

WHICH ONE DO YOU PREFER AND WHY? THINK ALOUD!

Sami Lahtinen, University of Art and Design Helsinki, Finland

Abstract

How participants compare and evaluate a pair of product pictures with sight? How they verbalize their attitudes and perceived product attributes? This study will illustrate gaze path and spoken language argumentation events, their sequences and interaction.

Participants were asked first to “View as you wish!” a pair of mobile phone pictures, then they were asked to view phones again with a task “Which one do you prefer and why? Think aloud!”. These two tasks were repeated with other pair of phones. Finally all four phones were shown together in one screen and subjects were asked to evaluate them aloud in scale 1 – 10. Their gaze paths and think aloud argumentations were recorded.

Spoken argumentations of subjects were transcribed and transcriptions were organized by segments of verbal idea units i.e. intonation units. Idea unit-coded transcriptions were time coded and then time synchronized with successive fixation segmented gaze path.

Gaze paths were analyzed with successive fixations analysis method and total fixation duration and count on left and right phone. Successive fixations analysis on the left and right phone picture reveal how subjects tend to fixate successively certain parts of one phone and then switch gaze at other phone to be able to compare earlier fixated information from first phone and so on. This gaze transition between phones can be seen as looping exploration and examination process. These successive fixations formed perceived product picture based attributes which were compared.

Visualized evaluation event examples of successive fixation units on the left and right phone with time synchronized transcribed verbal units will be presented in paper to support the findings.

Keywords: Design evaluation, visual perception, eye tracking, think aloud protocol, visual focus and verbal focus.

Introduction

Motivation of study

This research is done in a larger research project -p-o-d- (Perception of Design), which aims to study how people perceive design products.

The motivation of this study was to understand how people evaluate products and their visual appearance. How subjects compare and evaluate a pair of product pictures with sight? How they verbalize their attitudes and perceived product attributes? What attributes they are evaluating?

Other aim was to develop test and analysis methods for design proposal and product evaluation.

Eye tracking and think aloud protocol test

Participants were asked first to “View as you wish!” a pair of mobile phone pictures, then they were asked to view phones again with a task “Which one do you prefer and why? Think aloud!”. These two tasks were repeated with other pair of phones. Finally all four phones were shown together in one screen and subjects were asked to evaluate them aloud in scale 1 – 10. Their gaze paths and think aloud argumentations were recorded.

Test screen sequences and times were following: for the “View as you wish” screens viewing times were 10 seconds, for the “Which one do you prefer and why?” screens viewing times were self-determined and for “Evaluate products A, B, C & D on your own preference in scale 1 – 10. Say your evaluation aloud!” screens viewing were self-determined.

Participants and test operator were interacting verbally in minimal manner. Operator mainly answered questions of participants and accompanied verbalization of participants.

The gaze paths were recorded with a remote Tobii eye tracker and the verbalized evaluation protocols were recorded with video camera.

We had 32 participants in our test. 8 participants were left out from the analysis because of eye tracker calibration problems or because of gaze bath drifting or because of mistakes in the test protocol. Out of the 24 analyzed participants, 12 participants had design education background (5 females and 7 males) and 12 participants had some other education background (7 females and 5 males). In the following, I will speak about consumers and designers (groups).

Visual and verbal attention foci

Gaze path data alone did not show direct connection to preference and attitude to phones with fixation duration and count metrics. Synchronized gaze paths segmented to successive fixation units on the left and right phones and transcripts segmented to verbal idea units revealed some temporal connections between visual foci and verbal foci.

Successive fixations on one product can be seen as visual attention focus (visual focus) of participant when he is evaluating and comparing a product pair. Some of these visual foci i.e. perceived attributes were verbalized by participants and some were not. Participants verbalized also attitudes and attributes which based on their earlier experiences and personal values.

Participants looked phone picture or area of interest or certain details because their experience and value based attitude to these are positive, neutral or negative. Or viewed phone picture or area of interest or certain details generated admiring, insignificant or critical gaze in other words perceived attributes had an effect on participant’s attitude to product.

