

Eye-tracking analysis of face observing and face recognition

ABSTRACT

Images are one of the key elements of the content of the World Wide Web. One group of web images are also photos of people. When various institutions (universities, research organizations, companies, associations, etc.) present their staff, they should include photos of people for the purpose of more informative presentation. The fact is, that there are many specifies how people see face images and how do they remember them. Several methods to investigate person's behavior during use of web content can be performed and one of the most reliable method among them is eye tracking. It is very common technique, particularly when it comes to observing web images. Our research focused on behavior of observing face images in process of memorizing them. Test participants were presented with face images shown at different time scale. We focused on three main face elements: eyes, mouth and nose. The results of our analysis can help not only in web presentation, which are, in principle, not limited by time observation, but especially in public presentations (conferences, symposia, and meetings).

Andrej Iskra, Helena
Gabrijelčič Tomc

*University of Ljubljana,
Faculty of Natural Sciences and
Engineering, Snežniška 5,
1000 Ljubljana, Slovenia*

*Corresponding author:
Andrej Iskra
e-mail: andrej.iskra@ntf.uni-lj.si*

First received: 16.12.2015
Accepted: 20.01.2016.

KEY WORDS

Eye tracking, face recognition, fixation duration, face elements, area of interest, heat maps

Introduction

During development of internet various forms of information came along. Now days we have all sorts of element as web content. On basic level we can divide them as text and multimedia. Further on multimedia divides in:

- Images
- Animation
- Sound
- Video (Kyrnin, 2014)

There is a great deal of debate why is important to include images in presentation (either web presentation or publication presentation) (Patel, 2014). From all of these elements images carry most information. That also applies for face images. Web presentations of companies, institutes, universities usually also present employees. From the point of value of person presen-

tation, it is very important to include person's image beside all other personal and professional information. There are some general factors about face images to be attractive to observers (Nielsen & Pernice, 2010):

- Smiling face
- Facing the camera
- Authentic
- Simple background

Our research was done with eye tracking system. In general, an eye tracker is a device for measuring eye positions and eye movement. From this explanation two elements of eye tracking are defined: fixations and saccades. Fixation is when the eye gaze pauses in a certain position, whereas saccades are quick, simultaneous movements of both eyes in the same direction (Cassin & Solomon, 1990). Series of fixations and saccades is called scanpath. Almost all visual information during

saccades is made in the central two degrees of the visual angle (fovea). Outside of these areas (periphery vision) is less informative. Therefore the location of fixation provides us with information of location which was observed and processed during observation session.

On average, fixations last for around 200 ms during the reading of linguistic text, and 350 ms during the viewing of a scene. There are many researches according to the meaning of fixation duration. For example, in terms of reading longer fixation normally means difficult words (Pollatsek, Rayner & Balota, 1986), in terms of visual aspect refers to difficulty in extracting information (Hooge & Erkelens, 1998). At viewing face image, we deal with longer fixation duration than other image content (e. g. natural scene) (Guo et al., 2006). And which face elements attract most attention? Several authors have been exploring these most attracted face elements, which are also most important for face remembering. All researches have shown three major face elements that significantly stand out of others: eyes, mouth and nose (Henderson et al., 2001; Buchan, Pare & Munhall, 2007). Based on those three main face elements a comparison between free and restricted viewing learning condition (Henderson, Williams & Falk, 2005) was conducted.

There are also slight differences at observed time distribution between main three elements in dependence of face expression, gender and age (Cangöz et al., 2013). Heisz & Shore (2008) investigated how face elements scanning changes for familiar faces. Face elements are also very important in cross-face recognition (stimuli and test participants from different races) (Goldinger, He & Papesh, 2009; Hills & Pake, 2013).

We conducted two tests. In the first test the aim of research was to investigate how participants viewed main three face elements according to the available time and what is the distribution of observation time for eyes, mouth and nose. Are eyes really mostly the first visited face element? How long the face presentation time must be that observer return to the eyes again?

The purpose of second test was to get basic information how well observers remembered faces according to the time of face shown. We simply measure the percentage of correct recognition of previously shown faces.

Material and methods

Participants

Participants were recruited among students at Faculty of Natural Sciences and Engineering (Ljubljana, Slovenia). They were divided into three test groups (A, B and C), so each has 14 participants. They all have normal or corrected-to-normal vision. Age range was

from 19 to 22 years old. They all volunteered in return for a small bonus at laboratory practice grade.

