

Tracing Cognitive Processes for Usability Evaluation: A Cross Cultural Mind Tape Study

Jyoti Kumar¹, Janni Nielsen², and Pradeep Yammiyavar¹

¹ Department of Design, IIT Guwahati, India

² Department of Informatics, CBS, Copenhagen, Denmark
{jyoti.k,pradeep}@iitg.ernet.in, jn.caict@cbs.dk

Abstract. Cultural differences in cognitive processes and cognitive tools have been extensively documented. Design and use of culturally sensitive interfaces have been in demand in HCI for sometime. In this study the method of stimulated retrospective verbalization which is called here as Mind Tape study, has been used to capture cognitive differences of Danish and Indian users while interacting with chosen websites on a given task. The recording of the interaction captures screen activities and video of user. The replay of the recording is used as stimulus during a voice over interview. Using Mind tape, not only the sequence of activities during task fulfillment is observed, but also an insight into the user's cognitive processes, motives and intentions, regarding the choices made and activities done are recorded. The paper reports the cultural sensitivity and suitability of the mind tape method for cross cultural usability evaluations in light of the study conducted.

Keywords: Stimulated Retrospective Verbalisation, Usability testing, Cross Cultural.

1 Introduction

Verbalisation as a window to the cognitive processes of the user has been a well talked of method in the usability evaluation practice [1,2,3]. Concurrent and retrospective verbalisations have been compared and contrasted for their reliability and validity often [4,5,6]. Whereas the Concurrent Verbalisation (CV) suffers from sharing the cognitive resources with task fulfillment [3], the retrospective verbalisation has been accused of memory loss due to time lag or subsequent influences on STM. The validity of Stimulated Retrospective Verbalisation (SRV) or Mind Tape (for under the influence of the stimulus mind acts as a tape and unwinds the memory thread by thread) have been established by a few studies [7,8] and the quality of Mind Tape data also have been reported as compared to CV like Think Aloud (TA) [7]. On the other hand cultural differences in social setups have been reported [9] and cognitive basis of the cultural differences have also been argued [10]. Now the issue of culturally sensitive methods of usability evaluation is being raised in this paper. When there exist cognitive differences in cultures, do we also need to examine the suitability and sensitivity of usability valuation methods in cultural and

cross cultural contexts? In this study mind tape has been used as a method to examine cognitive processes cross culturally, namely with Danish and Indian participants, over a task of exploration and finding a place of interest in each country of interest to visit using national tourism websites of three countries, India, Denmark and China. The results suggest that Mind tape gives rich data for analyzing the cognitive processes and tools employed by the users in task fulfillment and the method is culturally suitable to both the cultures in terms of satisfaction reported and data gathered.

2 Method

2.1 Website Explorations

Websites of three countries were selected. The sites are the official tourist websites of the three countries. They all address the same target groups; potential tourists, and English version of all of them were available. The user studies were conducted with 7 pairs (a user and an interviewer) each from Copenhagen Business School, Denmark and from the Indian Institute of Technology Guwahati. A scenario framed the task which was I) to explore the three sites and II) to find a place of interest in the tourism websites to take back to a group of friends planning a holiday trip. All seven user/interviewer pairs came from interdisciplinary studies where computer science made up one part combined with another discipline. In Denmark students came from Copenhagen Business School, Department of Informatics. They were master students (beginning of 4th year) and had enrolled in a course in HCI. They had some idea of interface design and evaluation aspects. The students from India were bachelor students in their final 3rd year at Department of Design, IITG and they had a similar educational background in design and evaluation of interfaces. A few had visited the website of their home country, but none had visited or were familiar with all three websites. None had explored the sites extensively as they were requested to do during the task. The mean age of Indian participants was 21.57 with standard deviation of 0.73, the mean age of Danish participants was 26.14 with standard deviation of 2.29.

Table 1. Profile of Danish Participants

d	Male	Female	Age
d1	1		28
d2	1		26
d3		1	28
d4	1		26
d5	1		24
d6	1		29
d7	1		22
	Total 6	Total 1	Mean 26.14
			SD 2.29

Table 2. Profile of Indian Participants

	Male	Female	Age
I1	1		23
I2	1		22
I3	1		21
I4	1		21
I5	1		22
I6	1		21
I7	1		21
	Total 7	Total 0	Mean 21.57
			SD 0.73

The users were asked to correlate between cursor and the user’s eye as they browsed through the websites for the task fulfillment. To enhance this correlation the users went through a training session to learn to coordinate cursor and eye. During the first task, the exploration, the users were encouraged to get a feel for the country so as to be able to communicate to her/his friends. The second task was to find one place of interest in a website where the user would like to go with her/his friends for a weekend. To get around the problem with verbal overshadowing of TA, and to allow the visual interaction to unfold undisturbed, no requests for concurrent verbalization were made. The user worked at her/his pace and in peace during the whole session.

