An Eye Tracking Technique Towards Form Reasoning in Design Practices.

Kuo-Jung Hsieh

Chang Gung Chi

Eleanor Jayne Browne

University of London

Question – how do individuals perceive visual forms in the indication of daily life interactions? The recording of eye movements during task operation can indicate levels of cognitive behaviours that have been considerably difficult to detect in the past. Fixation and saccades data which display 'intention shift' and 'motor system drive' illustrate how individuals interact with visual stimuli. Two experiments were coordinated to examine the possibility of visual form reasoning investigation in design applications. The broad results indicated that successful communication and decision-making were very likely comprised of three major elements – 'stimulus presentation', 'context' and 'personal preference'. However, even if each element could be further developed for the achievement of efficient communication it is still not a guarantee for successful visual search. Instead, it is more decisive to consider a broader picture of interaction for such vision search behaviours – the effective relationship between the three elements. Humans have the ability to work in imperfect conditions, to skip untruths, to create shortcuts, to optimise their judgments and find alternative solutions. Human information processing does not shut down its actions because of any occurrence untruths. The strength of stimuli, the effectiveness of context and the possibility of user's diverseness can be measured accordingly for their impacts on the interaction result. The extensive question revealed from this study is "how important are the impacts of such indefinite relations in respect to human-product interactions?" and "what are the limitations of adoptability by users?" Also gleaned was that the successful conjunction between each element is crucial and could possibly be another field in which to investigate human-product interaction. This pilot study method reflected a cognitive approach towards multi-discipline design study and more research attentions must be applied before such a method can be implemented into design practice and design research communities.

Title: An Eye Tracking Technique Towards Form Reasoning in Design Practices

Abstract

Question - how do individuals perceive visual forms in the indication of daily life interactions? The recording of eye movements during task operation can indicate levels of cognitive behaviours that have been considerably difficult to detect in the past. Fixation and saccades data which display 'intention shift' and 'motor system drive' illustrate how individuals interact with visual stimuli. Two experiments were coordinated to examine the possibility of visual form reasoning investigation in design applications. The broad results indicated that successful communication and decision-making were very likely comprised of three major elements – 'stimulus presentation', 'context' and 'personal preference'. However, even if each element could be further developed for the achievement of efficient communication it is still not a guarantee for successful visual search. Instead, it is more decisive to consider a broader picture of interaction for such vision search behaviours – the effective relationship between the three elements. Humans have the ability to work in imperfect conditions, to skip untruths, to create shortcuts, to optimise their judgments and find alternative solutions. Human information processing does not shut down its actions because of any occurrence untruths. The strength of stimuli, the effectiveness of context and the possibility of user's diverseness can be measured accordingly for their impacts on the interaction result. The extensive question revealed from this study is " how important are the impacts of such indefinite relations in respect to human-product interactions?" and "what are the limitations of adoptability by users?" Also gleaned was that the successful conjunction between each element is crucial and could possibly be another field in which to investigate human-product interaction. This pilot study method reflected a cognitive approach towards multi-discipline design study and more research attentions must be applied before such a method can be implemented into design practice and design research communities.

Keywords: eye movements, cognitive behaviour, stimulus presentation, context, person preference, human-produce interaction

1.0 Introduction – Visual Reasoning

The main explanation for an integrated design method from which information processes from user states and designer states, is the lack of visual reasoning methods in connection with error handling within Human-Machine interaction while visual resources are considerably important within the whole information processing. Another fact stems from current HEI or ergonomics design which supports methods favouring the systematic approach or the task-orientated concept. Unfortunately the visual reasoning method, which potentially directly relates to design manipulation, has received little attention. Therefore, to propose the visual reasoning method, which contains strong cognitive evidence, is to enrich Human-Machine interaction in the broadest sense and to enhance applicable design methods.