Apparatus

As mentioned above, the key instrument for our investigation was eye tracking system. Eye tracking is mainly used for researches in the field of usability investigations in all sorts of research areas. That can be done for web sites, newspapers, magazines, games, stores (real stores in shopping centers), TV commercials, traffic, education, etc. (Horsley et al., 2014).

Tobii X120 eye tracking system was used in our experiment. It is stand-alone eye tracking unit designed for eye tracking studies of real-world flat surfaces. We used it in combination with computer's LCD monitor.

This system has sampling rate of 120 Hz for recording eye movements. The default setting for definition of fixation was 100 ms for 30 pixel area. That means if gaze stayed in the area 30 pixel for at least 100 ms it was concerned as one fixation (Goldinger, He & Papesh, 2009). If gaze left the region and returned within 100 ms, it was considered to be the same gaze. Recommended distance from eye tracking system is 60 cm (Goldinger, He & Papesh, 2009). All recordings and analyses were done using Tobii Studio 3.1 software.



» **Figure 1:** *Eyetracking testing setup*

Stimuli

We took and downloaded 60 high resolution images of faces on the web. Then we cropped them to the size 600 pixel × 600 pixels and assure all images have face elements at the same position. We categorized faces by sex, age and race.

Visual appearance age was categorized in three segments:

- Children and young people
- Middle age people
- Elderly people

Race was visually categorized and equally used as:

- White (Caucasian) race
- Black (Negro) race
- ellow (Mongolian) race

Analysis of how race influence face recognition would require further study because it has additional social and psychological aspects (Klama & Milton, 2012; Hills & Pake, 2013), so we made a general set of faces by race.

Procedure

We prepared three tests with different duration of face view. Each test contains 10 faces categorized as mentioned above. Faces in test “1 second” were shown for one second, at test “2 second” for two seconds, and at test “4 second” faces were shown for the period of 4 seconds. Each participant completed all three tests, but we changed the order of test for three groups in the shape of Latin square shown in Table 1. Latin square ensures test to be independent from participant’s concentration which is fading as time passes.

Table 1

Latin square of tests order

Test group	Order of tests		
	“1 second”	“2 seconds”	“4 seconds”
Test group A	“1 second”	“2 seconds”	“4 seconds”
Test group B	“2 seconds”	“4 seconds”	“1 second”
Test group C	“4 seconds”	“1 second”	“2 seconds”

All 10 faces were shown in continuous order with 1 second pause of black screen. The purpose of that was to neutralize position of eye gaze before appearance of new face.

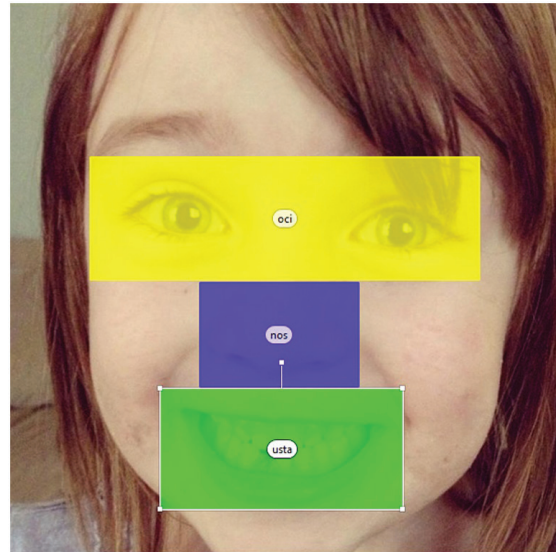
Experiment 1

We focused on three main face elements (Golding-er, He & Papesh, 2009; Hills & Pake, 2013): eyes, nose and mouth. To obtain required metrics of these face elements we set AOI (Area of Interest). Figure 2 shows example of our AOI on analyzed face image.

Each recording had sample percentage. Sample percentage shows how many sent signals from eye tracker were correctly recorded. For example, if participants look away from the monitor eye tracker couldn’t record his gaze. Other causes of problem with receiving signal were glasses, women’s eye makeup, ambient light, etc.