The nature of this paper is descriptive. I describe how and what participants compared and evaluated phone picture pairs with gaze. Another fascinating point is how participants verbalized their evaluation and what attributes they argued aloud.

Results and conclusions of this paper are preliminary, because of the experimental nature of this study and because some analyses are still on progress.

Analysis and results of verbalized evaluation attributes and attitude

The question “Which one do you prefer and why?” measure mainly attitude defined by Fishbein and Ajzen (1975, 6) as “a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object”. Only few participants expressed preference by selecting one phone rather than another for them self in their verbalization. Common expressions were like: “I like more the left one () because” or “Right one () because”.

Verbalizations of participants are mainly Think Aloud, level 2 verbalizations. Level 2 verbalizations of participants are those that must be transformed from images to words before being verbalized in short-term memory. Participants vocalized also level 3 verbalizations when their evaluation based on their earlier experiences. These earlier experiences are retrieved from long-term memory. (Ericsson & Simon 1984, 15-20; Boren & Ramey 2000, 262). I consider both level 2 and 3 verbalizations valid and reliable when studying how people evaluate products.

The question “which one do you prefer and why?” do not suggest any certain evaluation attributes. Attributes that participant vocalized reflect their own needs, desires and expectations. So the assumption is that participants evaluated phone pictures according to their own personal attitude and attribute expectations.

What kind of evaluation attribute and verbalization profiles revealed from verbalizations? I present some examples below.

Why some participants favored NOKIA 6630? How they argued their attitude and opinion?

p18: “And I like this right one ((NOKIA 6630)) ()
just because of practical reasons that
it has bigger ((display))
its functionality then
that pleases me more ()
and then the colour is nicer than that other one I don’t like black.”

Why some participants disfavored NOKIA 6630? How they argued their attitude and opinion?

p3: “Well it must be this left one ((NOKIA 6230)) ()
that right one looks like ()
bit bigger and looks like kind of bit stupid and it has awful big display ()
small keys
that left one appears more practical.”

Why some participants favored NOKIA 6230? How they argued their attitude and opinion?

p5: “Yes this left ((NOKIA 6230))
simply because of the reason that it probably fits better to pocket than that right
one block ()
although colours should just be other way round ()
in the left silver and in the right dark
but anyway the left is better.”

Why some participants disfavored NOKIA 6230? How they argued their attitude and opinion?

p7: Puuhh ()
well the left one is somehow nicer than the right ((NOKIA 6230)) ()
is too ()
sharp ()
in there below ()
oh well it depends on which appearance you have used to ()
I don't like either the keys of left one but ()
its sharpness doesn't jump out so bad ()"

Which phones did they favour?

Out of **24** participants **13** favored **NOKIA 6630** and **11** favored **NOKIA 6230**. In case of the second pair **12** participants favored **NOKIA 3220** and **11** favored **NOKIA 5140**. One consumer evaluated NOKIA 3220 and NOKIA 5140 as equal.

5 consumers favored **NOKIA 6630** and **7** consumers **NOKIA 6230**. **NOKIA 3220** were favored by **6** consumers and **NOKIA 5140** were favored by **5** consumers. One consumer could not tell which one he favored NOKIA 3220 or NOKIA 5140.

From the **designers 8** favored **NOKIA 6630** and **4** **NOKIA 6230**. **Both NOKIA 3220 and NOKIA 5140** were liked more by **6** designers.

In the last evaluation screen participants were asked to evaluate all four phones in scale 1 – 10. Average grades were following: NOKIA 6630 **6,29** and NOKIA 6230 **6,27** and for the second phone pair NOKIA 3220 **6,21** and NOKIA 5140 **6,04**.

Analysis and results of verbal foci of product evaluation

Spoken argumentations of participants were transcribed and transcriptions were organized by segments of (verbal) idea units (Holsanova 1999, 14) i.e. intonation units (Chafe 1980). Idea unit is the mental unit that is the object of investigation. (Holsanova 1999) An intonation unit is a sequence of words combined under single, coherent intonation contour, usually preceded by a pause. (Chafe 1987, 1996)

Idea unit-coded transcriptions were time coded and then time synchronized with successive fixation segmented gaze path.

