Indian Cultural Effects on User Research Methodologies

Jack Beaton
Nokia, Inc.
5 Wayside Rd.
Burlington, MA 01803 USA
jack.beaton@nokia.com

Ripul Kumar
Kern Communications Pvt. Ltd.
210 B, Swapnalok Complex
S.D. Road
Secunderabad, 500 003 India
ripul@kern-comm.com

CHI 2010, April 10 – 15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
Modern user research techniques such as Think Aloud usability testing were mainly designed and refined in Europe and North America. These techniques perform substantially differently in traditional Indian culture due to the participants' perception of social status differences between them and the moderator(s). Understanding and controlling these effects can make the difference between a successful research project and one that gains little reliable data. Examples are cited from India-based user testing and open-ended field research by Kern Communications for Nokia's Ovi Mail and Nokia Life Tools services in January 2009.

Keywords
User studies, user testing and evaluation, internationalization/localization

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): Evaluation/methodology.

General Terms
Experimentation; Human Factors

Introduction

In January 2009, Kern Communications and Nokia administered user research on two services now available in India: Ovi Mail, an email service with an emphasis on mobile users, and Nokia Life Tools, a service for rural users in emerging markets. The Ovi Mail testing used the Think Aloud protocol [1] in user tests of a mobile email setup process on the browser of Nokia S40 phones with 13 English-fluent and 12 non-English-fluent Hindi speakers in Hyderabad, none of whom were familiar with email or the Internet. The Nokia Life Tools research consisted of more user testing, one-on-one interviews, and a street survey, spread out through multiple villages and towns in Telugu and Hindi-speaking areas of central India. Kern Communications' accumulated experience and familiarity with testing in India is also drawn upon.

User research was affected by a number of interesting cultural factors. Primarily, the effects of the perceived differences in status between the moderators and participants affected the methodology and data received from formal testing. Although status is an issue across cultures, the effects on research in India were more extreme than in the USA. Additionally, factors in scheduling testing participants and supporting the necessary facilities are discussed below.

Related Work

Previous research has found that Indian moderators with Indian participants recovered significantly more data than Western moderators. [3] The "Bollywood Technique" is a form of culturally-localized user testing in which the participant adopts a fantasy scenario (eg., from a Bollywood drama) that may increase participant enthusiasm, but may not accurately reflect the participant's actual use cases for the interface, as they are pretending to be someone else during testing. [2]

Indian Social Status

Relative social status is extremely important to social interactions in India. Although the caste system as one indicator of social status is weakening, other means of identifying status superiority or inferiority are being used in its place. The Western concept of universal equality is foreign to India, where it's normal for people to feel insecure until they are certain what their relative positions are in comparison to the others around them.

Higher status may be interpreted from these indicators:

- Being rich, as shown by having valuable clothes, accessories, and possessions;
- High compensation in salary and international travel opportunities at work;
- Working for a well-known company (eg. Nokia);
- Being older;
- Being male (not as much an effect on the minds of highly educated and/or wealthier participants);
- White Americans and Europeans, or Indians who have traveled or live abroad;
- Employing another person, or being employed by and reporting to a high-ranking person;
- Fluency in spoken English;
- Possessing a degree from a American or European university;
- Demonstrating useful knowledge or skill that others do not (including how to use an interface). This gives participants a strong incentive not to expose confusion.

The differences of status are often too subtle for a Westerner to notice. However, the fundamental reality of the status system is soon made shockingly obvious to a Westerner who spends sufficient time in India, as one of the authors discovered when a stranger in a rural village spontaneously asked him (a white American wearing expensive clothes) to bless her grandchild while he was walking down the street. In this case, the status difference was relevant to the Hindu religion, in which current high status is seen as a cosmic reward for saintly morality in past lives. Status is therefore extremely important in India and strongly affects behavior.

Effects of Perceived Relative Social Status

Kern Communications has learned by trial and error that perceived status differences reduce the amount of quality data gathered during Think Aloud user testing.

If the moderator appears to be *high status*, the participant will "clam up" for fear of exposing a "wrong" reaction that may incite scolding from the moderator. In earlier testing with Nokia in Hyderabad, Kern discovered that participants remained only minimally responsive until a Finnish observer left the room. In another notable case, a street survey in Indore had to be abandoned because participants simply gave no answering information at all, even when the white Nokia employee moved down the street to allow an Indian Kern employee to ask questions alone. A stranger then walked up and politely said (roughly translated from Hindi): "India is a great country, but this is a dirty area. You're clean people. You don't belong here. Your area is several kilometers in that direction." In short, this was their part of town, and we were making the local inhabitants nervous by introducing our high-status personalities into their perceived low-status community, a difference that they openly acknowledged in their polite, but firm, request.