The data was collected by recording the entire interaction on the screen including a video image of the user. Immediately following each task, the interviewer replayed the recording and conducted a qualitative interview. The software used made it possible to record voice over the original recording. The interviewer paused and played the original screen recording asking the user questions like for e.g. “what are you looking for?” when the user’s mouse is seen wandering around on the screen for sometime without clicking, or “Why do you click there?” when the user clicks at some link. The answers from the users were developed upon to further probe into the user’s intentions and expectations. A questionnaire was applied at the end of all the three website explorations to get additional information about the overall view of the websites and the experiences with mouse eye coordination.

3 Results

3.1 Mind Tape Study

The Interactions with three websites were screen recorded and voice over interviews were conducted on them. Finally the voice over video was analysed for the users’ responses to the interviewers’ questions regarding what they were doing at specific instances during the website explorations. Some of the noteworthy observations are listed subject wise in Table 3 as an example of the kind of data that was obtained from the Mind Tape study.

Table 3. Observations from Mind Tape video

Subject	Observation	User’s response	Inference
d1	Indian Site: Mouse wanders in the beginning, checks the menus.	Looking for ‘Tajmahal’ for I have heard of only that from India.	Posit: Danish People/ In general people search by what they already know on an unfamiliar website
	Clicks ‘Heritage’ link	Expected that it will give me some pictures of Tajmahal.	Pictures are what information can be quickly and richly availed.
	Picture of Tajmahal comes on screen. Expression of dissatisfaction on the face of subject.	Got only one picture with little text so I started looking for some other link where I can get more info.	Need of many pictures.

Table 3. (Continued)

Subject	Observation	User's response	Inference
	Text of info comes	I am looking for pictures, I am not going to read 10 pages of text.	Lot of Texts is not preferred on a tourism website.
	Danish Site: Beginning... mouse wanders around	I am looking for something interesting	Posit: When the person is well aware of the place then one looks for something interesting (does it mean not known earlier)!
	Clicks link named ' <i>inspiration</i> '	I am looking for something interesting so I guess here is something...	The word ' <i>inspiration</i> ' promises for new and exciting on a tourist website.
	Further sub menus come upon clicking <i>inspiration</i> - sub menu - <i>culture</i>	I am not looking for so specific information when I click culture, I want a general picture.	There is a threshold of detailed information that one seeks while looking for a tourist place, at least initially.
	Chinese site: Beginning... Mouse wanders...	I thought Hong Kong is part of China, I am not able to get it.	Again search by know place on a less known site.
d2	Indian Site: Mouse static in the beginning.	Looking for some pictures to see what all places to visit in India, I do not know much about India.	Pictures as means of getting an image of the place.
	Selects <i>Beaches of India - Goa</i>	Because it has pics of beaches so I can go there	Probably familiar locations interest more
d3	Selects <i>places to visit</i>	I do not know anything about India so may be this is a good place to begin with.	Cognitive tools that aid in beginning to search is not names of places for they are unknown but the categories that represent them. This could possibly be a universal phenomenon.
	a list of places is shown	I do not know any of the places so this list doesn't give me desired information.	Further categories of places and then the list might have helped probably.
II	After a lot of trials on menu items	The purpose of this website is not clear... whether it is about introducing me to the culture.... Or it is also to help me get there...	Could it be much talked about- holistic thinking in east Asians.... Trying to get the bigger picture?

Table 3. (Continued)

Subject	Observation	User's response	Inference
I2	Looks at an image	It looks like from my very own place	Does this cultural identity phenomena relevant more to this individual or to the community?
I4	Gets a submenu filled with known items except one	These I know...OK... but what is this?.. let me click	Posit: In known territories, people explore the less known to them item.
I5	State wise organization of info	Why is it done state wise? I am interested not in states but the kind of holiday I want to have.	Information architecture to suit the motivation of the user was observed in user's of both the cultures.
I6	Highlights the text while reading	I always do it while reading, it helps me identify the text from rest	Cognitive tool used by most of the Indian participants while reading to focus on the text being read. Is it a cultural phenomena?

