No IM Please, We're Testing

Richard Boardman
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043 USA
rickb@google.com

Copyright is held by the author/owner(s).
CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract
This paper discusses the use of instant messaging (IM) as a communication tool during usability studies – primarily between the interview and observation rooms. The benefits and challenges associated with providing an IM link are discussed, based on feedback from a survey of study moderators and observers. Observers were much more positive about the use of IM than the moderators. A key concern to moderators was the potential distraction to themselves, participants and observers. In contrast, observers greatly welcomed the opportunity to ask questions and help deal with buggy prototypes. Guidelines are outlined for the effective use of IM within a usability context, and contexts outlined when an IM link is most appropriate.

Keywords
Usability testing, instant messaging, test methodology

ACM Classification Keywords
H5.3. Group and Organization Interfaces: Collaborative Computing, I.m Computing Methodologies: Misc.

Introduction
Instant Messaging (IM) or "chat" is fast becoming a pervasive digital communication tool. It has become an important communication tool in many domains, both business and personal, e.g. [1]. In this experience report, IM is considered as a communication tool in another domain, that of usability testing.

The report is driven by the personal experiences of the author, a usability researcher at Google. Initially, I employed IM whilst evaluating a multi-step sign-up UI that required "back-end" approval at various stages. The approvers were located in the observation room, whilst I moderated in the interview room (this is the typical usability arrangement at Google). So as not to distract the participant, I used IM from a laptop out of their direct view. IM proved to be essential in coordinating the participant's progress with the approval in the observation room. After these initial experiences, I experimented with IM in a number of subsequent studies. Some example transcripts are listed below illustrating cases where IM was useful.

1. Here, an observer reminded me to ask about a UI feature I had forgotten to cover:

Moderator: any more questions b4 we wrap up?
Observer: Add the pulldown quickly? (the checkbox)
Moderator: aha I forgot about that! thx
Observer: no thank YOU -- this guy is really chatty! (-;

2. In this example, IM was helpful in confirming some unexpected UI behaviour:

Moderator: is this right? am I on the right page?
Observer: that's a prototype error - if it's not too late, refresh the page

Whilst the IM link offered clear benefits to the team, it also lead to challenges for me as the moderator - primarily not becoming distracted from the main flow of the test whilst monitoring the IM window. Furthermore, in some cases I had to decline observers' requests, e.g. when they wanted me to prompt the user.

This paper is aimed at investigating this trade-off, between the advantages of improved communication, and the challenges that accompany it. There is little literature on the use of IM in usability research. Voida *et al.* report their experiences in performing interviews with remote participants *via* IM [2]. However, there has been no discussion of IM communication between the moderator and observers during a usability test.

Rather than focusing on the author's personal experiences, the paper is centered on a survey of usability study moderators and observers, regarding their use and opinion of IM as a usability tool. The next section presents the survey and results. Then the report concludes with some suggested guidelines as to appropriate use of IM during usability testing.

Survey of Study Moderators and Observers

I distributed the survey to a number of usability test moderators and study observers - at Google and 2 other internet companies. The survey asked about usage of IM during usability tests, and its perceived advantages and problems. Twenty-six responses were received - 11 moderators, and 15 observers (designers, product managers, engineers and tech writers).

The first question asked participants whether they thought IM was a useful tool during usability testing. This immediately highlighted a strong contrast between study moderators and observers (see Table 1). Whilst the observers were mostly positive about the idea of IM, usability researchers were much more skeptical.

The next two sections cover the feedback received from each side of the one-way mirror: (1) study moderators, and (2) study observers.

Role	# replies	"Is IM useful in usability tests?" (1 strongly disagree, 5 fully agree)
Moderators	11	Average: 2.3 (median 2)
Observers	15	Average: 3.9 (median 4)

Table 1: Summary of survey responses.

The Moderator View – "IM as a distraction"

Eleven moderators replied. All were usability researchers at Google and several other companies. As can be seen in Table 1, they were mostly negative about IM during usability, and tended to focus on the downsides of its use. Only 3 moderators had tried IM and only one was fully positive about it. Several discussed both pros and cons, but argued that the disadvantages outweighed any benefits. They identified the following problems (ordered by frequency):