In the average consumers (**average 25243** ms) used more time to evaluate the product pairs than designers (average **23170** ms). The longest evaluation duration was **84416** ms when consumer p10 evaluated NOKIA 6630 and NOKIA 6230 and shortest evaluation duration **7414** ms was when designer p19 evaluated NOKIA 5140 and NOKIA 3220.

One can see from the transcriptions that there are two extremes how participants verbalized their evaluations on product pairs: short-spoken ones and verbose ones.

Consumer p12 verbalized only one attribute about first product pair: "Yes yes I () like that right one more () it look like more modern" and for the second phone pair he did not express his

opinion about the phones: “Well these I don’t have () any opinion () if I should select a mobile either one would do”.

And designer p19 just uttered about NOKIA 6230 and NOKIA 6630: “Left one () because it is considerably more beautiful” and about NOKIA 5140 and NOKIA 3220: “Right one () because it is not nearly as ugly as that left one”.

The most verbose one consumer p10 verbalized 8 attributes for first phone pair and for the second pair 12 and the most verbose designer p20 verbalized 9 and 6 attributes for the phone pairs.

Following transcription example of consumer p10 presents how successive fixation units on the left and right phone and verbal evaluation idea units are temporally connected and how they interact. Bolded successive fixations units like “**6 SF**” mark that consumer p10 is speaking during those successive fixations units.

“2 SF NOKIA 6230” indicates that participant is looking at NOKIA 6230 and

“3 SF NOKIA 6630” indicates that participant is looking at NOKIA 6230 and so on.

Successive fixations units / verbal idea units of p10 during “Which on do you prefer and Why” NOKIA 6630 (left) and NOKIA 6230 (right)

1 SF		o: Okay now you can tell
2 SF	NOKIA 6230	
3 SF	NOKIA 6630	
4 SF	NOKIA 6230	
5 SF	NOKIA 6630	
6 SF	NOKIA 6230	p10: Somehow I’m more used to those normal
7 SF	NOKIA 6630	
8 SF	NOKIA 6230	
9 SF	NOKIA 6630	
10 SF	NOKIA 6230	() rectangle so those round shapes bit disturbs me
11 SF	NOKIA 6630	() although I haven’t used one so it could be quit handy but
12 SF	NOKIA 6230	
13 SF	NOKIA 6630	
14 SF	NOKIA 6230	() I like more those straight models
15 SF	NOKIA 6630	()
16 SF	NOKIA 6230	
17 SF	NOKIA 6630	
18 SF	NOKIA 6230	
19 SF	NOKIA 6630	
20 SF	NOKIA 6230	then ()
21 SF	NOKIA 6630	I can’t right now see where that power-button is and how easy it is to use
22 SF	NOKIA 6230	
23 SF	NOKIA 6630	() if it is there associated with C-button in that it is at least not
24 SF	NOKIA 6230	()
25 SF	NOKIA 6630	in that rounder model it is not at least in C-button which is a good thing () ehm ()
26 SF	NOKIA 6230	
27 SF	NOKIA 6630	then ()
28 SF	NOKIA 6230	both have navigation-button that I have used to use which is a good thing
29 SF	NOKIA 6630	
30 SF	NOKIA 6230	
31 SF	NOKIA 6630	()
32 SF	NOKIA 6230	
33 SF	NOKIA 6630	although that one has some other functions in there that rounder model ()
34 SF	NOKIA 6230	
35 SF	NOKIA 6630	and display is indeed bigger in that other which would be quit good
36 SF	NOKIA 6230	
37 SF	NOKIA 6630	

38 SF	NOKIA 6230	()
39 SF	NOKIA 6630	if you sometime play games so ()
40 SF	NOKIA 6230	that is noticeable difference
41 SF		()
42 SF	NOKIA 6630	
43 SF	NOKIA 6230	
44 SF	NOKIA 6630	
45 SF	NOKIA 6230	NOKIA stands out clearer and somehow nicely in that black phone
46 SF	NOKIA 6630	
47 SF	NOKIA 6230	()
48 SF	NOKIA 6630	
49 SF	NOKIA 6230	
50 SF	NOKIA 6630	then ()
51 SF	NOKIA 6230	
52 SF	NOKIA 6630	
53 SF	NOKIA 6230	
54 SF	NOKIA 6630	
55 SF	NOKIA 6230	it is difficult to say indeed without testing how those feels in hand but I based on these pictures I would select for myself that black
56 SF	NOKIA 6630	
57 SF	NOKIA 6230	()
58 SF	NOKIA 6630	model