3.2 Rankings of the Websites by the Subjects Under Different Criteria

The subjects were asked to rank the websites after the task fulfillment was over. The criteria given were ‘the website they liked’, ‘The website that was most easy to use’ and ‘the website which had most pleasing interface’. Table 5 lists the frequency of ranks allotted to the websites by the Indian, denoted by ‘Ind’ and Danish user’s denoted by ‘Dan’ under each criteria given for every website.

Table 4. No. of participants from India (Ind) and Denmark (Dan) who ranked the sites under the criteria of liked, easy to use and Good Interface

Rank	Indian Website						Danish Website						Chinese Website					
	Liked		Easy to use		Interface		Liked		Easy to use		Interface		Liked		Easy to use		Interface	
1st	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan
2nd	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan
3rd	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan	Ind	Dan
	5	3	4	2	2	1	2	3	1	4	5	5	0	1	1	1	0	1
	1	2	2	2	4	4	4	3	2	3	2	1	2	2	2	2	1	3
	1	2	1	3	1	2	1	1	4	0	0	1	5	4	4	4	6	3

3.3 Grading of the Websites by the Subjects

After the task fulfillment, the subjects were also asked to rate the websites on a 7 point scale for how ‘attractive to look’, ‘exciting to visit’ and ‘friendly to use’ each website was. The results have been tabulated in Table 5 with mean (with standard deviation), maximum rating and minimum rating that each website got from Indian (Ind) an Danish(Dan) participants.

Table 5. Mean of ratings of three sites on a 7 point scale under criteria of Attractive, Exciting and Friendly of Indian (Ind) and Danish (Dan) participants

Rank	Indian Website						Danish Website						Chinese Website					
	Attractive		Exciting		Friendly		Attractive		Exciting		Friendly		Attractive		Exciting		Friendly	
	Ind	Da	Ind	Da	Ind	Da	Ind	Da	Ind	Da	Ind	Da	Ind	Da	Ind	Da	Ind	Da
Mean	5.4	5.0	5.4	3.6	4.5	4.5	4.7	4.3	4.0	4.0	4.7	5.0	3.5	3.8	3.4	3.6	3.2	3.1
Std. Deviation	.7	1.4	1.2	1.2	1.5	.8	.9	1.2	1.0	1.0	1.7	.8	1.1	1.7	1.1	1.8	1.1	1.3
Minimum	5	3	4	2	2	3	3	2	2	2	1	4	2	2	2	1	2	2
Maximum	7	7	7	5	6	5	6	5	5	5	6	6	5	6	5	6	5	5

4 Discussion and Conclusion

4.1 Hand Eye Coordination

3 Danish participants reported ‘no problem’ using the hand eye coordination and that is was ‘natural’, 2 Danish participants reported that is was difficult when ‘scanning the pages’ and that ‘the eye moves faster than the hand’. 4 of Indian participants reported the difficulty in ‘scanning the page’ and 2 said ‘it was natural while reading as one always does that.

Inference: Hand eye coordination as a means to get the data about visual focus of attention on the screen may be natural to some and they otherwise also may have a tendency to take the mouse where their eyes went in normal interactions. Whereas, to some, it was intrusive in their normal task fulfilment activity. Anyhow, text reading was observed to involve the cursor movement along with the text being read naturally, many Indian participants had shown the behaviour of highlighting the text being read for better attention. It may be posited here that the mouse track data can be a good approximation to the eye gaze data in case of the reading activities like text reading, menu items reading etc. but it may not be reliable in case of image viewing or scanning webpages.

4.2 Quality of Mind Tape Verbalisation

Subjects were able to recall satisfactorily what they were thinking/doing at the time of the activity being replayed on the screen. The voice over interviews yielded considerable data on the why's and how's of the activity. The Indian participants specially, gave an extended set of logical explanations of what made them do the activity, some even presenting their views about the site in general while the activity was being interviewed. Participants from both the cultures were comfortable in the mind tape study and the information related to cognitive processes and tools applied for task fulfilment were satisfactorily reported.

Inference: The rich set of verbal data corresponding to each activity which was possible in mind tape study could have possibly interfered with their normal task fulfilment in the concurrent verbalisation. The satisfaction that was seen in the users sharing their *why's* and *how's* of activity due to a human angle to the verbalisation namely, the interviewer, could have possibly not been there in the monotonous concurrent verbalisation. This helped in getting deeper insights into the cognitive processes employed by the users. In general it can be said that mind tape is a culturally sensitive tool for usability evaluation tasks.

