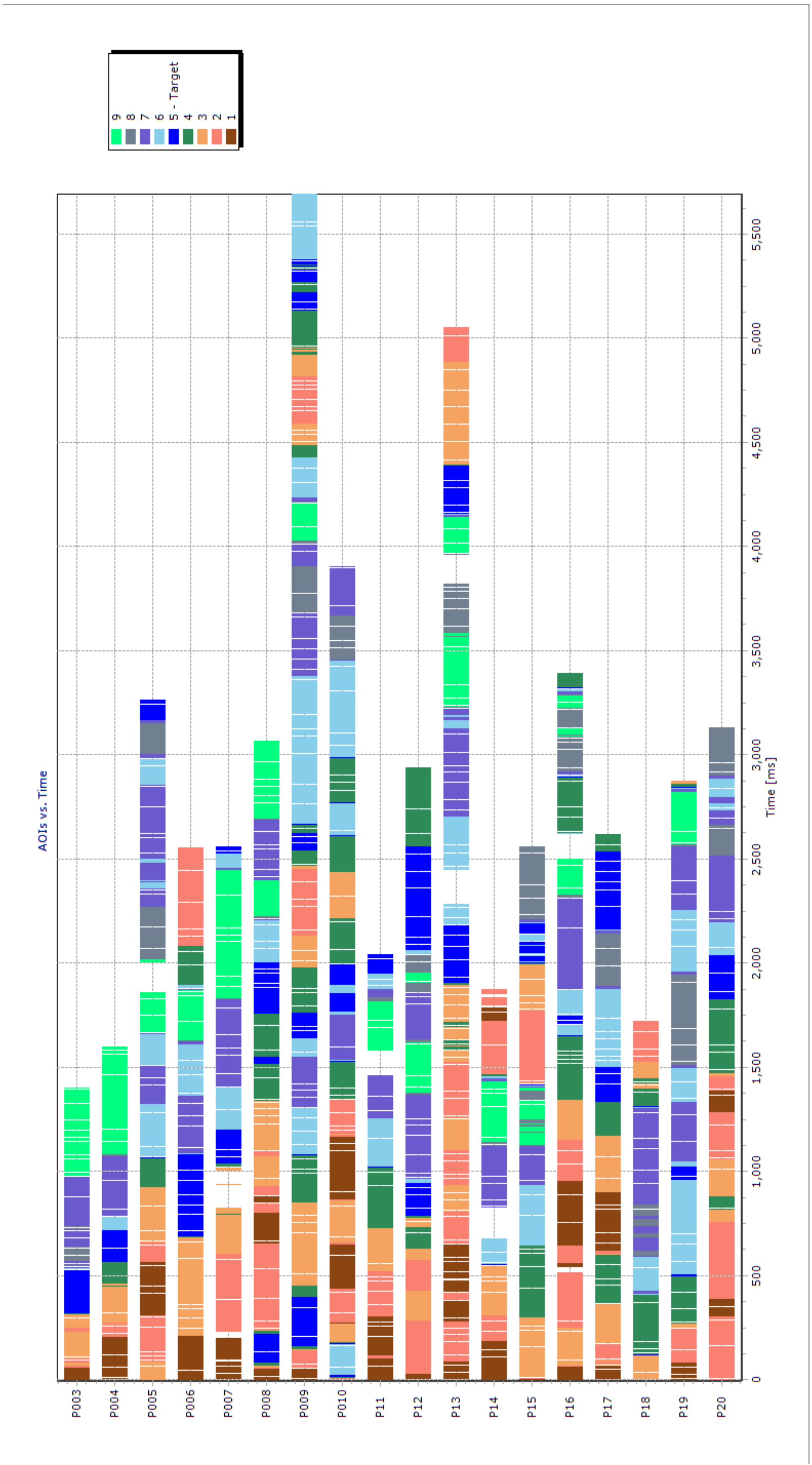

Figure C.100: AOI sequence chart for raw data M1x

C.3.3 M2

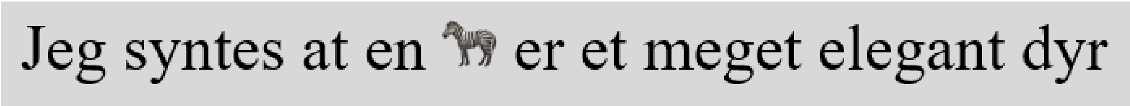

Figure C.101: Stimuli M2

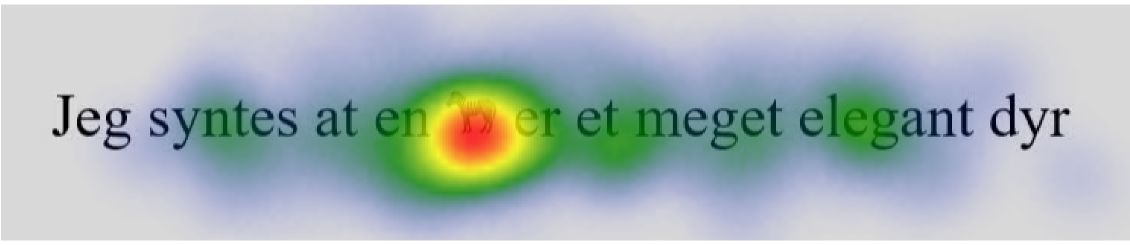

Figure C.102: Heat map M2

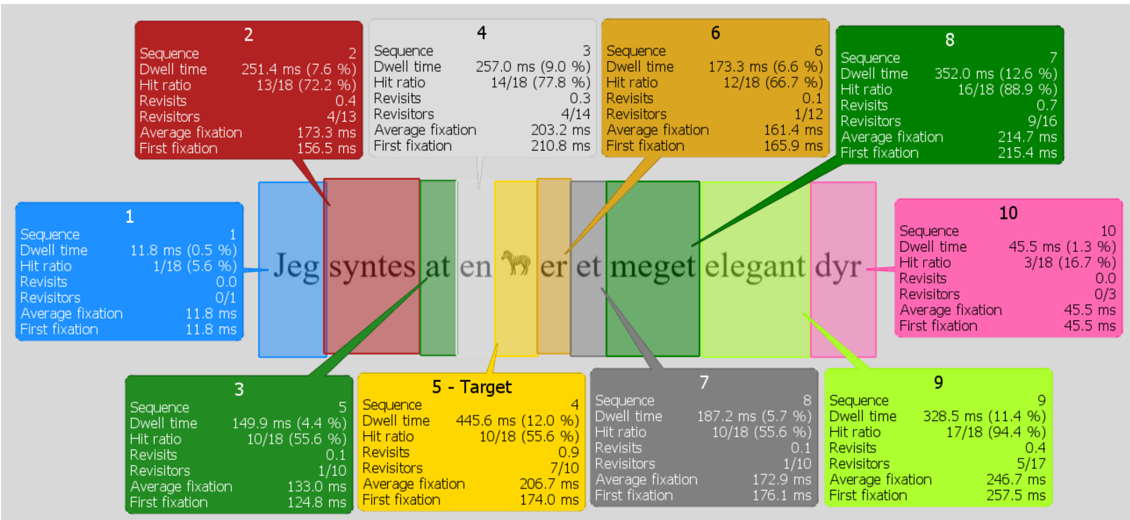

Figure C.103: Metrics M2

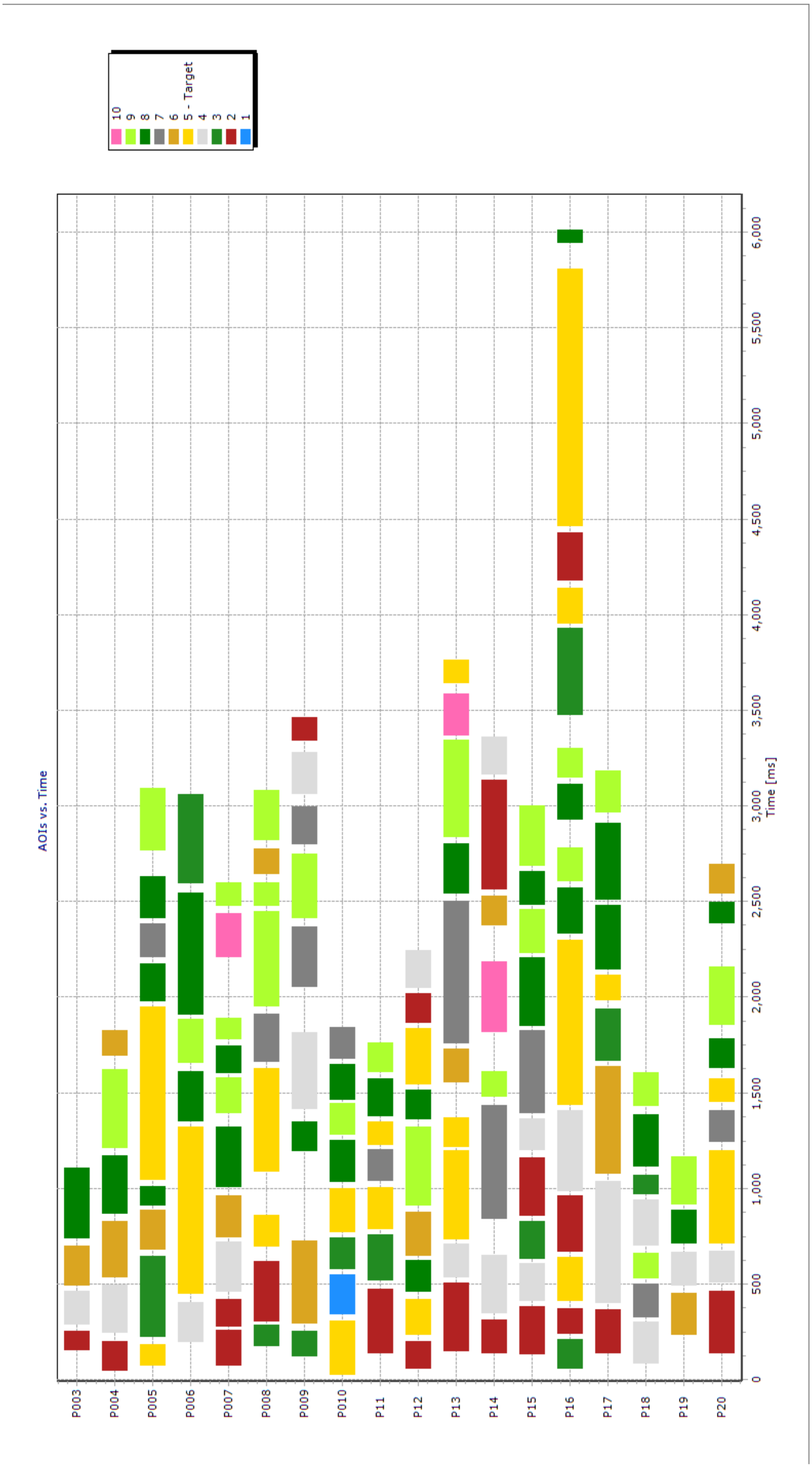

