Seeing the Forest and the Trees:
Towards a Comprehensive Usability Testing Method

Susan Hura | Principal, SpeechUsability
Peter Leppik | CEO, Vocal Laboratories Inc.
Mobile Voice Conference | April 22, 2010
Some background
Many ways of evaluating speech applications

Usability Testing

- Task Completion Rate
- Automation Rate
- Misroutes
- Customer Satisfaction
- Tuning
- Containment
- Focus Groups
- QA
- Recognition Performance
- Regression Testing
Usability testing is a method for...

• observing end-users individually,
• as they interact with an application,
• completing predefined tasks, and
• soliciting their opinions about the interaction you observed
Many Flavors of Usability Tests

- Application: prototype or production
- Users: actual end users or surrogates
- Data: real, dynamic data or created, fixed
- Venue: In-person or remote
Comparing Methods

Large-Scale Remote
- Participants interact with the system in their natural environment
- Participants complete a survey describing their experience
- Ability to test large numbers of participants in a relatively short time

In Person, Lab-Based
- Participant interacts in a lab setting
- One-on-one discussions with facilitator to describe experience
- Ability to explore participant’s reactions and opinions to specific parts of the interaction
Unique benefits to each

**Large-Scale Remote**
- Catches infrequent, but important, interactions
- Shows what’s happening in the interaction and with what frequency

**In Person, Lab-Based**
- Ability to ask “what were you thinking when...?” about choices participants make
- Shows how strongly participants feel about issues they encounter

**Reliability:** the issue occurs regularly

**Validity:** the issue is relevant to the interaction
Combining methods creates synergies

Remote Lab APPLICATION DESIGN Remote Testing Lab - Based Testing

Can be run simultaneously

Establish the importance of issues observed in remote testing. Discover user motivations and sources of confusion.

Validate frequency of occurrence of issues discovered in the lab. Discover issues missed in lab testing.

Both methods drive strategic changes to the application
Client Example

- Large IVR (1 million calls/month)
- Lab-based test with 15 participants, remote test with 500 participants
- Run simultaneously
- Issue: customers’ opinion of the voice of the system
Converging data

In the lab 2/18 participants reported that they disliked the voice.

Remote testing validated that 12% of participants disliked the voice AND that younger participants were more likely to dislike the voice.

Lab-testing revealed that participants disliked the voice because it sounded condescending or scolding.

Client was targeting a younger demographic.

Change the voice!
Higher quality data

• What’s happening and what it means
• Provides both a broad view of customer behavior and in-depth understanding of their motivations, needs, and preferences
• Allows organizations to make strategic changes to self-service applications with less risk
Thank You!

- Susan Hura
- 512.897.0350
- susan@speechusability.com

- Peter Leppik
- 952.941.6580
- pleppik@vocalabs.com