Analysis of cumulative fixations on the left and right phone and results

Gaze path is formed from sequences of fixations and saccades. Sequence when the eye steadily views a single gaze position is called fixation and saccades are rapid motions of the eye from one position (fixation) to another. Vision is possible only during fixations. The foveal zone (1,5 – 2,5 degrees) is the area of clear vision (the size of the fixation circles in pictures in this paper).

Total fixation duration and count on the left and right phone and favored phone

Analysis of individuals total fixation duration and count on the left and right phone did not reveal correspondence between the favored phone and gaze data. Individual's cumulative fixation heat maps on the left and right phone showed only what areas of interest participants fixated and in what density. The total fixation duration and count revealed how the total gaze spending time on left and right phone have divided. Participants spent longer durations gazing on both their favorite and unfavorite phones.

Statistical analysis of gaze data and results

Average fixation duration difference between consumers and designers during the tasks

One viewing difference between consumers and designers was found in the statistical analysis of the gaze data. The average fixation durations of designers (average **261,3** ms) were higher than the average fixation durations of consumers (average **235,0** ms) in all five tasks. Miall and Tchalenko (2001, 38) have observed in their studies that artists average fixation duration was higher than novices during drawing of a sketch of photographed faces.

It seems that one explanation could be that designers have learned to look differently during their education i.e. they have learned to look differently while practicing drawing. There can be other explanations.

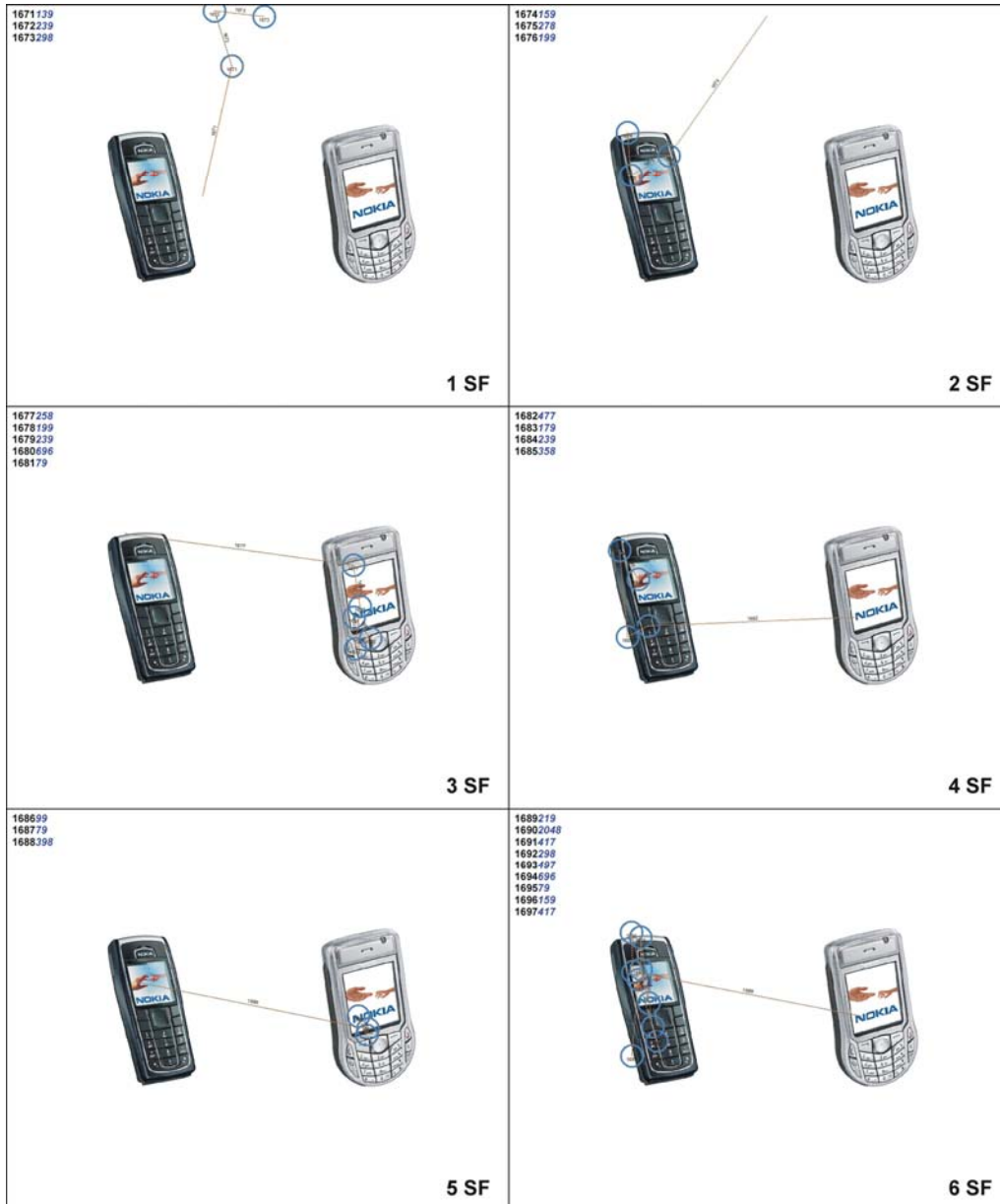
This average fixation duration difference can be interpreted so that generally designers were examining phones and certain areas of interest in more specific learned manner by gazing them with higher fixation durations. This assumption is preliminary. Examination and interpretation this average fixation duration difference is still on progress.