Figure C.104: AOI sequence chart for fixations M2

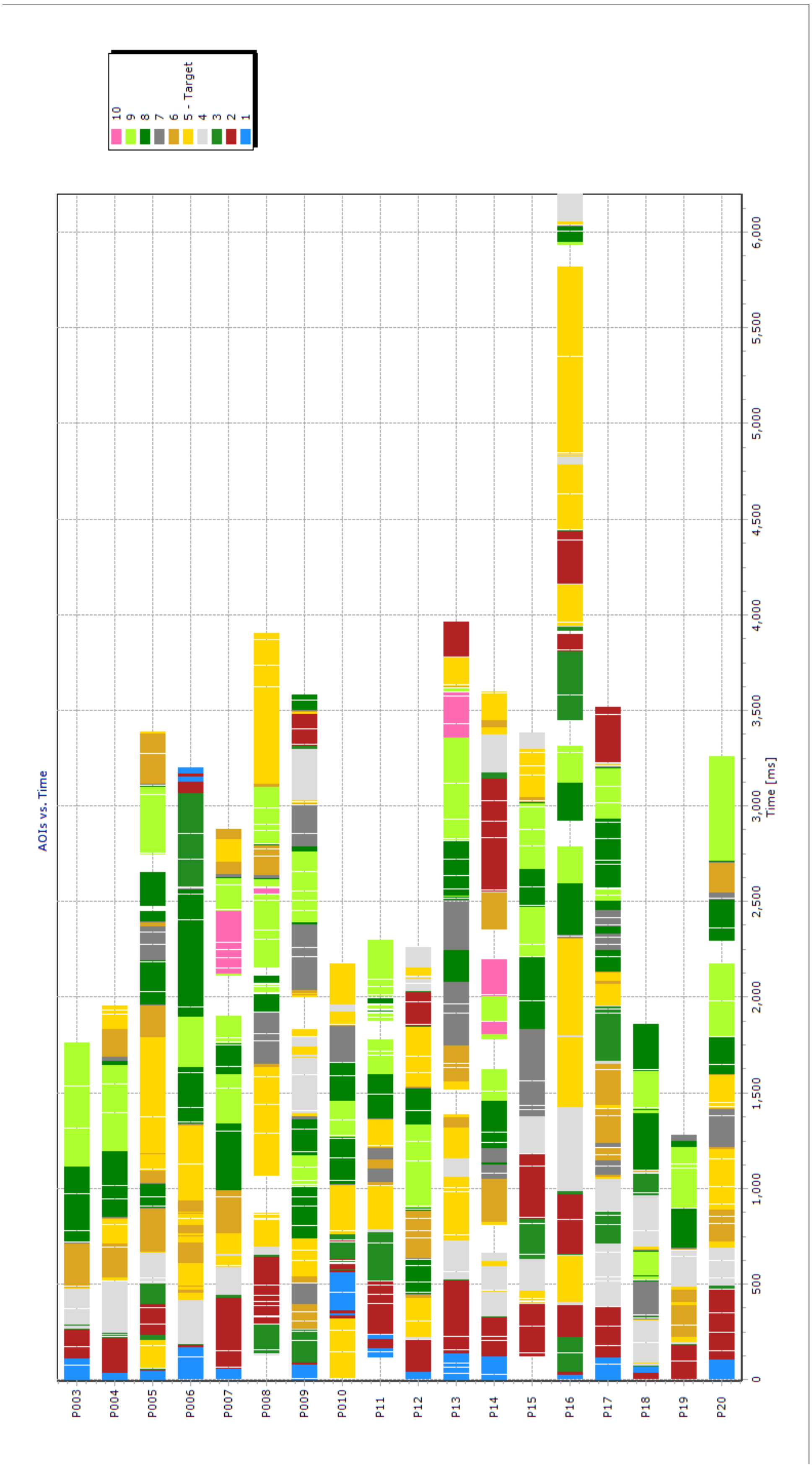

Figure C.105: AOI sequence chart for raw data M2

C.3.4 M2x

Figure C.106: Stimuli M2x

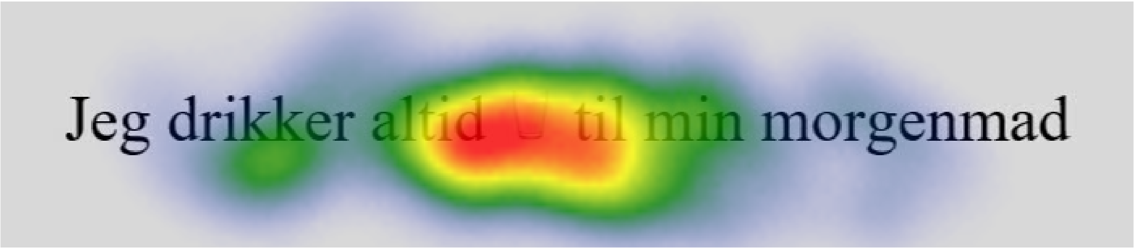

Figure C.107: Heat map M2x

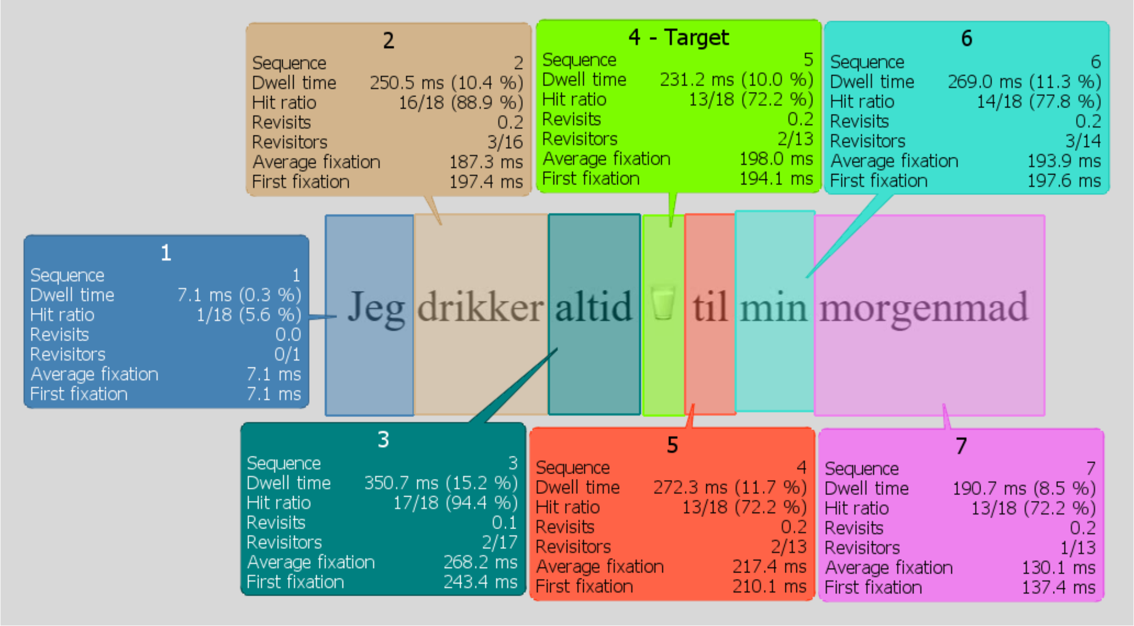

Figure C.108: Metrics M2x