The consistence of our environment acts as the information background and provides the context of understanding of any experienced event. (Minsky, 1975) Even during a naive event, the background information will be able to provide us with a close guess. Detecting differences from background information is to reason large visual information into an efficient input - to detect useful information. The perceiver can draw from his already stored knowledge; with minor alternation according to the useful information, the new schema can be reconstructed with only little effort. Rumelhart, thinking "*about data structures to represent generic concepts stored in memor"* (1975), and Minsky's concepts of frame and node (1975), explained a similar way of efficient organisation of our knowledge. It is assumed that visual information is processed in the same way. For more sophisticated and artificial situations, the acumination of new experiences is very much based on the old ones. To react in a new situation means firstly, the successful detection of different information and, secondly an effective match with old knowledge. Finally, 'goal orientation attention driven' is considered to be the major doctrine of human visual perception during daily life activities.

Attention serves as the valve for adjusting workload especially when there are limited mental resources to deal with large amounts of information. From investigating the relationships between eye and hand movements in preparing tea and sandwiches, as a simple daily activity, Land and Hayhoe (2001) pointed out the oculomotor system provided guidance, locating, directing, guiding and checking, for hand operation, and that eye movement was essentially task-relevant, presenting top-down scan behaviour. Visual perception here acts as an intended activity to confirm and secure one's operations, not merely a random or thoughtful system search. Even though visual perception is able to access and decode a complex task into a set of sub-tasks, Pelz and Canosa (2001) suggested the high-order perceptual strategy in the complex task, while major eye movements cope with immediate action some eye movement effort involves planning and organising how to look. The eye movement is a projection of the perceiver's mental activity and his/her plan. Such inference is based on eye movement driven by attention shift, and this to a certain degree represents the way we think.

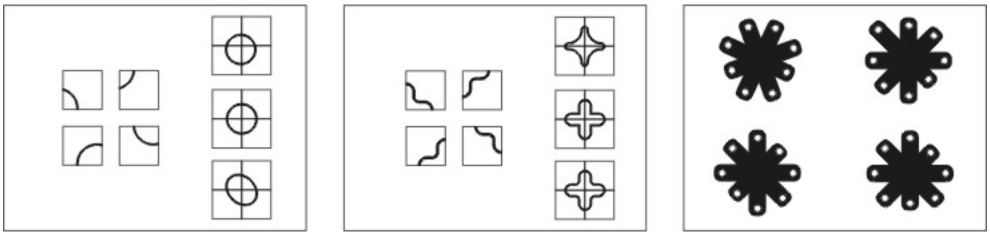
There are good reasons to explain the way we perceive the world and how visual perception interlinks with other inner mental activities. What we are looking at is rather a series of strategies incorporated with organism function rather than a random choice. As in Brunswick's assertion (1956) in probabilistic functionalism, the intuitive perceiver seeks for the most useful information for survival. Visual perception, in this respect, is more than just collecting information but takes its 'initiative' to support its master.

2.0 Case Study - 'Jigsaw Puzzle'& 'Similarity mapping'

The nature of this case study is to establish the experiment protocol for reasoning visual forms in design applications. The eye movement detecting method (EyeLink II system) has been adopted for this study and positive insights for reasoning visual forms are explored. Jigsaw puzzle and similarity matching tests were designed as a joint trial.

In the jigsaw puzzle trial separated images on the left hand side are puzzle units and three different completed images, only one corresponding to the left, was presented on the right hand side. (see figure 1) Each separate puzzle unit was rotated 90 degrees to increase mapping task difficulty. The focal aim for this trial was to observe how participants worked out the jigsaw puzzle and it was assumed that there might be many ways to achieve this. The possibility of such visual reasoning behaviours may provide relevant hints for 'design for error & alternative usage' in product design applications.

The third section, 'similarity matching' trial, presented four images on each corner. Two identical images were located on the upper left and lower right corners at differently rotated angles. The remaining two similar images, with different arm lengths, were located on the upper right and the lower left corners. This similarity trial's aim was not only a means of observing comparison behaviours but also a means of identifying other elementary factors (for reasoning forms), such as length, position or structure. The white spot on each image served as an additional visual preference.



Section 1 (Puzzle I) Section 2 (Puzzle II) Section 3 (Similarity)

Figure 1: Jigsaw puzzle & Similarity matching

2.1 The Trial Process and Participant Description

7 design students, aged between 22 and 25 years of age were selected, comprising three male subjects and four female. All were final year students from the BA Industrial Design course, therefore it can be surmised that subjects may have had greater sensitivity with regards to visual shape. The duration of this process from calibrating the trial process to collecting the trial data was approximately 15 minutes per participant. Each trial section was set up every 20 seconds, so individual data collection time was 60 seconds. A pre-trial was shown to the participant during the trial brief. This inclusion aimed to prevent attention loss and to indicate what participants should look at.