In general, sample rate tells us the quality of recordings and according to instructions (Tobii Technology AB, 2012) we eliminated recordings under 90 % recorded samples (we treated them as bad recordings). Normally, longer face presentation had better recorded samples. This 90 % criteria was passed by 34 participants in “4 second”

test, 29 participants in “2 second” test and 27 participants in “1 second” test. In faster visual information change (shorter time of face presentation) the velocity of eye movements was higher, so, consequently, the possibility for eye tracker to lost track of eyes was also higher.



» **Figure 2:** Our AOI at face image (eyes, nose and mouth)

Within these three AOI we investigated:

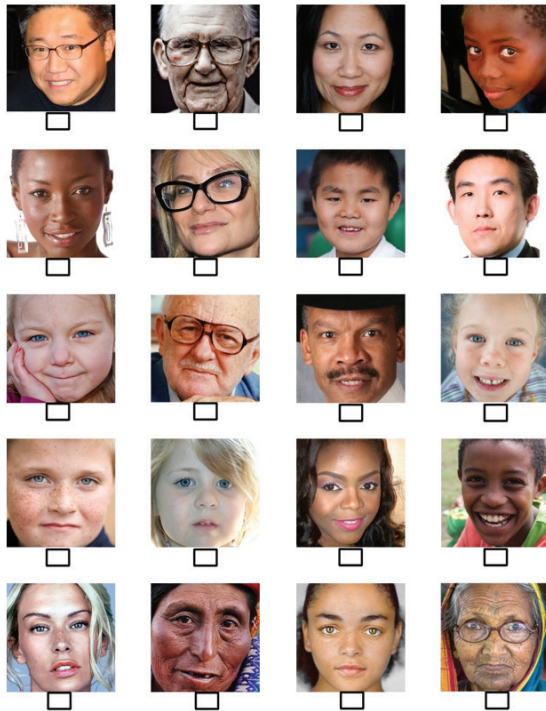
- Fixation time for different tests;
- How many participants returned to the AOI of eyes after they left it (“regression level” or “rate of return”) (Golding-er, He & Papesh, 2009);
- How many participants first looked at the eyes (first fixation);
- What percentage of all time participants spent looking at eyes (face element that attract most attention);
- How many participants looked at the mouth during time of face image presentation and if they did, how long they observed it;
- How many participants looked at the nose during time of face image presentation and if they did, how long they observed it;

Experiment 2

In second experiment we examined the capability of memorizing faces by test participants. The aim was to identify correct remembrance of faces in correlation with time of face presentation (Leyk et al., 2008).

Memorizing faces is very important in many fields of face identification (as an eyewitness, at border control, teachers in schools) (Divyarajsinh & Brijesh, 2013). In doing so, we tried to figure out the difference between the results of the memorizing faces, depending on the time display. Procedure was the same as in experiment 1. Participants were given presentation of ten faces

which were changing every one second (“1 second” test), two seconds (“2 second” test) and four seconds (“4 second” test).



» **Figure 3:** Sample of sheet of faces

At the end of the all presentations participants got a sheet of twenty faces. Ten were shown in the test and other ten were not in the test. Participants were required to mark those which they thought that they had seen in the test.

Results

Experiment 1

Table 2 shows results for AOI of eyes, mouth and nose for all investigation mentioned before.

Experiment 2

Results of successful memorization of faces are shown in Table 3. We tested 42 participants divided in three groups, who have different order of tests (“1 second”, “2 seconds”, “4 seconds”). Each group had 14 participants, thus 140 presented faces.

By changing order of tests we tried to find out whether results of recognition is worse if the test came later (person concentration drops with time).

Table 2

Results of observation of three main face elements (eyes, mouth and nose)

		Test name		
		“1 second”	“2 seconds”	“4 seconds”
	Average fixation time (ms)	256	281	320
Eyes	Time of observation (s)	0.693	1.116	2.288
	% of time spent observing eyes	69.3%	55.8%	57.2%
	Number of return to the eyes	47	189	314
	Max. number of return to the eyes areas	270	290	340
	Percentage of return (regression level)	17.4%	65.2%	92.4%
	First fixation made on eyes area	232	218	251
	All number of first fixation	270	290	340
	Percentage of eyes area as first fixation	85.9%	75.2%	73.8%
Mouth	Time of observation (s)	0.11	0.34	0.62
	% of time spent observing mouth	11.0%	17.0%	15.5%
	Number of visits to area mouth	106	198	276
	Max. number of visits to the mouth area	270	290	340
	Percentage of visit to the mouth area	39.3%	68.3%	81.2%
Nose	Time of observation (s)	0.15	0.25	0.43
	% of time spent observing nose	15.0%	12.5%	10.8%
	Number of visits to area nose	121	173	267
	Max. number of visits to the nose area	270	290	340
	Percentage of visit to the nose area	44.8%	59.7%	78.5%
	Total time of observings all three areas (s)	0.953	1.706	3.338
	Percentage of observing all three areas (s)	95.3%	85.3%	83.4%