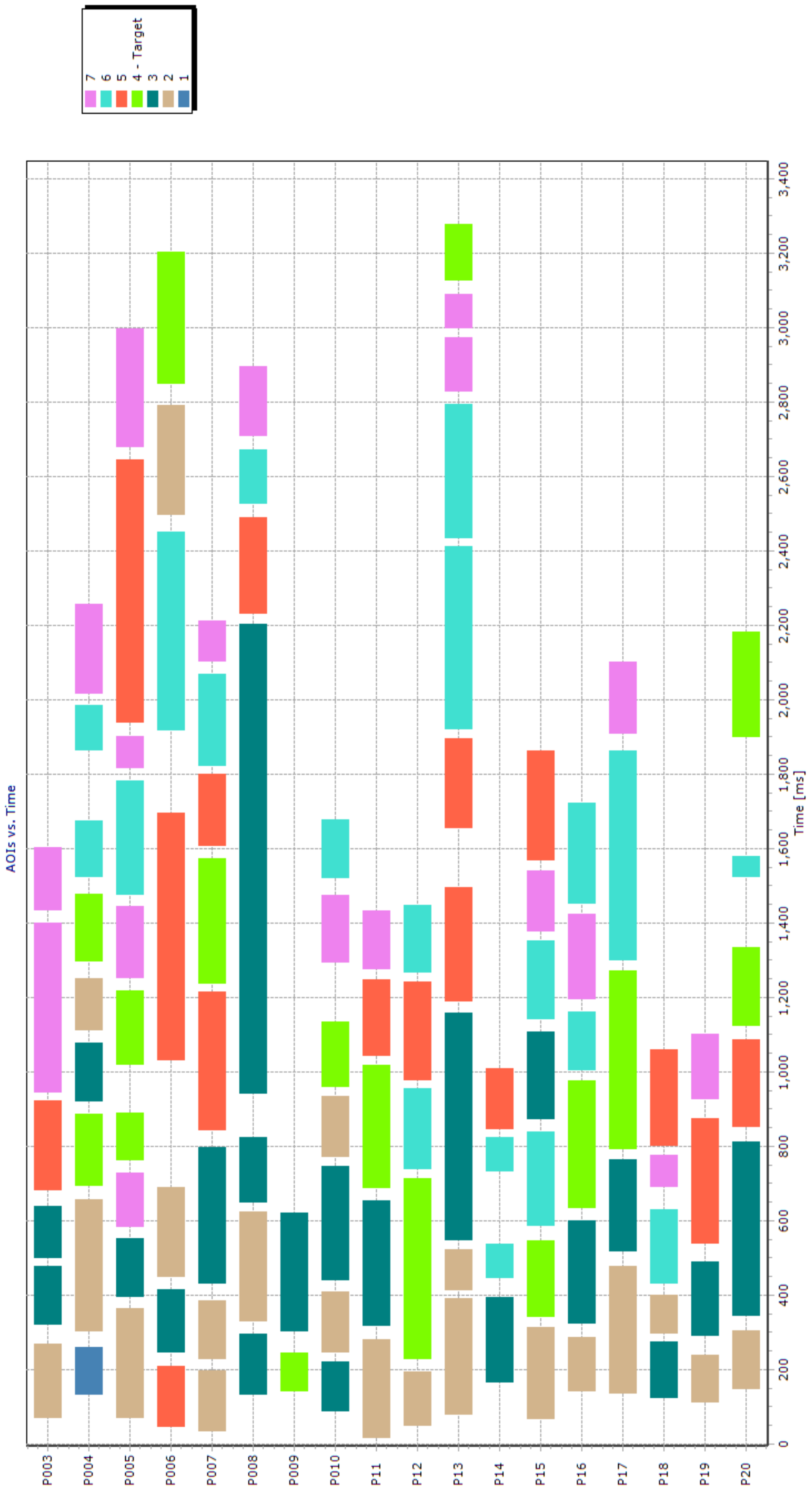

Figure C.109: AOI sequence chart for fixations M2x

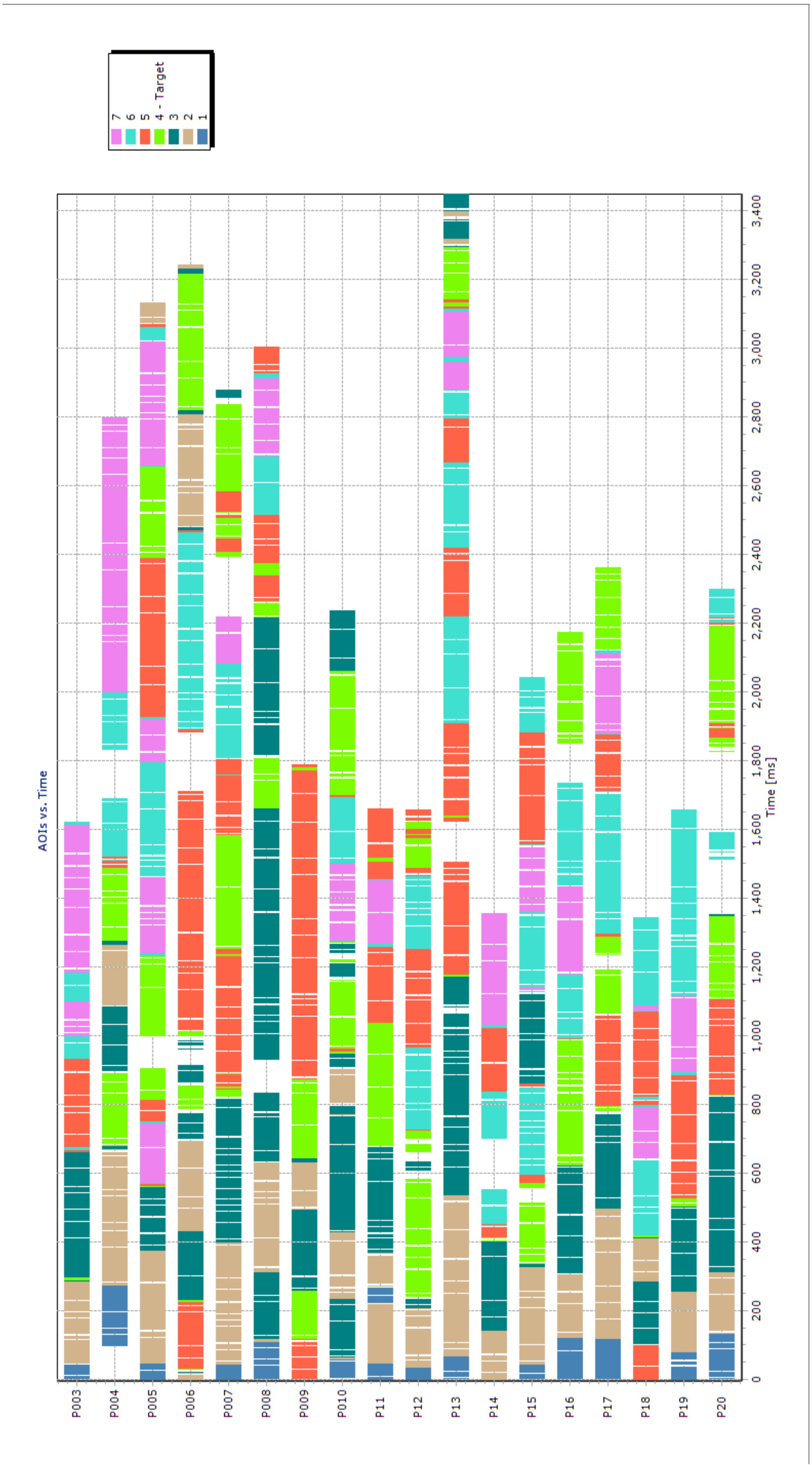

Figure C.110: AOI sequence chart for raw data M2x

C.3.5 M3

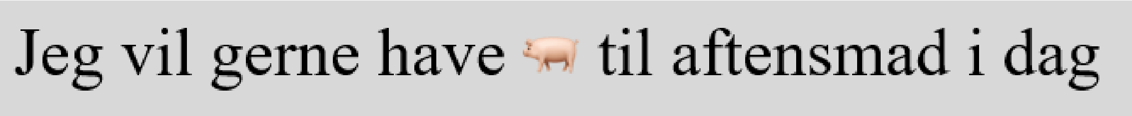

Figure C.111: Stimuli M3

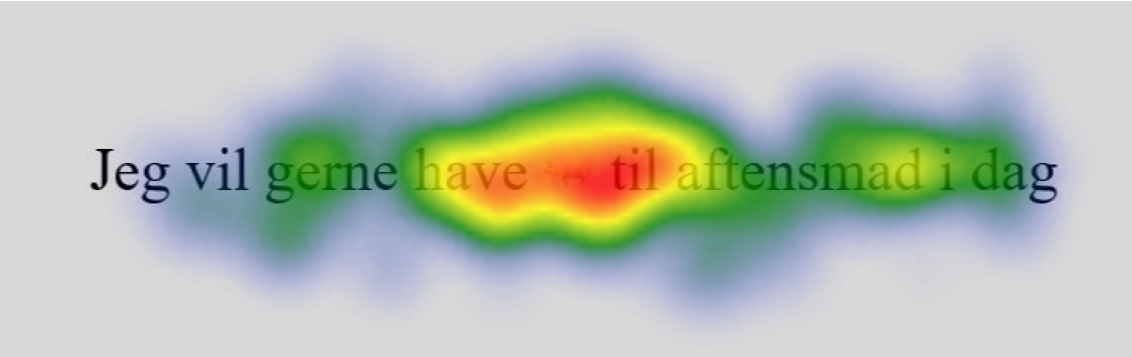

Figure C.112: Heat map M3

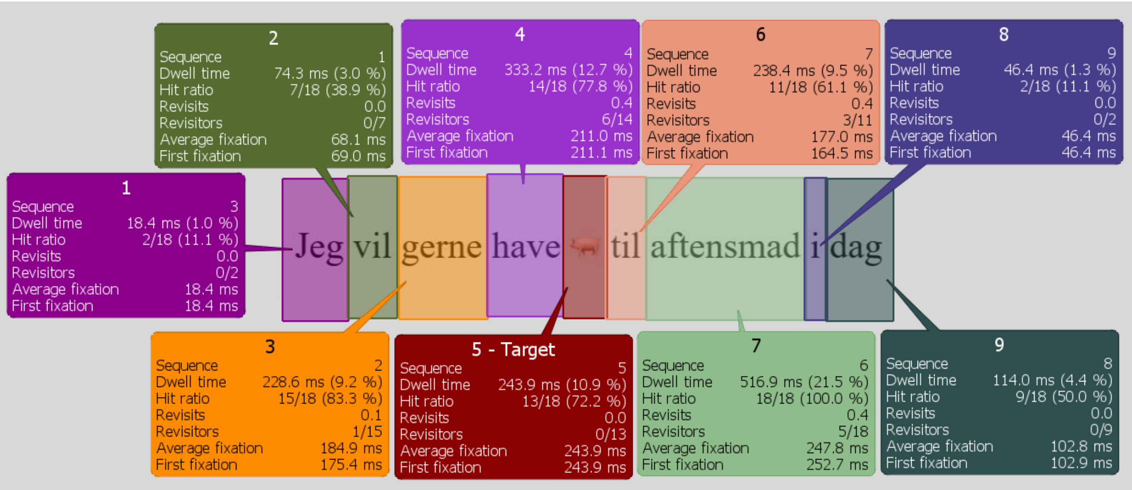