Analysis and results of visual foci of product evaluation

Successive fixations on the left and right product

Usually participants started their visual observation during the “View as you wish!” task with quick exploration on the both phones. They seemed to first fixate in the space between the phones or to fixate the left or right one with few fixations and then the other one with few fixations to identify what they are evaluating (See figure 1, 1 SF, 2 SF & 3 SF). After identification they continued gazing between the phones with a few more fixations on both phones. These comparative perceptions i.e. visual foci between phones continued so that the amount successive fixations on phones seems to increase. The increase of fixation amount varied (See figure 1, 4 SF, 5 SF & 6 SF). Participants also re-fixated certain areas of the phones to re-examine them. In figure 1 you can see the successive fixations on the left and right phone of participant p20 during “View as you wish!” task.

FIGURE 1: “View as you wish!” task, designer p20, all successive fixations sequences on the left and right phone.



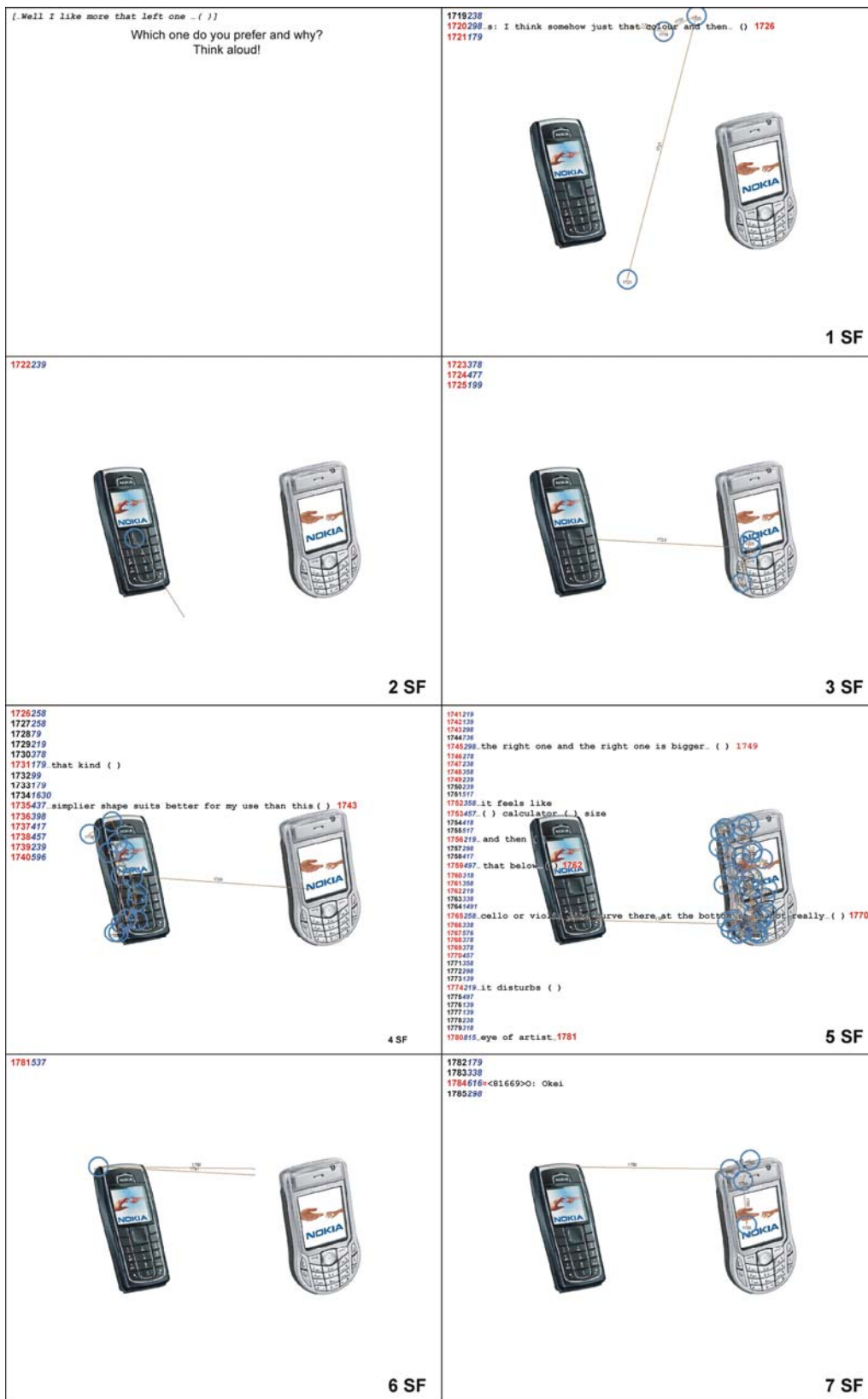


FIGURE 2: “Which one do you prefer and why?” task, designer p20, first 7 successive fixations sequences on the left & right phone.

During the “Which one do you prefer and why?” task these same perception phases can be recognized also (See figure 2).

In the later parts of this (“Which one do you prefer and why?”) evaluating perception some participants started to examine one or the other phone with a long duration sequence of successive fixations on single phone. The verbal idea units of participants synchronized with these long duration sequences of successive fixations on certain phone reveal that participants are then examining and evaluating a single phone and its attributes (See figure 2, 5 SF). The attention of participants is focused on attributes of a single phone and they are not comparing the phone pair during this sequence. Then again participants moved their gaze and attention to other phone and continued comparative evaluation.

These phases of exploring and evaluating perception of a phone pair could be called identification, exploration and specific examination phases i.e. visual attention focused from general to specific in reiterative way. This result is in line with studies of Yarbus (1967) and Holsanova (1999).

Conclusions and future work

Total fixation duration and count on the left and right phone and favored phone

The individual total fixation duration and count on the left and right phone revealed how the total gaze spending time on left and right phone have divided. Participants spent longer durations gazing on both their favorite and unfavorable phones.

Average fixation duration difference between consumers and designers during the tasks

The average fixation durations of designers were higher than the average fixation durations of consumers in all five tasks. Further examination and interpretation need to be done so that we can understand the nature of average fixation duration difference between consumers and designers.

Visual foci and verbal foci of product evaluation

Successive fixations on the left and right phone reveal how and what participants compare and evaluate with their gaze. This visual attention transitions between the phones and their visual contents do not alone tell us why and what participants are exploring and comparing these visual areas of interest or what attributes they are evaluating.