Table 3

Results of face recognition

	All faces within test	140	140	140
Test group A (1, 2, 4)	Recognized	102	123	129
	% success recognition	72.9%	87.9%	92.1%
Test group B (2, 4, 1)	Recognized	92	131	133
	% success recognition	65.7%	93.6%	95.0%
Test group C (4, 1, 2)	Recognized	99	125	134
	% success recognition	70.7%	89.3%	95.7%
	Cumulative % success recognition	69.8%	90.2%	94.3%

Discussion

Experiment 1

First we investigated average fixation time for different time of face presentation. It is clear that test with shorter time of face presentation had much shorter average fixation time. The differences in average fixation time in quite significant (256 ms, 281 ms and 320 ms). These results confirmed that eye quickly adapted more dynamic presentation with shorter fixation time (Dorr et al., 2010).

At the test "4 seconds" fixation time 320 ms was very consistent with normal fixation time for observing faces (Guo et al., 2006). We can say that this longer face presentation is most natural way of viewing faces. As we predicted most important face element (eyes) attracted most attention. From the time of observation area of eyes and the time of face presentation we calculated percentage of time spent at area of eyes. For test "1 second" this percentage was the highest among all tests (69,3 %), because in this test faces changed so rapidly that participants gaze rarely left the area of eyes and visited the area of other parts of the face.

At other two tests participants had enough time for observation other face elements, so the percentage of time observing eyes was much smaller (55,8 % and 57,2 %).

Test "4 seconds" had slightly higher percentage than test "2 seconds", due to the fact that participants more often returned back to the area of eyes after observing other face elements.

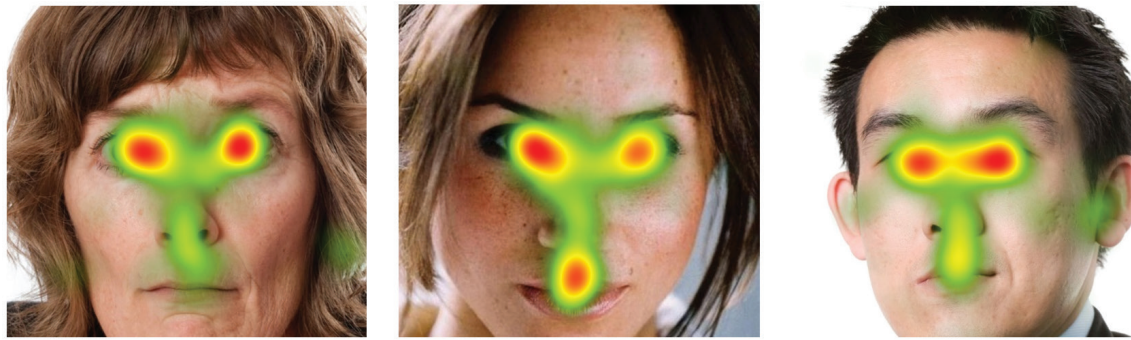
We also investigated "regression level" ("rate of return") for the area of eyes. That is percentage of return to the area of eyes after leaving it to observe other face elements (Goldinger, He & Papesh, 2009). Eye tracker metrics for that information is number of visits to the AOI. If that number is more than 1, test participant had left the area and also returned back one or more times. At longer time of face presentation participants more often left the eyes area, observed other face elements and come back to the eyes.

Results showed that switching faces at one second period was much too fast for participants to perform procedure described above. So regression level was only 17,4 %. Regression level of 92,4 % at "4 seconds" has shown us that four seconds period of face presentation was long enough to observe other parts of the face (also cheeks, chin, forehead, ears and hair) and most of the test subjects have returned to the eyes area. Due to the higher regression level, the percentage of time observing eyes area in the test "4 seconds" was a little higher than in test "2 seconds", because at test "4 seconds" participants more often returned to the area of eyes.