Figure C.113: Metrics M3

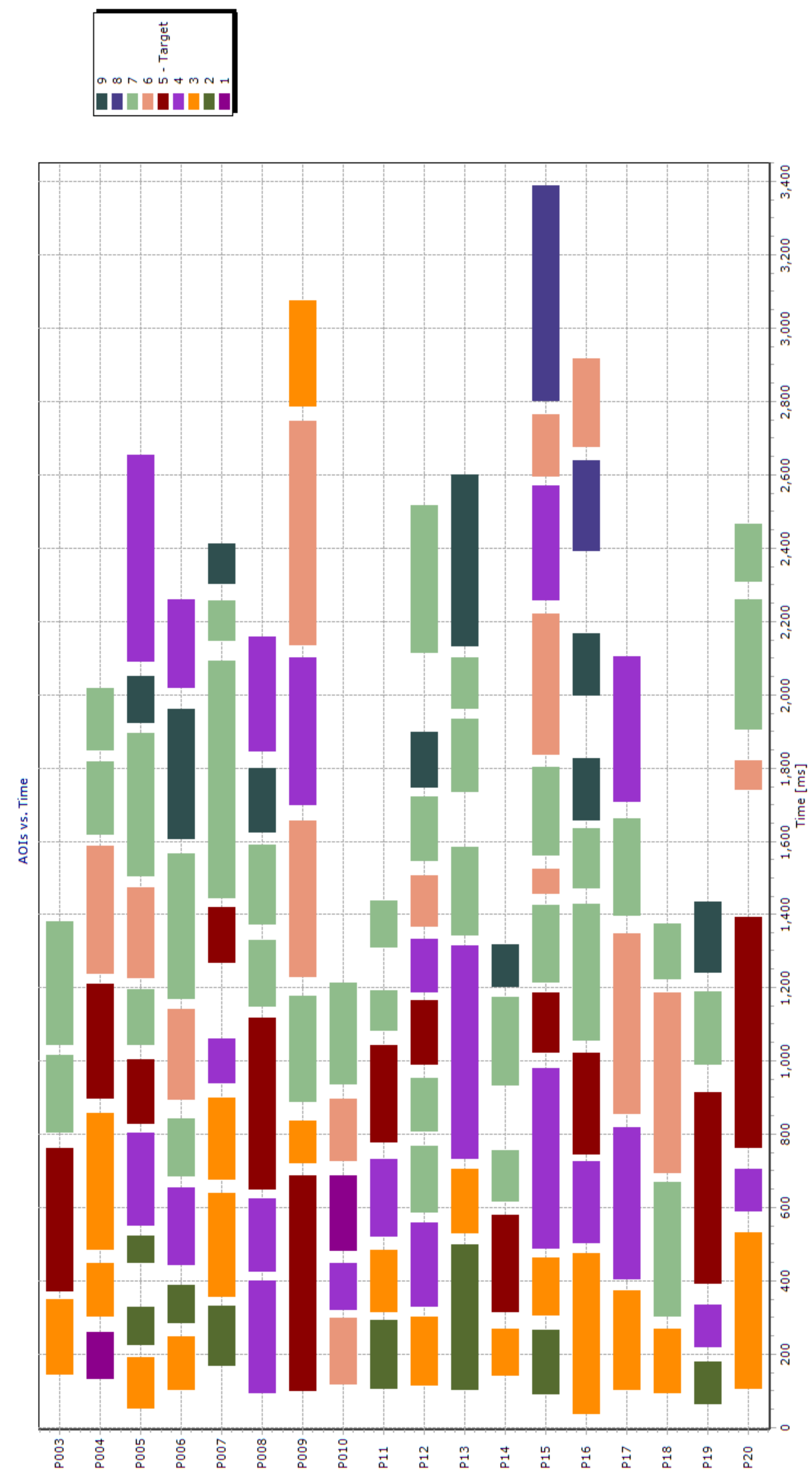

Figure C.114: AOI sequence chart for fixations M3

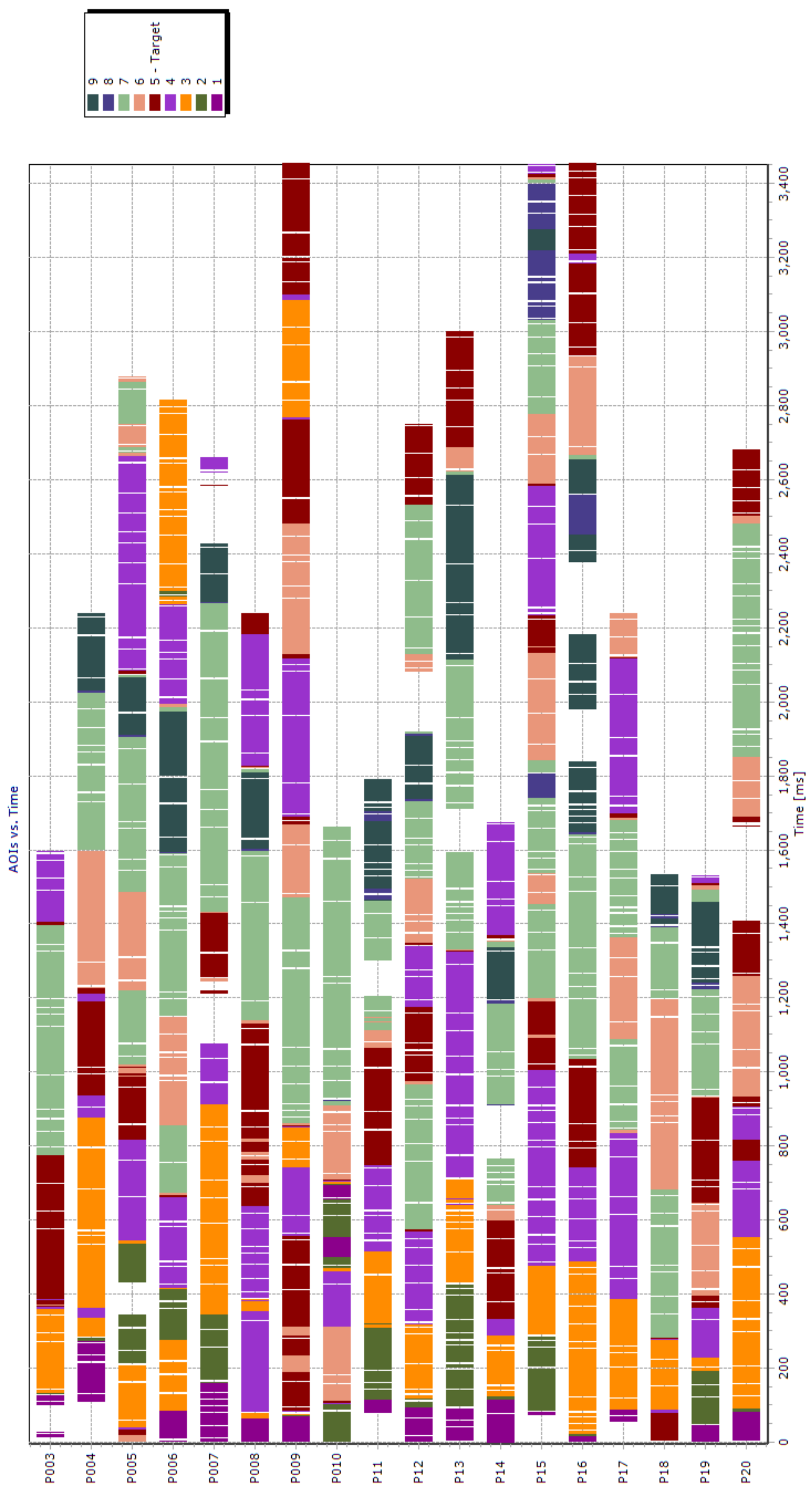

Figure C.115: AOI sequence chart for raw data M3

C.3.6 M3x

Figure C.116: Stimuli M3x

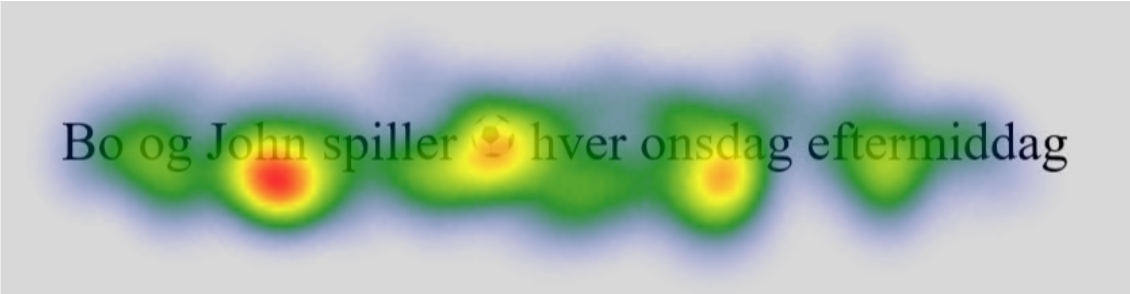

Figure C.117: Heat map M3x