- ***Distraction of the moderator (7 Mods)*** – The most common criticism of IM within a usability context was that it would distract the study moderator, e.g. *Mod5: "It's hard to pay attention to what the user is doing and what is happening with your IM window"*. Several had tried it but been put off for this reason, e.g. *Mod6: "I missed several critical (fast) interactions that the user was taking."*
- ***Distraction of participant (6 Mods)*** – Six moderators were concerned about distracting the participant whilst typing for IM (or otherwise), e.g. *Mod5: "In general I think it would distract some users to hear me typing on my laptop, even if they thought I was only typing notes".*
- ***Back seat driving by observers (4 Mods)*** – Four moderators were concerned that IM would allow observers to interfere with study tasks and hence reduce the validity of findings, e.g. Mod3*: "When I worked at X, the usability analyst sat behind the mirror with the PM [Product Manager]. While the analyst was the only person allowed to talk in the microphone, we often had problems of PMs trying to interfere with the study ("tell him to click 'ok,' ask about this, ask about that"). These suggestions usually corrupted the credibility of the activity."*
- **Distraction of Observers (3 Mods)** – A third area of distraction was to observers trying to communicate with the moderator, e.g. *Mod3: "I also worry that if observers think they can communicate with me, they may spend less time paying attention to what the user is doing. I'd rather they spend that energy taking good notes."*
- ***Encourages observer laziness (2 Mods)*** – A final area of concern was that IM could encourage various forms of laziness on the part of observers! Two moderators worried that it may discourage the observer from preparing beforehand, e.g. *Mod2: "using IM during a test allows the PMs/designers to go into a test less prepared. When I have told designers/PMs that they can't talk to me during testing, they immediately start trying to think of all the questions they want me to ask"*. One of them was also concerned that some observers may not even turn up, *Mod2: "The designer (or PM if they created the product) must be in the room watching every study (instead of at their desk w/ AIM up)."*

Whilst moderators were generally negative, several considered potential benefits of IM as a usability tool, or situations where it may be appropriate:

- ***Training new moderators (3 Mods)*** – The area of clearest benefit was for training new moderators, using IM to give "on the fly" feedback regarding use of time, and choice of questions. For example, Mod2 stated: *"When we have new grads, it can be helpful for us to give them feedback during a test … it seemed to really help her when she was having trouble wording questions or redirecting participants that had gone off on a tangent."*
- ***Allowing observers to ask additional questions (3 Mods)*** – Three moderators mentioned this benefit, e.g. *"The PM/designers may come across new questions they want to ask based on what the user does".* However, many felt there were other channels such as talking before or after studies, e.g. *Mod3: "This does not mean that the PMs and designers are not involved; I always run the tasks/script by the designer and PM beforehand, so there are no surprises."*
- ***Dealing with buggy prototypes (2 Mods)*** – One moderator mentioned a recent study where IM proved helpful in dealing with an early prototype, Mod5: *"if the prototype broke in a way I didn't understand, I could communicate with the front-end engineer about it (e.g. he could tell me to steer the user away from using a certain feature, or suggest that we return to the home page)."*
- ***Supporting ad-lib studies (2 Mods)*** – Two moderators suggested that IM would be useful when there had been little time to prepare a script and/or get input from the team, e.g. *Mod6: "There are some situations where IM may be useful … if there is little or no time to prepare for the study and the analyst is forced to run without a script".*

Overall however, moderators argued that the problems caused by IM outweighed any potential benefits. Only one moderator (besides the author) was fully positive, focusing on the benefits of allowing observers to ask questions. However, he noted that there had to be ground-rules, *Mod4: "There has to be good etiquette. We have to be able to respond to requests with a simple 'no' if we don't agree with the request".*

The Observer view – "IM as a tool"

The feedback from the fifteen observers highlights a major contrast to the moderator group. Twelve were highly in favour of an IM link. The following benefits were mentioned (ordered by frequency):

- **Asking additional questions (12 Obs)** – Several observers recalled studies where they had wanted to ask questions but were not able to, e.g. *Obs4: "During my tests, I had questions/comments that I wanted to ask, but couldn't and left those questions unanswered"*. Many welcomed the ability to ask questions without the need for a follow-up test, particularly when a user had behaved unexpectedly, e.g. *Obs6: "Its most useful is when the user is reacting unexpectedly and we need to drill down on why, I can ask questions directly. The new data from an interesting session segue saves a lot of second-guessing in follow-up and may prevent the need for a follow-up test."*
- ***Dealing with a broken UI (5 Obs)*** – Five observers argued that IM was useful when a buggy

prototype was being tested, e.g. *Obs5: "It helps to give the usability analyst a hint about how to make a 'bug' in a mockup behave correctly".* Two also mentioned occasions when they had to interrupt a test when it had progressed to a buggy or undesirable area of the interface, e.g. *Obs6: "sometimes there's an urgent communication needed – Don't show them that area, it's got confidential information."*

- ***Watching remotely (1 Obs)*** – One observer stated that IM allowed him to attend sessions remotely by receiving a "blow by blow" account of events from the moderator or another observer. This provides some supporting evidence to the moderator concern that IM may encourage some observers not to attend sessions!
- ***Avoiding noise in the observation room (1 Obs)*** – One observer talked about another use of IM - to communicate *within* the observation room, with other observers, thus avoiding noise.