Eyes are usually the first area for the user's gaze. So we also explored eyes area as the first fixation. In all three tests first fixation at most of the faces were made at eyes area. The highest percentage of first fixation on eyes area was at test "1 second" (85,9 %). This can be explained with the fact that at very rapid face change participants gaze didn't even left the area of eyes before the next face image appeared, so at the time of appearance the next face, the gaze was already at that area (despite that we used black screen for 1 second). In other two tests this percentage was lower (75,2 % and 73,8%), but still eyes definitely attracted most attentions.

Comparing results of mouth and nose showed us some interesting aspect. In "1 second" test number of visits to the target areas and time of observation was higher at nose area. One second time change was very fast for participants, so if they even left the area of eyes, their gaze usually did just a small saccade down (area of nose). In general, mouth are more attractive face elements than nose, so in our case, when participants had enough time for observation ("2 seconds" and "4 seconds" test), percentage of visits and specially time observation of those area were in favor of mouth.

Last calculation was about all three tested areas together. As those are three most attractive face elements, participants observe other parts of face only when face presentation was long enough. We already found out that one second was too short and sum percentage of observation of three areas are very high (95,3 %). In other words, less than 5 % of all fixations were made outside of these three main face elements.



» **Figure 4:** Samples of heat maps for different tests

Results for other two tests with longer time of face presentation are very similar. Manual review showed that in “2 seconds” and “4 seconds” tests majority of participants had enough time to also look other face parts of faces (cheek, forehead, ears, etc.)

Figure 4 shows samples of heat maps for one face in every test (“1 second”, “2 seconds” and “4 seconds”). In first face heat map almost all fixations were made in the area of eyes, second face heat map presents that mouth was also strongly observed and at third face heat map we can see that participants had enough time to look also at the ears and cheek. Those are most representative heat maps for different tests.

Experiment 2

As we said participants were divided in to three groups (A, B and C) and order of tests we organized in Latin square. Each set contain 14 participants; thus 140 faces were presented. Table 3 shows results for each test group. Results of correct recognition increased with longer time of face presentation. But the difference between “1 second” test and “2 seconds” test was much bigger then between “2 seconds” test and “4 seconds” test. When we compared same tests in different order results also met our expectation. Recognition results for test “1 second” were the best when this test was performed first (test group A) and worst when it was performed last (test group B). Fall of concentration had some influence to the ability of memorizing. Same pattern can be seen for test “4 seconds”. Test “2 seconds” also had best recognition results when it was performed first. Cumulative results of correct recognition for tests “2 seconds” and “4 seconds” were very good (90,2 % and 94,3 %), whereas test “1 second” had much lower recognition success (69,8 %).

Conclusion

In our study we investigated how people look at the face elements in dependence from time of face presentation. We confirmed that eyes are by far most attractive element. Only if participants had more

time available, their gaze also went to the mouth and nose, whose were next two face elements that attracted attention. Even in our longest test (4 seconds) other face element were rarely observed.

Results has shown that for memorizing faces using short term memory time of face presentation of one second was too short and two seconds was already long enough. Longer face presentation doesn’t improve results in such great level. Further research could investigate how different numbers of presented faces influence on face recognition success.

We can also conclude that from user experience perspective, any presentation on internet should include person image. It is also advisable that person must face the camera.

Important factor in process of face recognition is also emotion of the faces. Authors have slightly different set of basic emotion (happy, angry, sad, fear, surprised, annoyed, etc.). Our future research is to investigate how people see and remember faces with different emotions (face expression) and what are most attractive face elements for different emotion.

References

- Buchan, J.N., Pare, M. & Munhall, K.G. (2007) Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*. 2 (1), 1-13. Available from: doi: 10.1080/17470910601043644 [Accessed 6th September 2014].
- Cangöz, B., Altun, A., Aşkar, P., Baran, Z. & Mazman, S. G. (2013) Examining the visual screening patterns of emotional facial expressions with gender, age and lateralization. *Journal of Eye Movement Research*. 6 (4), 1-15. Available from: doi: 10.16910/jemr.6.4.3 [Accessed 7th September 2014].
- Cassin, B. & Solomon, A. B. S., Melvin, L. R. (ed.) (1990) *Dictionary of Eye Terminology*. 2nd edition. Gainesville, Triad Pub Co.