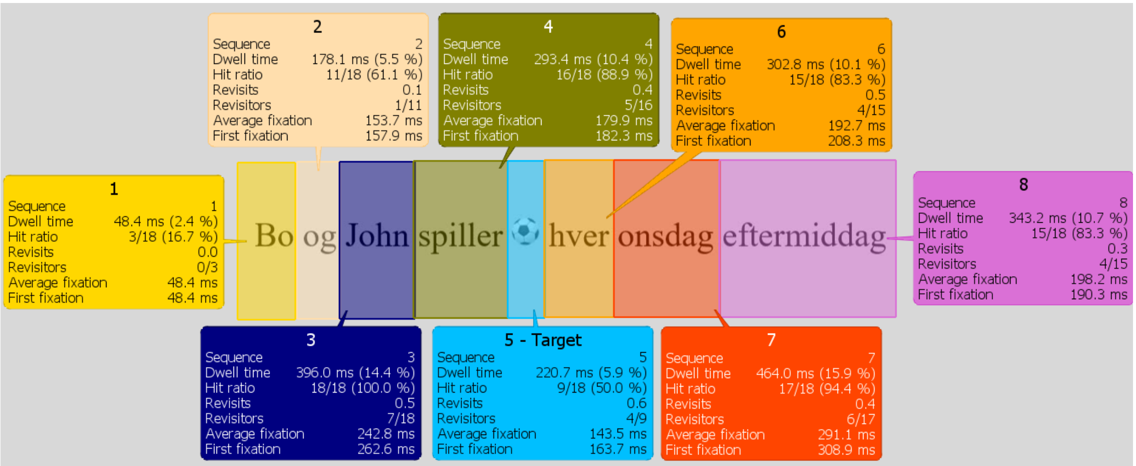

Figure C.118: Metrics M3x

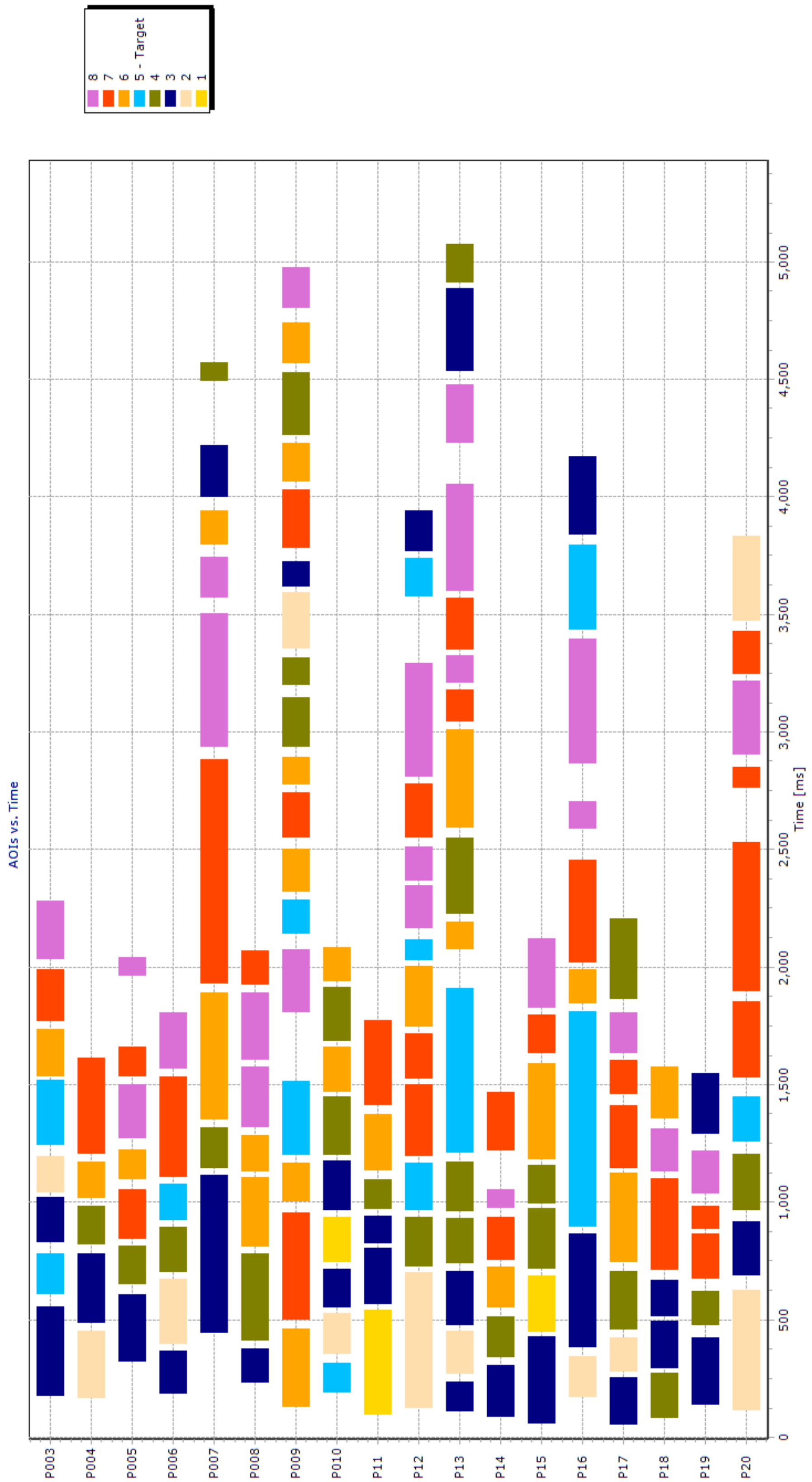

Figure C.119: AOI sequence chart for fixations M3x

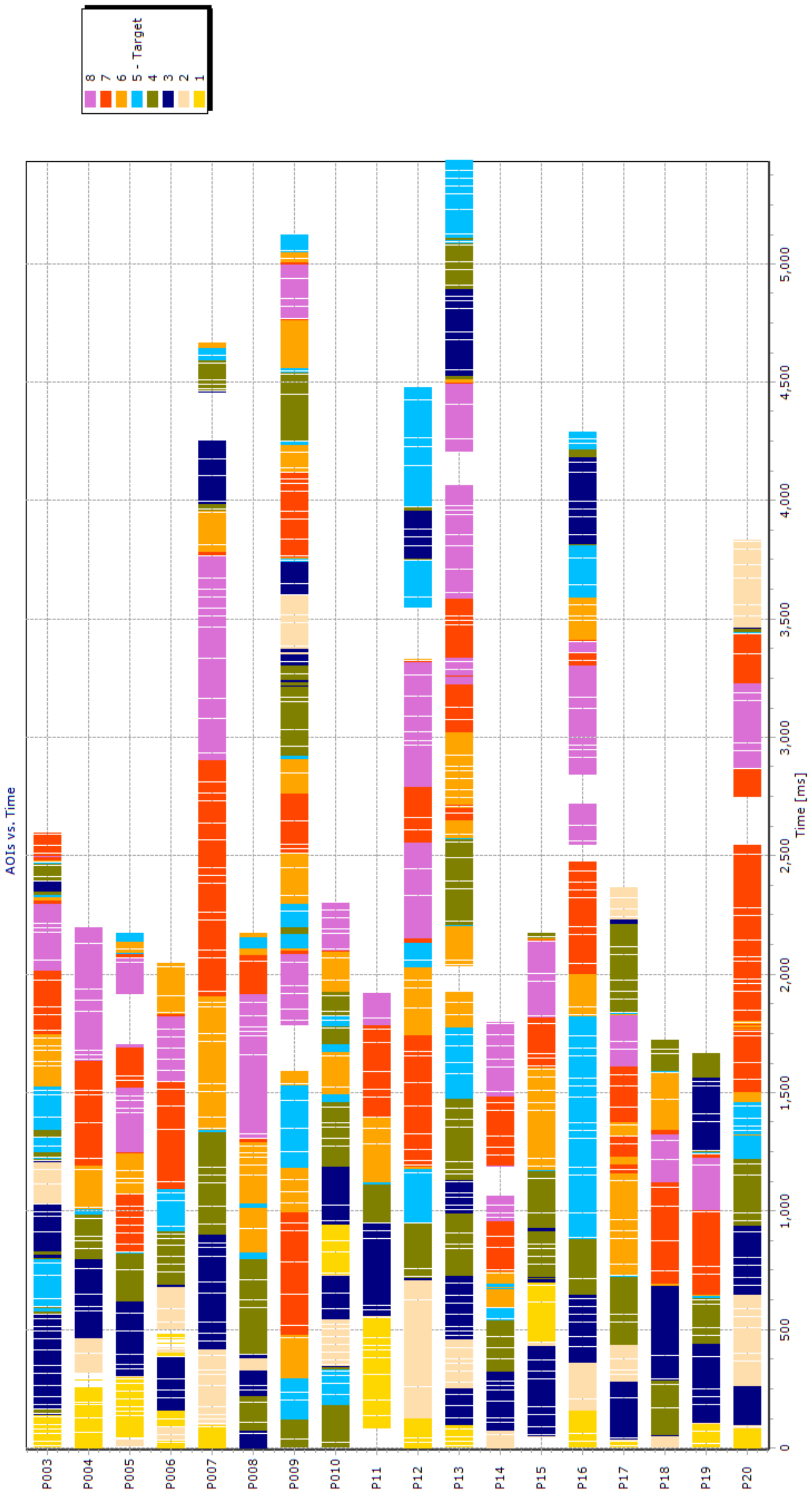

Figure C.120: AOI sequence chart for raw data M3x

C.3.7 M4

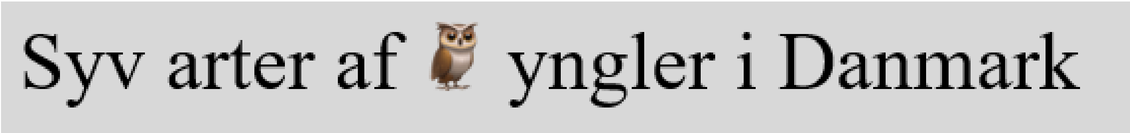

Figure C.121: Stimuli M4