2.2 Results and Discussion

In P1 trial (Jigsaw Puzzle trial section 1), 4 participants offered the correct answer (57% accuracy), and in P2 (Jigsaw Puzzle trial section 2), all gave the correct answer (100% accuracy). In S3 (Similarity matching trial section 3) no correct answers were given (0%). From the trial sample shown in figure 1, this high level of accuracy was obviously due significantly to P2's curved characteristics. While P1 presented similar curves, participants required complex comparisons of various subtle resources such as proportion and size. When judgment relied on more than one source, then one may have impacted on the other thus resulting in bad judgment. According to resource theory (Norman & Bobrow, 1975; Moray, 1969), overloaded working memory has essential influences on a subject's performance by abandoning least priority information. There is another significant saccade behaviour difference between P1 and P2 trials - during P2, participants seemed to perform reductive observation. Once they picked up the clue from the left side, most saccades and fixations were located on the incorrect answers as opposed to the correct ones. This meant that the shape of the curve had been picked as a strong reference, and might be explained as the comparison target being registered as a sort of knowledge schema, visual-schema. Their response time was quick, ranging from 4-8 seconds. However, such eye movement behaviours did not happen in P1. In P1, participants seemed to perform deductive observations from which their eye movement shifted between each image. Their responses were due to a lack of confidence and the response time was slow, ranging between 10-16 seconds.

Sub.	Sex	P1	P2	P1,2 Note	S3	S3 Note
1	M	3 (false)	3 (true)		3-4 (false)	
2	M	1 (true)	3 (true)	Curve	3-4 (false)	Length
3	F	1 (true)	3 (true)	Curve	1-2 (false)	Length
4	F	2 (false)	3 (true)	Segment	3-4 (false)	
5	F	1 (true)	3 (true)	Proportion	2-4 (false)	
6	F	2 (false)	3 (true)		2-3 (false)	
7	M	1 (true)	3 (true)		1-2 (false)	Length
Result		4/7	7/7		0/7	

Note: 1. TG trial & TB trial: V – perform; / - not perform

2. P1: Puzzle trial section 1; the verbal responses are shown

3. P2: Puzzle trial section 2; the verbal responses are shown

4. S3: Similarity trial section 3; the verbal responses are shown

Figure 2: Trial Summary

The most surprising aspect was the S3 trial – 0% accuracy. This was possibly due to task complexity and time pressure. A number of participants were not able to complete the whole comparison before the trial concluded. If they could neither select unique elements nor elucidate any structural rules, they had to perform a more complex ‘pair comparison’. There were 6 ‘pair comparisons’ to complete in 20 seconds which was very doubtful to achieve. Of those who did conduct comparisons, they showed preferences for horizontal and vertical pair comparisons rather than diagonal pair ones. For the reason that the correct answer was a 1- 4 match of diagonal pairs, participants may not have even performed this pairing before running out the time. However, most participants relied heavily on length difference as the main clue.

Despite the rough conclusions drawn from observation and questionnaires, it is more essential to reveal how subjects worked out their answers from fixations and saccades. To analyse any particular subject eye movement behaviour is easier than to conclude a total review. Although two exciting techniques for testing Scanpath Theory (Noton and Sterk, 1971) have been conducted recently, Markov process (Pieters etc., 1999) and String-Editing analysis (Brandt and Stark, 1997; Josephson and Holmes, 2002) the differing aims of this experiment are to provide clear outlines in understanding how visual intentions shift from one to another and, what kind of visual elements are a subject's concern. Hence, the following analysis is based on deducing what could be the important aspects of eye movements in relation to visual targets.