- Divyarajsinh, N. P. & Brijesh, B. M. (2013) Face Recognition Methods & Applications. *International Journal of Computer Technology and Applications*. 4 (1), 84-86. Available from: <http://www.ijcta.com/documents/volumes/vol4issue1/ijcta2013040114.pdf> [Accessed 7th September 2014].
- Dorr, M., Martinetz, T., Gegenfurtner, K. R. & Barth, E. (2010) Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*. 10 (10), 1-17. Available from: doi: 10.1167/10.10.28 [Accessed 30th August 2014].
- Goldinger, S. D., He, Y. & Papesch, M. H. (2009) Deficits in Cross-Race Face Learning: Insights from Eye Movements and Pupillometry. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 35 (5), 1105-1122. Available from: doi: 10.1037/a0016548 [Accessed 6th September 2014].
- Guo, K., Mahmoodi, S., Robertson, R. G. & Young, M. P. (2006) Longer fixation duration while viewing face images. *Experimental Brain Research*. 171 (1), 91-98. Available from: doi: 10.1007/s00221-005-0248-y [Accessed 5th September 2014].
- Heisz, J. J. & Shore, D. I. (2008) More efficient scanning for familiar faces. *Journal of Vision*. 8 (1), 1-10. Available from: doi: 10.1167/8.1.9 [Accessed 25th August 2014].
- Henderson, J. M., Falk, R. J., Minut, S., Dyer, F. C. & Mahadevan, S. (2001) Gaze control for face learning and recognition in humans and machines. In: Shipley, T. & Kellman, P. (eds.) *From fragments to objects: Segmentation and grouping in vision*. Amsterdam, Elsevier Science, pp. 463-481.
- Henderson, J. M., Williams, C. C. & Falk, R. J. (2005) Eye movements are functional during face learning. *Memory & Cognition*. 33 (1), 98-106. Available from: doi: 10.3758/BF03195300 [Accessed 23th August 2014].
- Hills, P. J. & Pake, J. M. (2013) Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition*. 129 (3), 586-597. Available from: doi: 10.1016/j.cognition.2013.08.012 [Accessed 15th August 2014].
- Hooge, I. T. C. & Erkelens, J. C. (1998) Adjustment of Fixation Duration in Visual Search. *Vision Research*. 38 (9), 1295-1302. Available from: doi: 10.1016/S0042-6989(97)00287-3 [Accessed 28th August 2014].
- Horsley, M., Eliot, M., Knight, B. A. & Reilly, R. (eds.) (2014) *Current Trends in Eye Tracking Research*. Berlin, Springer.
- Klama, E. K. & Milton, F. (2012) Differences in eye movements between same and other race face recognition. In: Miyake, N., Peebles, D. & Cooper, R. P. (eds.) *Building Bridges Across Cognitive Sciences Around the World: Proceedings of the 34th Annual Meeting of the Cognitive Science Society, 1-4 August 2012, Sapporo, Japan*. Austin, Cognitive Science Society. pp. 1804-1809.
- Kyrnin, J. (2014) *What is Web Content?* Available from: <http://webdesign.about.com/od/content/qt/what-is-web-content.htm> [Accessed 23th August 2014].
- Leyk, D., Sievert, A., Heiss, A., Gorges, W., Ridder, D., Alexander, T., Wunderlich, M. & Rütger, T. (2008) Validation of a short-term memory test for the recognition of people and faces. *Ergonomics*. 51 (8), 1125-1136. Available from: doi: 10.1080/00140130802094371 [Accessed 5th September 2014].
- Miner, M. & Park, D. C. (2004) A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments & Computers*. 36 (4), 630-633. Available from: doi: 10.3758/BF03206543 [Accessed 7th September 2014].
- Nielsen, J. & Pernice, K. (2010) *Eye Tracking Web Usability*. Berkeley, New Riders.
- Patel, N. (2014) *8 Powerful Takeaways from Eye Tracking Studies*. Available from: <http://www.quicksprout.com/2014/04/16/8-powerful-takeaways-from-eye-tracking-studies/> [Accessed 6th September 2014].
- Pollatsek, A., Rayner, K. & Balota, D. A. (1986) Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*. 40 (2), 123-130. Available from: doi: 10.3758/BF03208192 [Accessed 20th August 2014].
- Tobii Technology AB. (2012) *User Manual Tobii Studio*. Version 3.2. Danderyd, Tobii Technology AB.