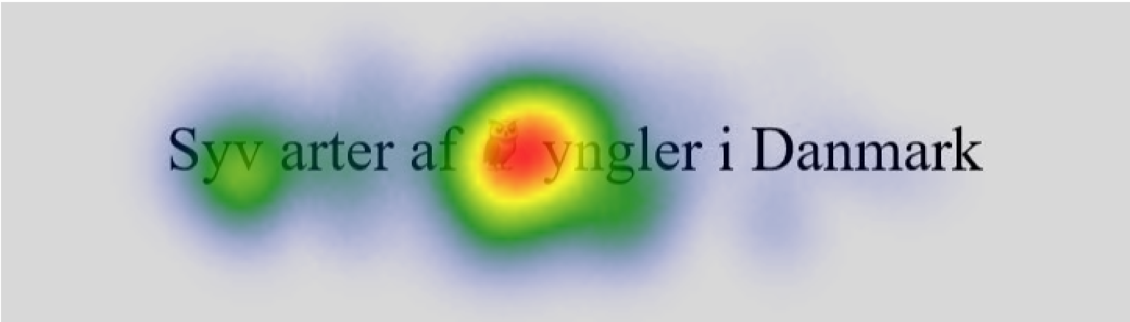

Figure C.122: Heat map M4

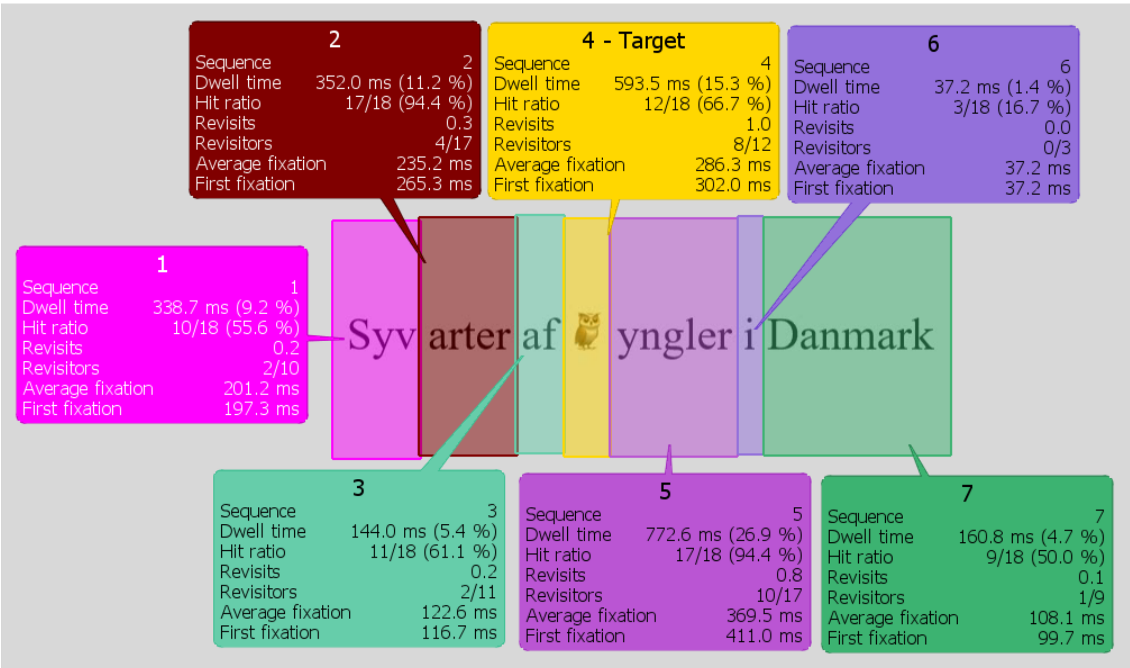

Figure C.123: Metrics M4

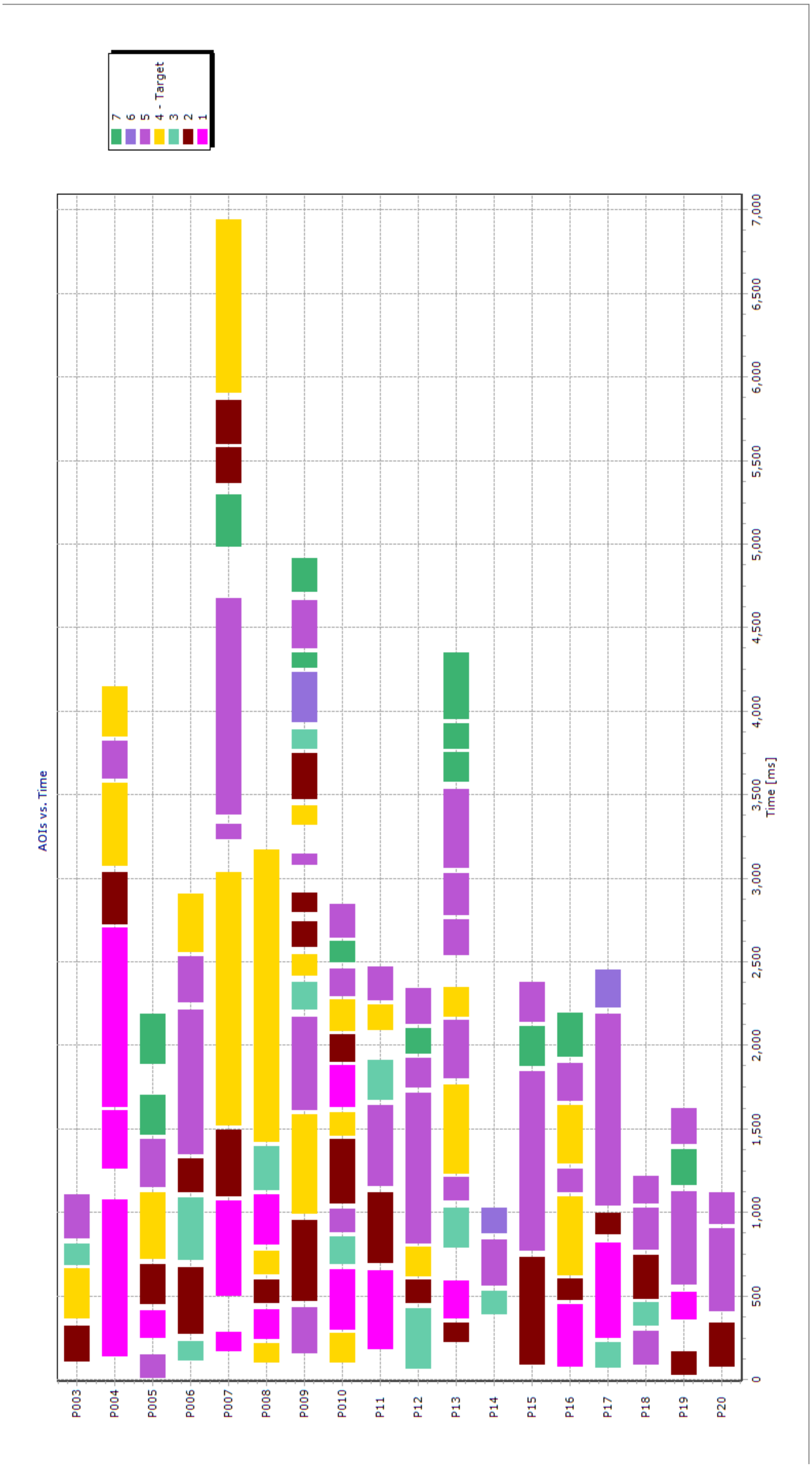

Figure C.124: AOI sequence chart for fixations M4

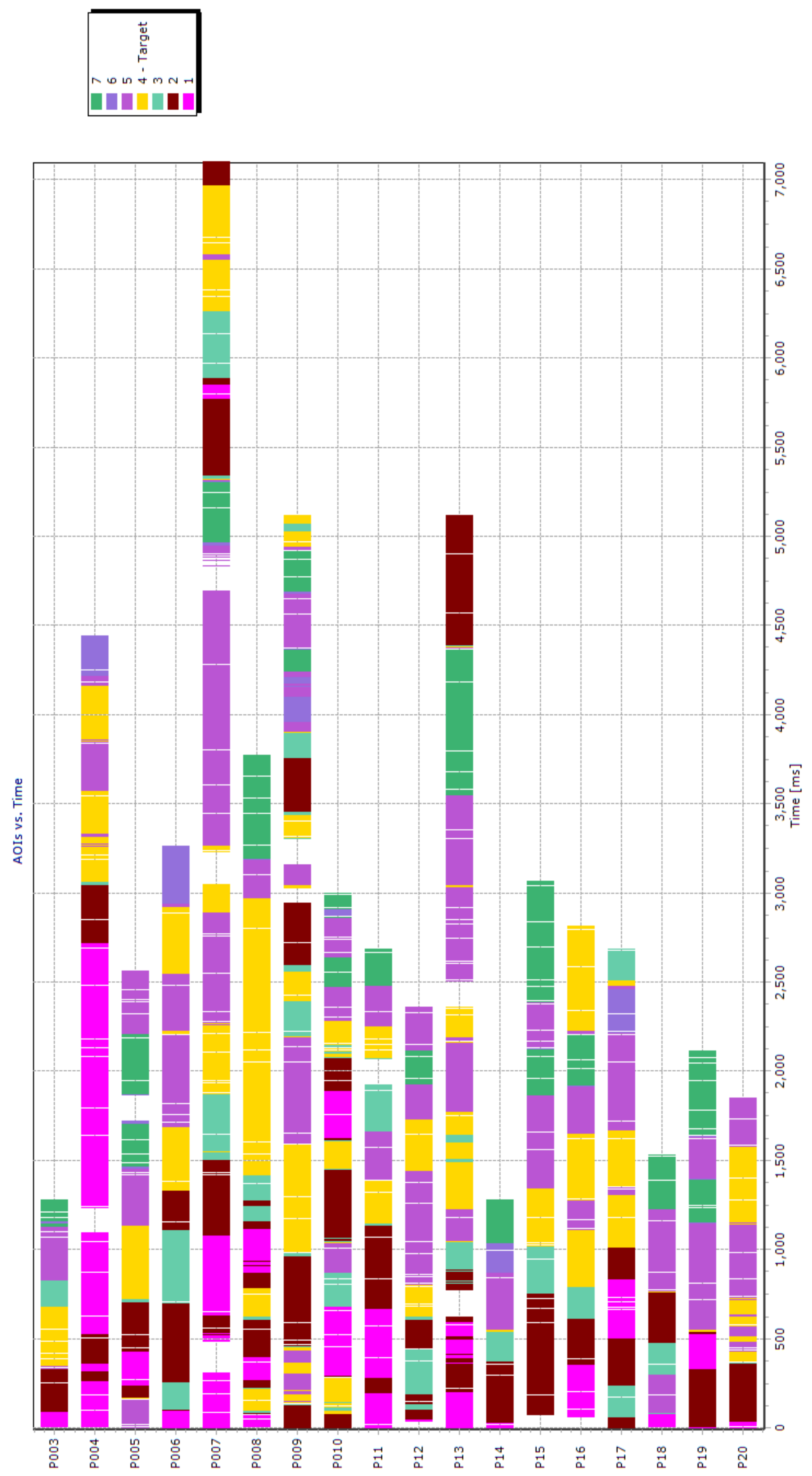

Figure C.125: AOI sequence chart for raw data M4

C.3.8 M5