2.2.1 Jigsaw Puzzle Trial Section 1

Average response time ranged from 10-16 seconds and 57% participants gave the correct answer. (see figure 3) According to fixation count, subjects placed equal attention to the left and right hand sides. Average fixations were 17.8 on the left and 16.1 on the right side. Similarity matching appeared to be the primary effort; nevertheless, subjects sought additional clues such as space, proportion or scale. This was especially the case within this trial because all curves shared similar characteristics. The third target on the right side received few fixations from each participant, except subject 1. It is assumed that the oval shape provided a stronger clue for deciding either a true or false match, therefore there was no need to give it much attention, in other words – less fixation points. Fixation can only represent a subject's attention and does not necessarily correspond to their answer. However, in this trial, all subjects showed the correspondence of the high fixation target as their answers. The second assumption, relating to most of the study about attention and fixation, is that there was a probable direct link with fixation and decision-making.

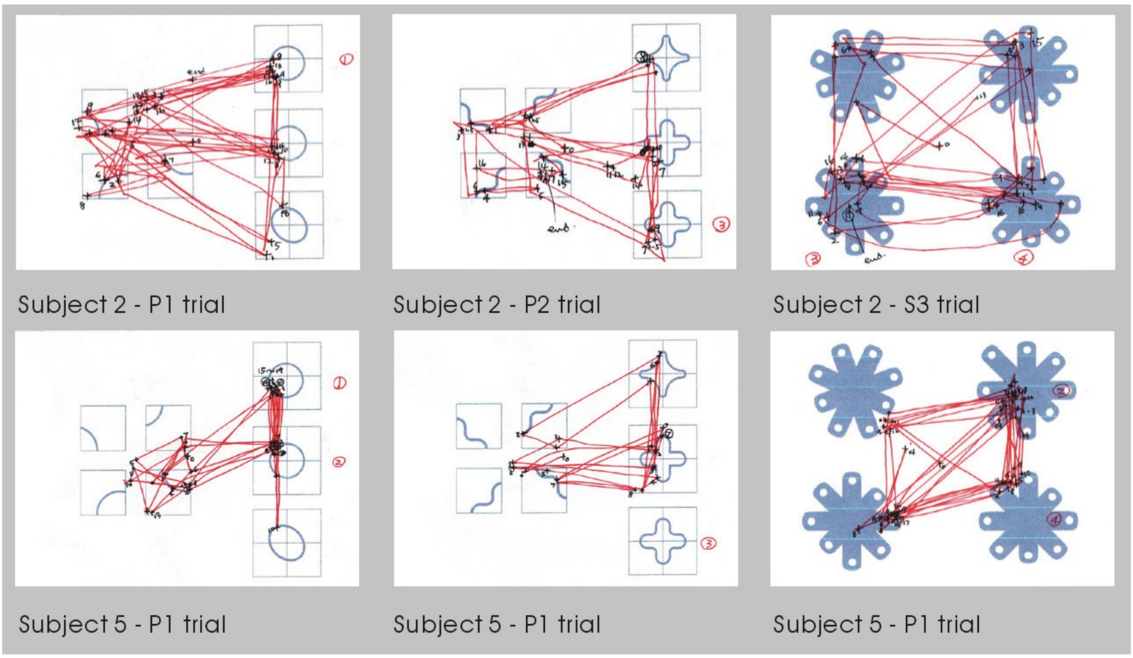
	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Average time / Correct rate
P1 Trial	(15)	16	16	(7)	14	(10)	15	13.3 sec. / 57%
P2 Trial	4	8	8	6	7	5	5	6.1 sec. / 100%
S3 Trial	(7)	(16)	(20)	(20)	(20)	(20)	(20)	17.6 sec. / 0%

Total trial time was 20 seconds. Numbers below each subject indicate their trial time response.
( ) indicates that subjects gave the false answer.

Figure 3: Trial result and response time.

2.2.2 Jigsaw Puzzle Trial Section 2

Average response time was 6.1seconds - much shorter then the previous trial and there was a 100% correct response rate. Obviously, the curved image with its stronger characteristics was more likely to register in working memory for association with target images. The average fixation count on the left hand side (sample images) reduced to 8.6, while the one on the right (target images) was 12. The reduction of attention to the left hand side image was significant even though a few participants did not scan all of the images before supplying their answers. In addition, there was one major difference - the high fixation target was not necessarily for the correct answer. In fact, target image 3 received 0 fixation out of 3 trials conducted. This is paradoxical as judging from previous results there is a correspondence of attention (fixation) and the subject's answer. It is assumed that the subject would apply different mapping strategies according to any given condition or changed context. The adopted strategy might aim to reduce mental load and perform better outcomes. (see figure 4)