4.3 Cross-Cultural Similarity in Cognitive Processes Employed

Search by familiarity: All participants, when searching in little known countries to them (like Denmark and China in case of Indian participants and vice versa in case of Danish participants) ordered their search from more known to less known places. Like in case of Danish participants exploring Indian website, 3 out of 7 participants, started their search from looking for 'Tajmahal', which was the only place in their prior knowledge (as reported), but upon finding no images corresponding to Tajmahal, moved to what interests them, like some of them searched for beaches in India. Whereas subjects looking for places when confronted with known set of places, looked for the one that was little known to them. But in finally deciding about the places, people based their decision on the combination of prior knowledge and supplements of info from the website.

Inference: The search by familiar could be phenomena common to the two cultures, or it could very well be universal phenomena, in case of travel websites. On the other hand the inquisitiveness for the odd one in the list of known places could either be an attempt for the mere information sake.

Role of images in decision making: Almost all participants from Denmark as well as India complained for the lack of pictures in Indian website. They articulated the role of images in getting a feel of the place to visit. They also closed very quickly those pages of the site that had no images.

Inference: This may speak of the similarity in the cognitive processes and tools of the users from both the cultures or it may be a universal phenomena. This needs to be further investigated.

4.4 Cross-Cultural Similarity in Cognitive Processes Employed

Query of cost as an aid in decision making in Indian participants: 5 out of 7 Indian participants searched for the prices of the facilities and used the information as a

primary aid in deciding about the places to visit. This behaviour was observed only in 2 of the Danish participants though.

Inference: As the sample under study is very small to generalise the inference statistically, still it becomes a significant suggestion towards further exploration into how do people from the two cultures employ cognitive tools in decision making.

Online reading habits: Indian participants (5 of 7) were found to select the text with mouse as they read, in the mind tape study they reported it as their normal habit while reading. None of the Danish participants had this habit.

Inference: Could it be possibly due to differences in cognitive tools people employ while seeking information online, specially through reading? The holistic versus analytical cognitive processes (in East Asians and westerners respectively) reported by Nisbet et al [1] are in action here? Further specifically designed experiment to study this phenomenon in web based information seeking behaviour could be conducted to verify/substantiate it.

4.5 Rankings of the Websites (Table 4)

As depicted in the table, 5 of the 7 Indian participants liked the Indian website most and said that they found it 'Organised', 'had images with relevant info', 'Concise with important guides', 'had Nice colours' and 'had relevant chunking of information' while 3 Danish participants liked the Indian website (giving reasons 'not confusing', 'information was grouped well') the most and 3 liked the Danish website the most (giving reasons, 'easier to navigate', 'had light colours' and 'was structured'). The one Danish participant who liked the Chinese website gave the reason of it having a lot of pictures. 4 of Indian subjects found the Indian website most easy to use (reasons elicited were 'front page had sufficient information, 'grouping of info was good', 'could locate places more easily' etc.) while 4 of Danish participants found the Danish website most easy to use (because of 'lots of useable links on the front page', 'Clear separation of information', 'menu made it easy' etc.). Importantly, two of the Indian participants who had liked the Indian website most and the Chinese website the least found the Chinese website most easy to use for it 'had a linear structure' and 'had nothing to search'. Both the Indian and the Danish participants (5 in both) found the Danish website having most pleasing to Interface and the reasons given were 'had best colour codes', 'had good selection of fonts, colours, photographs', 'had a neat and clean look' (Indian participants), 'had simplicity', 'was clean' (Danish participants).

Inference: a) Both the Danish and the Indian participants found the Danish website clean and simple which can help us hypothesise that the cognition of neatness and cleanliness depend on similar visual cues in both Danish and Indian cultures. b) A lot of pictures (and only pictures in eyes of d4, I4 and I5) on the Chinese website helped in deciding about the place in that country but it couldn't win for the most liked website for both groups of participants, from this we could hypothesise that though pictures become the most important element for deciding about places in tourism websites for the users of Danish and Indian origins (as reported by both groups of participants), but it doesn't win the site most liked award for user's of both countries prefer 'organisation' and 'neatness'. c) The ease in use reported by both groups for

their own culture website may owe to either the familiarity of the information available on the website or it may be due to cognitive difference the user groups have in reality. Further study need to be conducted for validation.