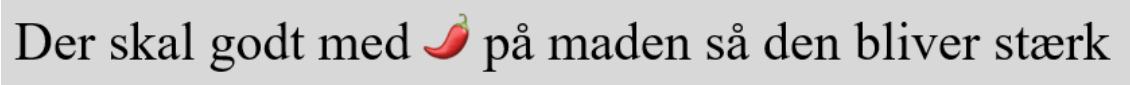

Figure C.126: Stimuli M5

Figure C.127: Heat map M5

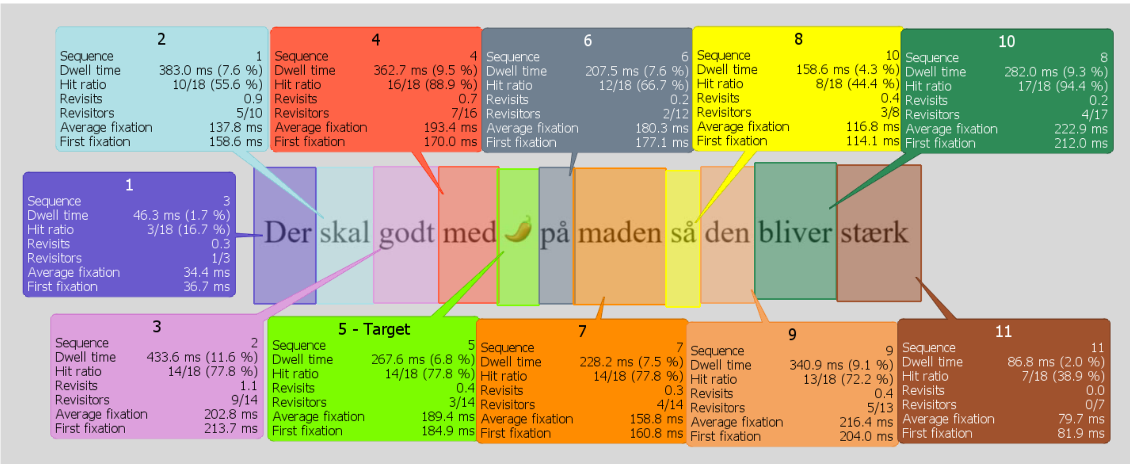

Figure C.128: Metrics M5

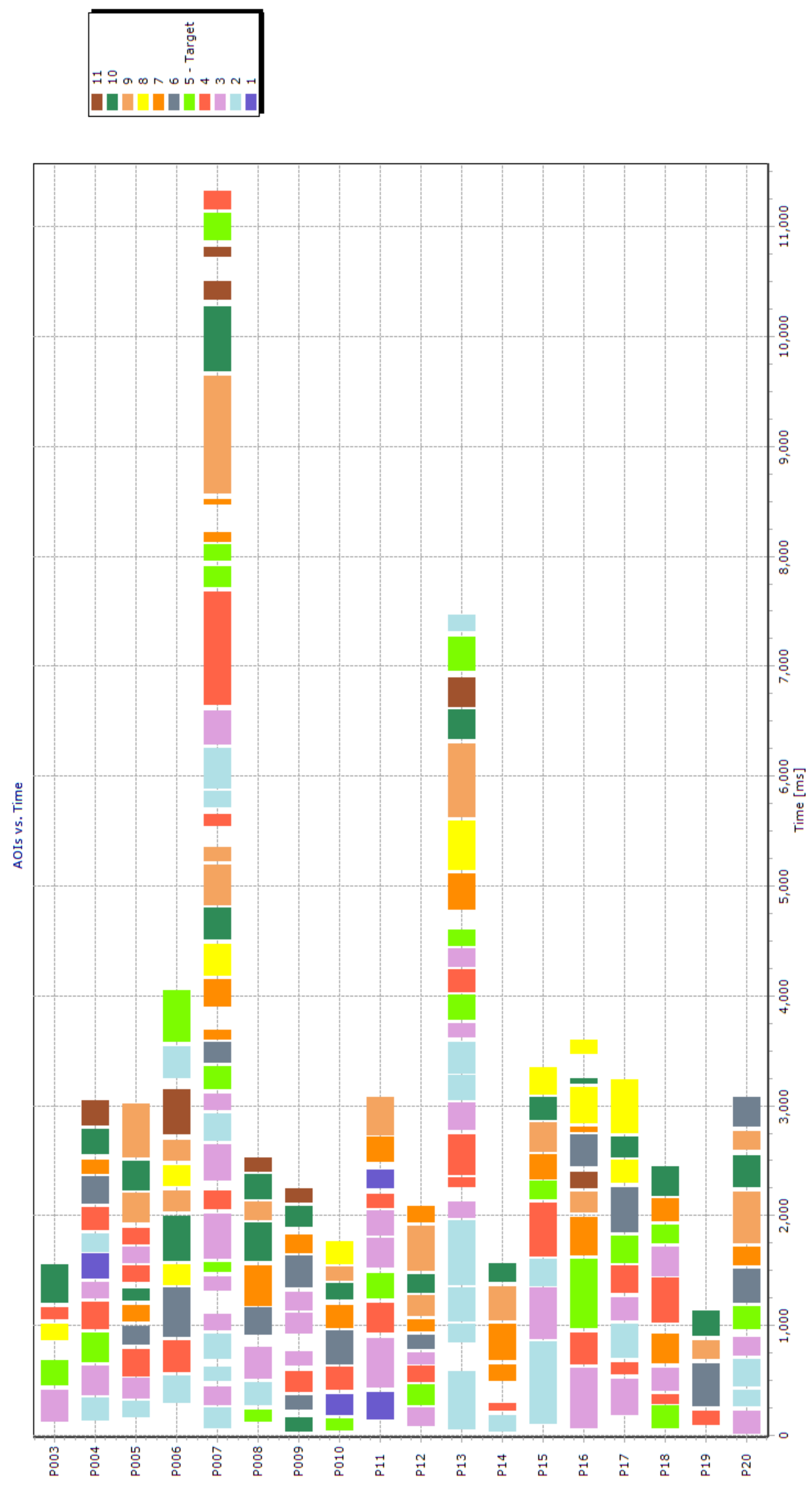

Figure C.129: AOI sequence chart for fixations M5

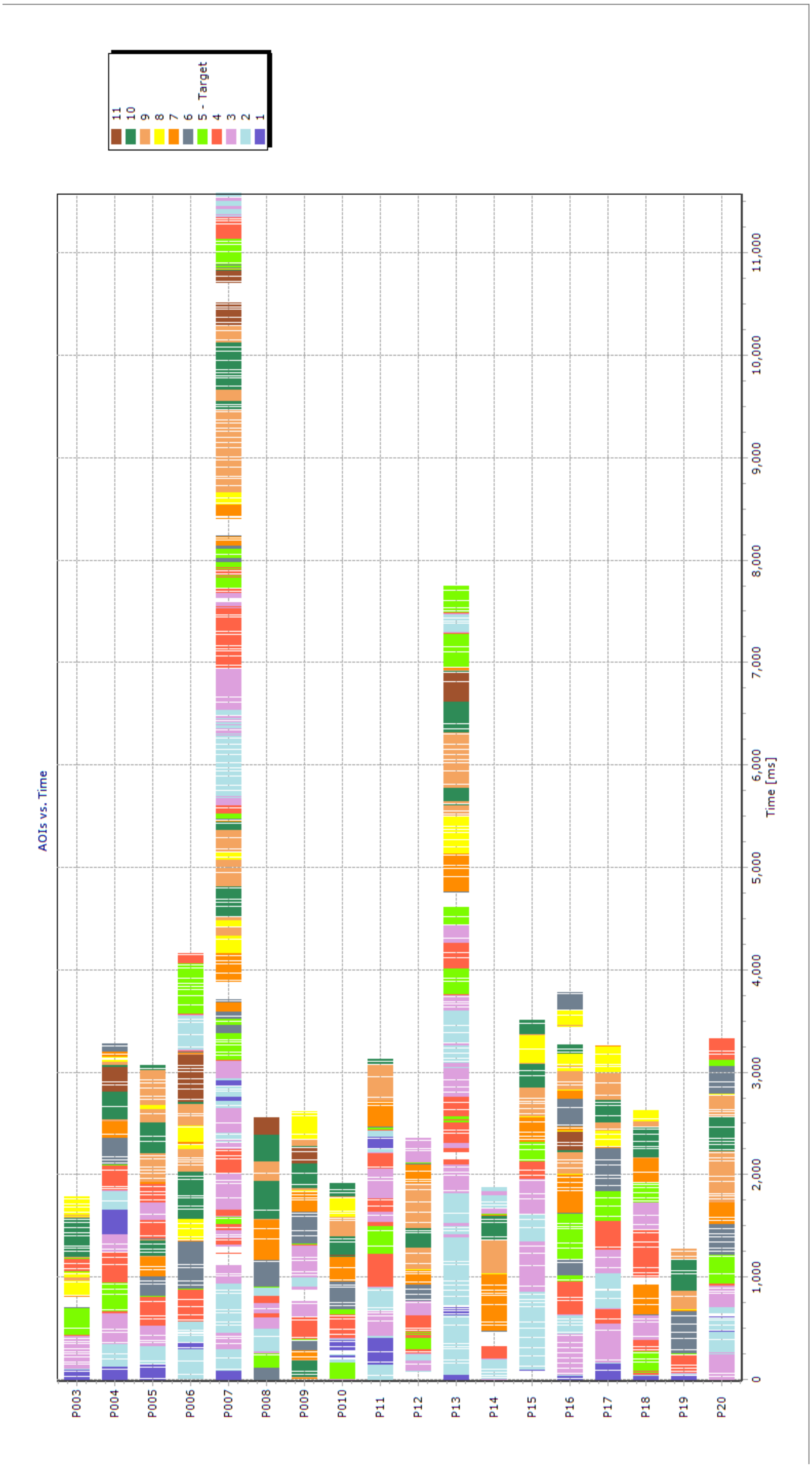

Figure C.130: AOI sequence chart for raw data M5

C.3.9 M6