If the moderator appears to be *low status*, the participant may refuse to acknowledge the questions or instructions of the moderator. We observed this when, during a street survey of a rural town, a young female Indian Nokia employee asked an older Indian man in a clean white shirt, riding a shiny (and for that town, expensive) motorcycle a question about his mobile phone. The Nokia employee, by comparison, had "dressed down" so as not to intimidate the rural testing participants. The older man ignored her question and instead demanded to know "who are you to come to our town and question us?" This was a much more hostile response than those from lower-status survey respondents in that town just minutes before.

In addition, status effects were observed to affect street survey results. High status individuals react, in body language, more approachably to high status strangers, and this makes it easy to bias street survey samples, especially if working towards a quota. Participants most comfortable speaking with English-fluent interviewers, especially white male American interviewers, are usually relatively-well-off men. It takes a stronger force of will from an interviewer, and more time to defuse tensions on arrival, to walk up to and question a less approachable person.

While the street convenience survey mentioned above did not get a single good respondent in the lower-status neighborhood of Indore, we quickly listed several dozen in Indore's higher-status Treasure Island Mall. However, the dramatic contrast made us suspicious of

the value of collecting such a large number so quickly. This caused us to extend the survey to the rural town.

Participant 'Status Loss' During Testing

Although controlling for initial status is possible for brief interactions, this control quickly wears off as a user test progresses. As a result, the Western-designed Think Aloud protocol visibly and irrevocably degenerates.

Confusion over technology is an indicator of lower education and lower social status in Indian society. Assuring participants that "we're testing the product, not you" has little effect. This may be because products are more expensive to Indians than to Americans, and if an Indian cannot figure out how to do something with the product they have, they cannot do it at all. Blaming the designer and switching products is not an option. Admitting and explaining confusion over technology lowers Indian participants' perceived status, but it is also the basis for Think Aloud testing. For this reason, testing sessions seemed to have three parts:

1. Introduction

Participant narration of thoughts lasts until their first breakdown. They stop narrating at that point.

2. Interrogation

The participant is confused and is trying to pump the moderator for hints, but wishes to keep the fact of their confusion ambiguous and deniable so as not to lose status. The participant may:

- Silently stare at the prototype until the moderator is forced to intervene to continue testing, sometimes long after the screen has timed out and gone dark;
- Ask indirect questions about the prototype, trying to avoid explicitly saying they are confused;
- Stare motionlessly at the moderator, looking for hints, possibly while slowly interacting with the prototype to check for moderator facial reactions in response to their movements.

3. Submission

The participant has been forced to admit they do not understand the interface, and is now in a lower-status role. This is confirmed if the moderator gives a hint to cause them to continue, confirming that the moderator is higher status because he/she knows more about the interface. He or she will not volunteer information and will wait for instructions, with minimal exploration. If the participant's belief in their status survives this experience, they are still less likely to explore, for fear their status will suffer more if they are confused again. Additionally, efforts to "draw out" the participants once they reached this passive state seemed tiring to the moderators, which may contribute to moderator errors.

Project stakeholders are generally not willing to make expensive design changes unless there is clear and indisputable evidence that the participant does not know something, which is exactly what the participants did not want the moderator to learn, for fear of losing status. As a result, the English-language test session recordings had to be reviewed many times to interpret what truly confused the participants. Compared to similar time investments with American participants using Think Aloud, less data was collected.

The contrast was especially obvious when participants who had dramatically failed to complete a task rated it

as being "very easy" on their post-task Likert scales. This data had to be thrown out as invalid.

Controlling for Perceived Social Status

Dress to match the socioeconomic status
Kern employees regularly wear different clothing to work and more or fewer items of gold jewelry (a common accessory in India) depending on the expected wealth level of the participants.

White internationals should stay out of sight
Previous research [3] has found that an Indian moderator can collect significantly more data that a white American moderator can collect from Indian participants. We have found that large status differences may drastically increase this, and may be caused by a non-Indian observer merely being visible.

