Evaluation

CS 376
Crystal Fong
April, 22 2008
Methodology Matters: Doing Research in the Behavioral and Social Sciences

• The research methods used are just as important as the research itself.
• No single research method is perfect
• Optimize results by using a combination of different methods
• Ideally, we would like to Maximize: Generalizability, Precision & Realism
Methodology Matters: Doing Research in the Behavioral and Social Sciences

Figure 2: The strategy circumplex (adapted from Runke & McGrath).

Maxima for each criterion:
A = Generalizability
B = Precision
C = Realism
Methodology Matters: Doing Research in the Behavioral and Social Sciences

- Analyzing the design of the study
- Comparison Techniques
- Accurate Baserate Information
- Correlation/Differencing
- Validity
Methodology Matters: Doing Research in the Behavioral and Social Sciences

- Measures
- Self report, Observation, Trace measures, Archival records
Practical Guide to Controlled Experiments on the Web: Listen to Your Customers Not the HiPPO

- Controlled web experiments allow us to quickly evolve our systems
- “... data trumps intuition ...”
- Small changes can make big differences in ROI
Practical Guide to Controlled Experiments on the Web: Listen to Your Customers Not the HiPPO

- Controlled Experiments
- Users randomly distributed
- Variants should only differ between Control and Treatment
- Sample size
Practical Guide to Controlled Experiments on the Web: Listen to Your Customers Not the HiPPO

- Online Experiments
- Treatment - incremental ramp up
- catches errors early on small populations
- automate to auto abort on underperforming experiments
- Beware of day of week / time of day / seasonal effects
Practical Guide to Controlled Experiments on the Web: Listen to Your Customers Not the HiPPO

- Limitations of Controlled Experiments
  - Generate a lot of data but very little user comments on their own.
  - usability labs
  - Coming up with a good OEC
    - One that includes long term goals
  - Beware of the Hawthorne Effect
Discussion Questions

- Where would Controlled Experiments on the web fall on McGrath’s pie chart?
Discussion Questions

• Do you believe McGrath’s conjecture that you can’t optimize all 3 properties (generalizability, precision and realism) at the same time?

• McGrath is a proponent of using multiple strategies to gain data. He doesn’t really go into how to combine the data. From experience, has anyone had to combine with data from multiple techniques?

• What are good ways of determining which techniques to use when designing a study?
Discussion Questions

• There is a lack of qualitative feedback in controlled experiment tests, how might one address this in their study? (User satisfaction?)

• Do you think the Hawthorne effect is really an issue in studies conducted on the web?

• “Apple has demonstrated that divine intervention can trump data”

• What other areas of research could benefit from the ideas presented as controlled experiments?