Figure C.131: Stimuli M6

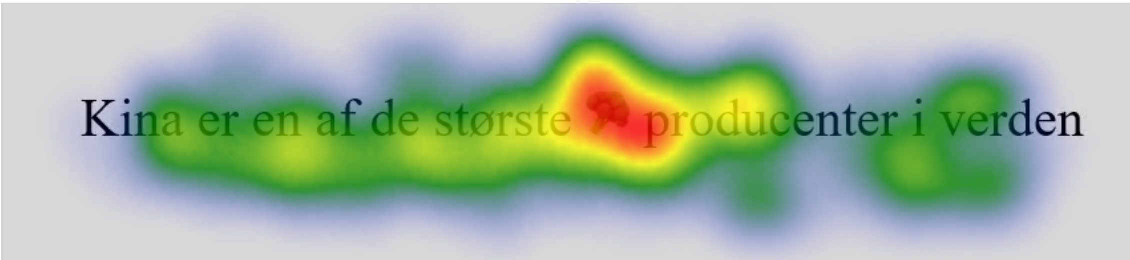

Figure C.132: Heat map M6

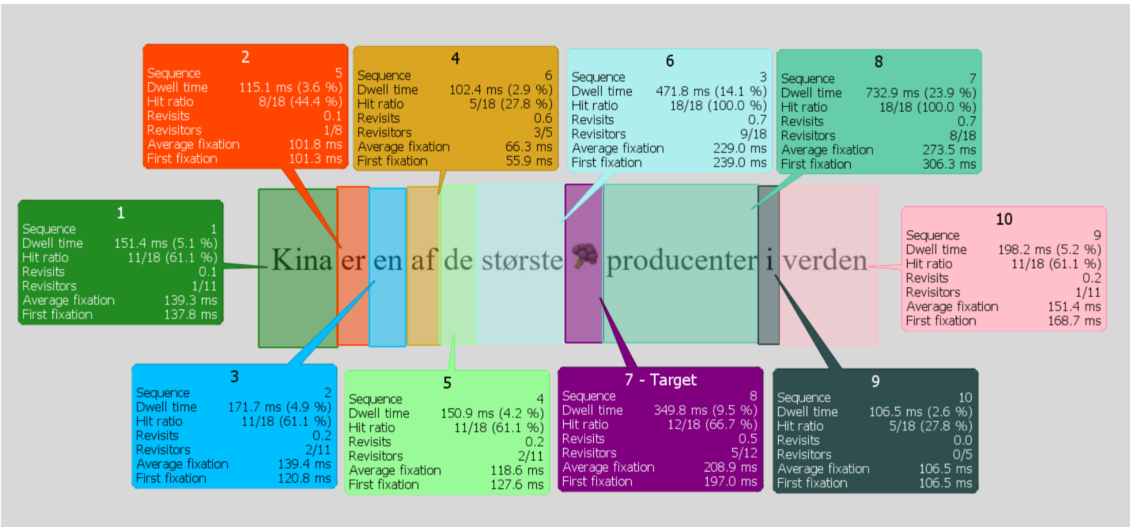

Figure C.133: Metrics M6

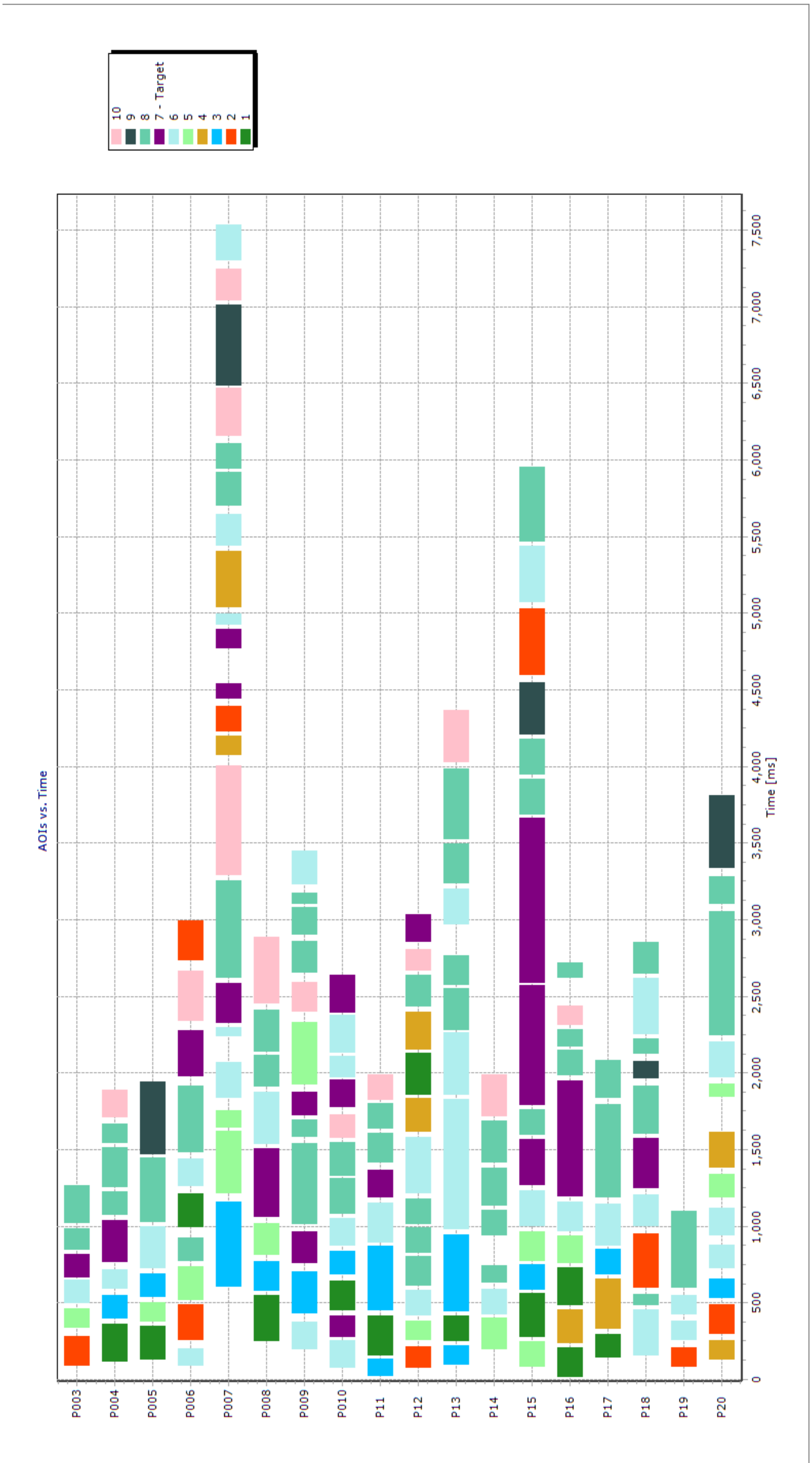

Figure C.134: AOI sequence chart for fixations M6

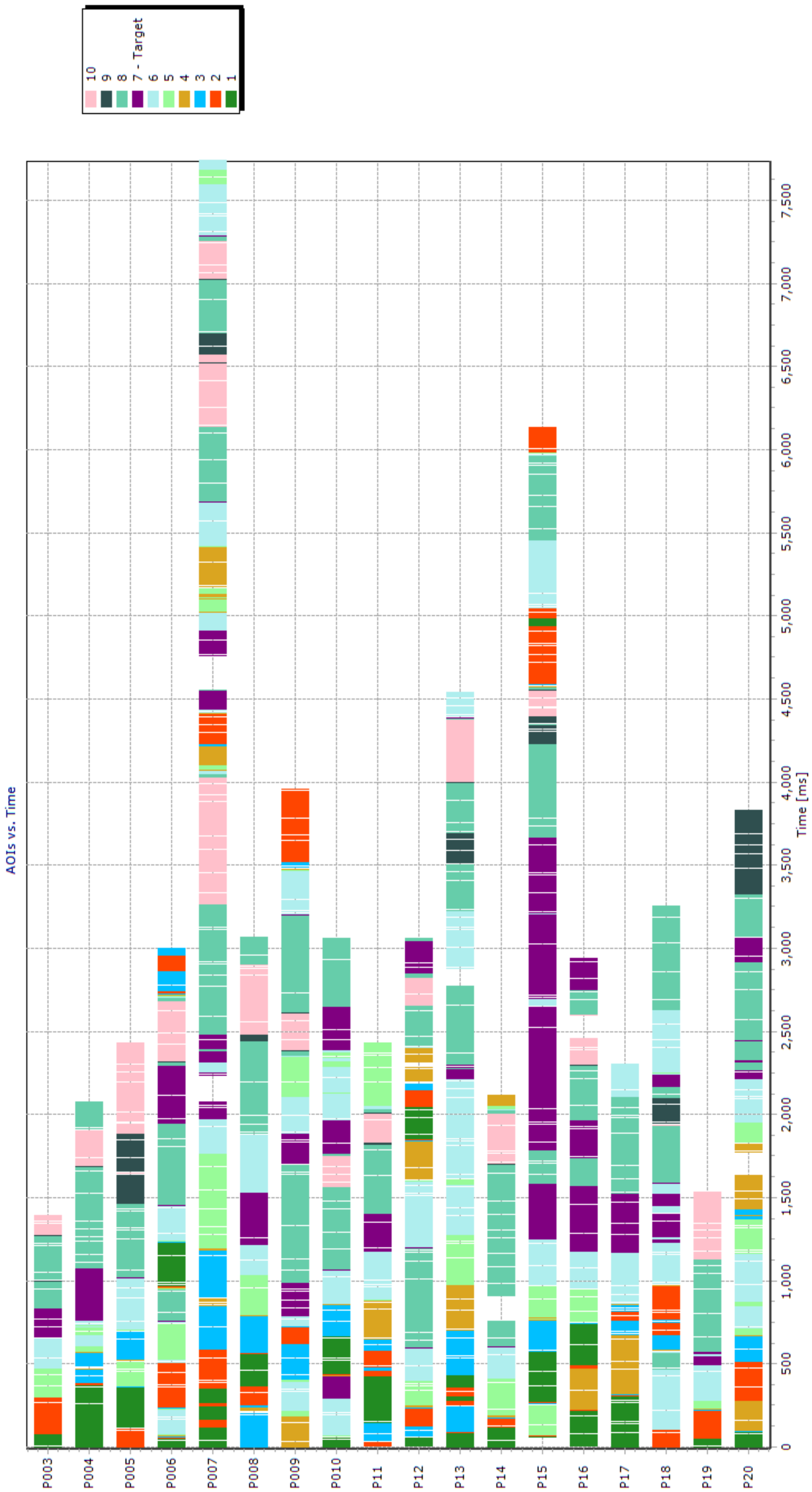

Figure C.135: AOI sequence chart for raw data M6