4.6 Ratings of the Websites (*Table 5*)

Indian website was reported as most attractive, getting an average rating of 5.4 by Indian participants and 5.0 by Danish participants on a 7 point scale whereas the Danish website was reported to be more friendly to use by both the groups (4.7 and 5.0 respectively). The divide in opinion of the two groups in terms of which was the most exciting to visit is clear when Indian group has favoured it's native site (5.4) in comparison to Danish site (4.0) whereas the Danish subjects have rated both the sites almost similarly (3.6 and 4.0) respectively.

Inference: The Indian site appearing attractive could possibly be attributed to it's bright orange colour in the layout, and plethora of selected images on the header which though one Danish participants disliked and many Indian participants ranked the Danish site's interface more clean and with soothing light colours, still during rating the Indian site has got more points under attractive attribute. Low ratings of Chinese website may be due to the unobvious position of links (which was on the images). Those subjects (like d4) who could figure out the links have rated it relatively higher because they could find a source to a lot of images which has been reported as one of the most sought after information sources.

To conclude, mind tape study does reveal insights into the cognitive processes of the users by developing upon and probing into the user's responses to the questions related to the activity they had just finished during task fulfilment. Furthermore, the human angle in the form of interviewer makes it easier and more meaningful to have a dialogue about the intentions and motives of the user in employing the cognitive tools, in form of choices, aids in decision making, preferences for colours, forms, images etc. while they perform tasks. The cultural suitability of this method to the two cultures under study has also thus been established. The study has revealed the common cognitive tools in two cultures like 'search by familiarity' and 'role of images in decision making', 'similarity in concept of clean and neat site' and the differences in form of the 'online reading habits' and 'search for holistic impressions'. The mental models of 'attractive site' and 'friendly site' have also been found to match for the two cultures. This study also advocates for the cultural sensitivity of this method as dialogue oriented cultures, it is posited here that Indian culture is one (on the basis of *relational, dialectical* and *person attribution* in Peng K [11]), as well as task oriented cultures, it is again posited here that western cultures are task oriented (on the basis of *non-contradictional* and *event attribution* in Peng K [11]), will find it suitable to have an interviewer to speak out their motives than just one way, monotonous verbalisations as in case of concurrent verbalisations like Think Aloud etc. These findings could help formulate further studies using mind tape method to explore the cross cultural cognitive process and tools differences and similarities in more detail.

5 Future Work

Further studies in finding role of the interviewer in the elicited data, impact of cross cultural pairs in voice over interviews for usability testing, probes useful in specific cultures, whether there are culture specific probes, could be conducted to expand and explore the possibilities of application and validity of mind tape method. The collected data itself is being further analysed for cultural cues for the method.

Acknowledgements. This study was co-funded by the Danish Council for Independent Research (DCIR) through its support of the Cultural Usability project.

References

1. Nisbett, R.E., Peng, K., Choi, I., Norenzayan, A.: Culture and Systems of Thought: Holistic Versus Analytic Cognition. *Psychological Review* 108(2), 291–310 (2001)
2. Nielsen, J.: *Usability Engineering*. Cambridge, MA, AP Professional (1993)
3. Ericsson, K., Simon, H.: *Protocol Analysis – Verbal Reports as Data*. MIT, Cambridge (1993)
4. Boren, M.T., Ramey, J.: Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication* 43(3), 261–278 (2000)
5. Nielsen, J., Clemmensen, T., Yssing, C.: Getting access to what goes on in people's heads? - Reflections on the think-aloud technique, Nordi CHI (2002)
6. Nielsen, J.: Reflections on concurrent and retrospective user testing. In: *Proceeding of India Human Computer Interaction Confr.*, Bangalore (December 2004)
7. Van den Haak, M., De Jong, M., Jan Schellens, P.: Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behaviour and Information Technology* 22(5), 339–351 (2003)
8. Zhiwei, G., Shirley, L., Cuddihy, E., Judith, R.: The Validity of the Stimulated Retrospective Think- Aloud Method as measured by Eye Tracking, CHI 2006 proceedings (2006)
9. Geert, H.: *Cultures consequences*, 2nd edn., Sage publications
10. Nisbett Richard, E.: *Geography of thought*, Free press
11. Peng, K., Ames, D.R., Knowles, E.D.: *Handbook of Cross-Cultural Psychology*. Oxford University Press, Oxford (2000)