Subject 2 - P1 trial | Subject 2 - P2 trial | Subject 2 - S3 trial

Subject 5 - P1 trial | Subject 5 - P1 trial | Subject 5 - P1 trial

Figure 4: Eye tracking data from subjects 2 and 5

2.2.3 Similarity Matching Trial Section 3

Average response time was 17.6 seconds, and most participants gave their answer after the trial and there was surprisingly a 0% correct response. Clues for this similarity matching were very subtle, including slightly changed arm length, and possible answers had 6 different combinations. Despite the time pressure, most of the participants focused on searching for obvious clues for pair comparisons; their saccades were very quick and shifting between targets and fixations were short. It is believed that the best strategy for this trial was to work out a target's total characteristics (i.e. the position of shorter arms) then to compare it with others. The results revealed from data, that it was unlikely that participants preferred such a search method. 'Partial comparison' was preferred. When obvious clues were detected, participants seemed to happily adapt and perform the task. In this trial, targets 2, 3 and 4 shared similar figures and angles, which mislead participants as the strongest hint. This echoes working memory theory (Reason, 1990) that within limited resource capacities, one would try to decrease mental load and only pick out only meaningful information. Secondly, from previous eye acuity test results, subjects preferred horizontal and vertical searches as opposed to diagonal ones. Although this trial was a failure, there are few noteworthy points worth mentioning: 1> the issue of predicting how and what individuals will look at? i.e. how designers can decide to locate relevant information to facilitate quality human-product interactions 2> the limitation of complex images vs. meaningful images for effective memory registration.

3.0 Conclusion

It is more certain to assume that visual search is highly dependent on strength of information rather than sequential search. If incoming information is situated within a more meaningful way, it is easier to trigger related schemas. A per-conceptual knowledge schema was formulated and related information stood by in working memory. The more significant the information is; the more effectively the subject can react. Puzzle trials and similarity matching trials have demonstrated direct influences of subject's responses from visual stimulus. Visual stimulus can be either easy or difficult in the transformation of referable information; in other words, there are simple clues and there are difficult ones. Secondly, despite the nature of the clue itself, the success of a visual search also relies on the context of the task, such as 'layout' that might facilitate visual search or increase limitations. How to present visual stimulus can therefore have significant affect on mapping behaviours. Visual search shortcuts are inevitable and misjudgments often involve bad layout. Finally, individual habitual behaviour, the most unpredictable section during trials, characterised diversified tendencies. There were great propensities for participants to have particular manners of search behaviour, significantly different from others, through trials P1and P2. Although S3 trials also indicated various behaviours, it was more likely to assume that participants did not really have specific searching strategies. Most of the data, short fixations and quick saccades between four stimuli, in S3 trials showed lack of confidence and it was natural to adopt ‘quick scan’ in order to pick up any hints. Participants showed high possibilities of implying precise searching strategies regarding explicit clues rather than implicit ones. Moreover, such strategies were greatly influenced by habitual behaviours.

3.1 Visual Stimulus, Context and User Approach

To achieve a successful comparison implies the effectiveness of three major aspects, stimulus, context and user approach. However, even if each element could be further developed for the achievement of efficient communication it is still not a guarantee for successful visual search. Brandt and Stark (1997) pointed out that eye movements are not random; they are influenced by the content of the visual scene. The location of stimuli is decisive for the first few saccades that attract attention. The nearer the target the greater the signal strength (Araujo, etc. 2001). It is suggested that nearby targets receive higher visibility and resolution from the fovea and therefore provide stronger signals to the attention; within visual search, saccades behave the certain path that prefers the line of sight to the nearby location. (Kowler, etc. 1995; Motter & Belky, 1998; Melcher & Kowler, 2001). It is therefore more decisive to consider a broader picture of interaction for such vision search behaviours – the effective relationship between the three elements. Human information processing is not as logical as machines or computer programs. Humans have the ability to work in imperfect conditions, to skip untruths, to create shortcuts, to optimise their judgments and find alternative solutions. Human information processing does not shut down its actions because of any occurrence untruths. The strength of stimuli, the effectiveness of context and the possibility of user’s diverseness can be measured accordingly for their impacts on the interaction result. The extensive question revealed from this study is “ how important are the impacts of such indefinite relations in respect to human-product interactions?” and “what are the limitations of adoptability by users?”