Whilst the observers were generally in favour of an IM link between the observation and interview rooms, several were aware of potential downsides:

- ***Distracting the moderator (4 Obs)*** – Four commented that an IM link could both add to the moderator's overload and interfere with study validity, Obs6: *"We should make it clear what diversions can be introduced. You can't divert too much off the task list because then comparisons between tests are more difficult and less reliable".*
- ***Distracting the observer (4 Obs)*** – Some observers were also aware of the distraction to themselves, *Obs6: "Its hard to focus on the participant while focusing on questions or follow-ups you want to get in there".*

Overall, there was a clear vote in favour of IM during usability tests. Still there was a perceived need for appropriate usage of IM, to avoid invalidating studies and overloading moderators. Two observers suggested some form of guidelines, e.g. *Obs4: "I'm sure we should identify some sort of process/guidelines so that we can't bother you with trivial questions, but I think it would provide a very positive benefit."*

Guidelines for using IM during Usability

The survey findings highlighted a contrast between the needs of moderators and observers behind the one-way mirror. Although this is a relatively small sample size, I envisage that the results will hold more widely in the industry. This section presents some guidelines for the use of IM during usability testing. They are based on comments gained in the survey, and the author's personal experiences. They represent an attempt to provide benefits to both moderators and observers, whilst mitigating the challenges outlined above.

1. ***Set observer expectations*** – Firstly, observers should be reminded that the moderator is in charge! It should be made clear that the moderator may not be able to respond to all requests, or may decide that a request is inappropriate. IM should not be seen as a tool to direct the study, but one to contribute to it for the benefit of the entire team.
2. ***Employ one*-way IM** – In general, observers should not expect a response to all IMs, especially with a complex UI or task. Enforcing a one-way IM channel from the observation room to the interview

room is one way of minimizing moderator overload. The reverse channel can remain available for emergencies (e.g. a major bug).

3. ***Avoid distracting the participant*** – Distraction can be minimized if the moderator uses a laptop (angled away from the participant) for IM. One-way IM as discussed above will also reduce distraction. Finally, the moderator should be sure to turn off all notification sounds in their IM client!
4. ***Nominate one IM contact in the observation room*** – This avoids the situation where multiple observers are trying to converse with the moderator. This is also a useful strategy for stimulating discussion in the observation room.
5. ***Employ a secure usability-specific IM account*** – Use a secure client and create a dedicated usability account. This will ensure confidentiality and avoid interruptions.
6. ***Only IM at the end of a session*** – Distractions to the moderator, observer and participant can be minimized by only using IM at the end of the study to communicate last-minute question requests.

IM may be particularly useful during the following usability test contexts:

1. ***Training a new moderator*** – IM can allow an experienced moderator to give mid-study guidance and tips. However, care should be taken not to overload the new moderator.

2. ***Brainstorming questions*** – IM can also be useful when there has been little time for the moderator to prepare a script, or during a pilot session when a script is being developed.

3. **Dealing with a buggy prototype** – A key strength of IM in a usability context is to allow observers to debug a prototype, or suggest alternative paths from behind the scenes.

4. **Federated studies** – A federated study is one where one participant is shown multiple prototypes during a single session. IM can allow the moderator to confirm the correct product teams are present before moving on.

5. **Improving observer attendance** - One final benefit of setting up an IM link is that it can improve observer engagement by giving them the ability to ask questions.

Conclusion

This report highlights a contrast between study moderators and observers regarding the utility of IM within usability tests. It is hoped that the guidelines may help teams benefit from IM within usability tests when they deem it appropriate. However, the author acknowledges – based on his first-hand experience – the increased overhead from IM usage for the already busy moderator. The final decision regarding the use of IM should rest with the moderator alone.

Acknowledgements

Thanks to the Bay Area usability researchers and observers who took part in the survey.

References

[1] Grinter, R., Palen, L., Instant messaging in teen life IM everywhere, Proc. CSCW 2002, p.21-30.

[2] Voida, A., Mynatt, E., Erickson, T., Kellogg, W., Interviewing over IM, Proc. CHI 2004, p.1344-1347.