Transcribed and to idea units segmented evaluative verbalizations (verbal foci) synchronized with successive fixations (visual foci) on the phones seems to reveal partly what attributes participants are comparing and evaluating with their gaze. On other words the verbal focus of attention sets some light why participants are looking at certain areas of interest on the phones. Verbal foci i.e. verbalized idea units can not always be connected to visual foci (successive fixation units on the left and right phone). Some successive fixation units on the left and right phone can not be connected to verbalized idea units, mainly because participants continued gazing during the pauses of their verbalization.

There is temporal synchrony between visual and verbal foci in product picture pair evaluation in individuals. Verbal focus follows visual focus with a delay. Participants fixated first, thought (verbally encoded) and verbalized later.

To identify and describe different types of temporal relations (configurations) between visual foci and verbal foci during visual product picture evaluation and verbal argumentation demands more examination in the future. Are there different perception strategies when people compare and evaluate product picture pairs similar to the narrow, the holistic and the combined first impression perception strategies reported in Koivunen et al. (2004).

Test and analysis method refinement

The eye movement and think aloud protocol test and analysis method presented here needs to be refined, because it is very laborious and time consuming to analyze what attribute participants are evaluating with their gaze during certain sequences. The strength of this test method is that the open task question “which one do you prefer and why?” reveals through verbalization which attribute expectations phones raise and what attributes participants evaluate.

One other way to conduct a product or product proposal evaluation test is to ask participants to look at a product proposal pair and evaluate one beforehand defined attribute at the time. In this way one can easier detect the connection between attribute and what product details or areas of interest they are comparing and examining with their gaze. One task question could be for example “Which one is more ergonomic and why?” We conducted a product evaluation test with beforehand defined evaluation attributes in spring 2005. This test will be reported in the future.

After refinement *the eye movement and think aloud protocol test and analysis method* described here can be used during product development processes, for example by asking future users to evaluate and compare different design proposals. Successive fixations on design proposals analysis will show how and what visual details and areas of interest subjects compared and evaluate. And think aloud protocol analysis will tell what perceived attributes design proposal pictures reveal. Attributes of subjects based on earlier experiences and personal values can show us how the design proposals meet the raised expectations of subjects.

Acknowledgements

I would like thank my fellow researchers Kimmo Koivunen, Sami Kukkonen, Harri Rantala, Toni Ryynänen and Selina Sharmin for fruitful co-operation and also test participants for their helpfulness.

References:

- Boren, M. T. & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice, IEEE Transactions on Professional Communication, vol. 43, no. 3, 261-278.
- Chafe, W.L. (1980). The deployment of consciousness in the production of a narrative. In: W.L. Chafe (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Ablex: Norwood, NJ, 9-50.
- Chafe, W.L. (1987). Cognitive Constraints on Information flow. In: Tomlin, Russel S. (ed.), *Coherence and Grounding in Discourse*. Benjamins: Amsterdam/Philadelphia, 21-51.

- Chafe, W.L. (1996). How consciousness shapes language. *Pragmatics & Cognition* Vol. 4(1), 1996. Special issue on language and consciousness, 25-54.
- Ericsson, K. A. & Simon, H. A. (1984). *Protocol Analysis. Verbal reports as data*, Cambridge Massachusets, 15-20.
- Fishbein, M. and Ajzen, I, (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Addison-Wesley, Reading, MA.
- Holsanova, J. (2001). *Picture Viewing and Picture Description: Two Windows on the Mind*. Doctoral dissertation. Lund University Cognitive Studies 83.
- Koivunen, K., Kukkonen, S., Lahtinen, S., Rantala, H. and Sharmin S. (2004). *Towards Deeper Understanding of How People Perceive Design in Products*, Proceedings of the Computer in Art and Design Education (CADE), Copenhagen, Denmark.
- Miall, R. C. and Tchalenko, J. S. (2001). *The Painter's Eye Movements*. *Leonardo*, 34 (1), 35-40.
- Yarbus, A. L. (1967). *Eye movements and vision*. Plenum Press.