3.2 Alternative View of Human-Product Interaction

This pilot study’s insights points toward the potential development in human-product (human-machine) interaction in respect to domestic technological products. Within low risk appliances, the primary design concerns of human-product interaction may not be embedded in seeking the perfect model, but the adoptable formula with total consideration to stimulus, context and user approach. The contradictory technological mechanisms characterised by rigidity of operation do not really allow much flexibility. Hence, the nature of human behaviours consists of analog types of interaction (communication) in which the content of resources, strategies and operation manners are more flexible and adoptable according to the context of tasks. To resolve the difficulty of human-product interaction is therefore not merely to facilitate the quality of each individual element but more importantly is to engage a broader consideration and to seek a compatible platform for consolidating digital interaction and analog manners.

Figure 17: Alternative view of developing interaction formulae. The principal consideration places emphasis on the inter-relation rather than individual element.

Stimulus – form, material, sign, icon
| The transformation of information. During this process, information might be perceived differently.
Context – contextual study, procedure planning
| Decision-making. This can be various due to context limitation and user behaviours.
User approach – user center approach, cognitive studying

Whilst there are still unresolved problems present, it is hoped that this pilot study can provide some insights for multi-disciplinary studies with regard to human-product interaction.

References

1. Araujo, C., E. Kowler, et al. (2001). "Eye movements during visual search: the costs of choosing the optimal path." Vision Research Vol 41(No. 25-26):

3613 - 3625.

2. Brandt, S. A. and L. W. Stark (1997). "Spontaneous eye movements during visual imagery reflect the content of the visual scene." Journal of Cognitive Neuroscience Vol. 9(1): 27-38.
3. Brunswik, E. (1956). Perception and the representative design of psychological experiments. Berkeley, University of California Press.
4. Josephson, S. and M. E. Holmes (2002). Visual Attention to Repeated Internet Images: Testing the Scanpath Theory on the World Wide Web. ETRA, New Orieans, Louisiana, ACM.
5. Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). "The role of attention in the programming of saccades." Vision Research, 35, 1897-1916.
6. Land, M. F. and M. Hayhoe (2001). "In what ways do eye movements contribute to everyday activities?" Vision Research Vol 41(No. 25-26): 3559 - 3565.
7. Melcher, D., & Kowler, E. (2001). "Visual scene memory and the guidance of saccadic eye movements." Vision Research, 41, 3597-3611
8. Minsky, M. (1975). A framework for representing knowledge. The psychology of computer vision. P. Winston. New York, McGraw-Hill.
9. Moray, N. (1969). Attention: Selective Processes in Vision and Hearing. London, Hutchinson.
10. Motter, B.C., & Belky, E.J. (1998). "The guidance of eye movements during active visual search." Vision Research, 38, 1805-1815
11. Norman, D. A. and D. A. Bobrow (1975). "On data limited and resource limited processes." Cognitive Psychology Vol 5: 44-64.
12. Noton, D. and L. Stark (1971). "Scanpaths in saccadic eye movements while viewing and recognizing patterns." Vision Research vol 11(No. 9): 929-932.
13. Pelz, J. B. and R. Canosa (2001). "Oculomotor behavior and perceptual strategies in complex tasks." Vision Research Vol 41(No. 25): 3587-3596.
14. Pieters, R., E. Rosbergen, et al. (1999). "Visual attention to repeated print advertising: A test of scanpath theory." Journal of Advertising Research Vol. 36: 424-438.
15. Reason, J. (1990). Human Error. Cambridge, Cambridge University Press.
16. Rumelhert, D. E. (1975). Notes on a schema for stories. Representation and understanding: Studies in Cognitive Science. D. Bobrow and A. Collins. New York, Academic Press.
17. Internet resources: SR Research, "Eyelink II system configuration"

http://www.eyelinkinfo.com/body_system_config.htm , access date:25/09/2003