Be suspicious of fast, low-cost surveys with a large N
The cheapest and easiest way to achieve a large population of Indian street survey respondents is to quickly log large numbers of high-status individuals from urban malls, skipping low status communities. Phone and Internet surveys are even more likely to reach only the richest members of Indian society.

Consider role-playing exercises instead
Conflating a test of both the applicability of the use case and a design's usability may be incompatible with Indian culture. The Bollywood Technique, in which participants pretend to be movie characters using the same interface, seems to prevent sensitivity over the participant's personal status. [2] However, because the use case is intentionally false, it may not inform whether or not the interface would be appropriate for the participants if they were "being themselves". That said, use case information should generally be captured by more open-ended techniques earlier on, such as concept validation, not by formal user testing.

Account for effects on Indian moderators in analysis
The culture of the moderator is also a factor and cannot be ignored. We believe our young Indian 18-25-year-old moderators were affected by the greater age of participants, causing them to give more hints during tests. Figure 1 shows an example of task completion rates. The 36-50-year-old participants failed; this can be explained by unfamiliarity with technology. The younger 26-35-year-old participants all succeeded with help. However, the 18-25-year-old participants, those in theory most familiar with technology, generally also failed. We therefore theorize that the strange "hump" in task success is due to a combination of enough youth for technological familiarity and enough age to successfully extract help from the moderator.

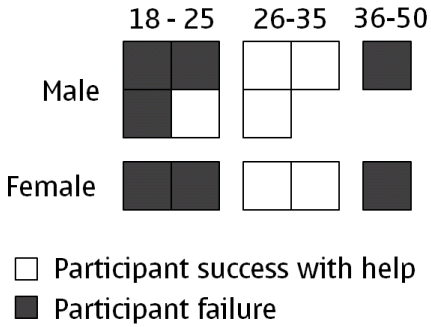

Figure 1. Failure rates of 13 English-fluent Indians without any previous experience using email, for completing an email setup wizard on a mobile phone.

Logistics and Scheduling

In this section, we will discuss the recruiting and scheduling of Indian test participants, and the logistics of ensuring a working test environment in India.

There are many reasons for Indian recruiting to be more difficult than recruiting in America and Europe. In Indian culture it is difficult and rude to disappoint people who ask things of you during (but not outside of) conversation. This includes being recruited for user tests. Recruits will continually say they will attend and sometimes that they are even on their way to user testing, but do not appear. They may stop answering their phones and go missing, or give an outrageous excuse; real examples include a puja (religious ceremony) scheduled at exactly that time that they forgot about, or that they were frightened of terrorism. It is normal in India for job applicants to accept multiple job offers, and only show up at one, and so participants are not shy about skipping testing. Also, Indian roads are often very crowded, and so it may be that they *are* on their way, but don't want to disappoint by saying they will be late, even if they truly will be.

For these reasons we found it is best to keep "as recruited" and "as participated" participant numbers separate in the records. Rather than scheduling specific people at specific times and seriously expecting that is when testing will occur, it is more realistic to view Indian recruiting as trying to raise the general probability that a participant will be in the testing room at any given time until the participant quota is reached. This may require continuous recruiting in parallel with ongoing testing, as opposed to setting a test schedule in advance. Without aggressive and continuous recruiting, Indian show rates may be as low as 0%.

In cases where continuous recruiting cannot be done during testing, participants can be double-booked. This is especially productive because typically adequate compensation (one hour for 1000 rupees, or US$22) is cheap enough that an extra participant can be used for valuable one-on-one interview time in parallel with user testing, should both participants attend.

India is constantly developing its infrastructure. Expect power, wired Internet, and cell phone signal to go down for periods of hours multiple times every day. It is best to have both a UPS and a generator on location.

Conclusion

As India develops and Indian consumers continue to purchase electronics, interface usability will grow in importance. However, this does not mean that they will psychologically shift to adopt Western mindsets. Just as Western products must change to reflect Indian cultural norms, testing methods developed in Western countries must also adjust to match those norms.

Acknowledgements

We thank the Nokia Ovi Mail, Life Tools, and User Research & Usability teams for funding the testing that made this learning possible.

References and Citations

[1] Ericsson, K.A., Simon, H.A. Verbal Reports as Data. *Psychological Review*, 87 (3), 1980, pp. 215-251.

[2] http://www.humanfactors.com/services/bollywood.htm

[3] Vatrapu, R., Perez-Quinones, M.A. Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. *Journal of Usability Studies*, Vol. 1, Issue 4, August 2006, pp. 15-170.