Evaluation, When and How

Tamara Munzner
University of British Columbia

Evaluation: How Much Evaluation Is Enough?
Panel, VIS13

Evaluation: broadly interpreted

• in private: ponder on case-by-case basis
  - as an author
  - as a reviewer
• in public: series of 4 meta-papers
  - as a reviewer
• in public: series of 4 meta-papers
  - as a reviewer
• as an author
• as a reviewer

“Of course...”

• you should evaluate your work
  - use appropriate methods!
• you should not have a user study in every paper
  - avoid litmus test and cargo cult thinking

Victories and challenges: I
• evolving sophistication: the user study pendulum swings
  - no user studies at all
    - a few dubious ones, lacking rigor
    - some good ones appear
  - rigorous studies are common
  - but pushes to change culture often overshoot...
    - some reviewers expect all papers to have user studies
    - some authors do user studies without understanding why or how

User studies: tricky to do right

“If you don’t know where you are going, any road will get you there.”
- Lewis Carroll
• many ways to go wrong that are not obvious from reading finished papers
  - ideal: collaborate with experts... several times!
• good resources
  - How to Design and Report Experiments, Andy Field & Graham Hole, SAGE
  - Experimental Human-Computer Interaction, Helen Purchase, Cambridge Univ Press

Victories and challenges: II
• significance testing with controlled experiments
  - we’ve moved beyond “my friends liked it”
  - new frontier: multiple regression for comparison
  - cognitive measurement of graph aesthetics.  Ware, Purchase, Colpoys, and McGill.
  - new frontier: thinking beyond time and error
  - qualitative vs quantitative
    - different axis from lab/field
  • BELIV workshops
    - 06 AVI, 08 CHI, 10 CHI, 12 VisWeek

Of course...

• you should evaluate your work
  - use appropriate methods!
• you should not have a user study in every paper
  - avoid litmus test and cargo cult thinking

Victories and challenges: III
• post-deployment studies with target users
  - we’ve moved beyond “I’m the only one who’s used it”
  - new frontier: post-adoption studies
  • Seven Scenarios: only 5 out of 800!
    TVCG 18(9):1520-1536, 2012.]
  • what happens after you get that first paper out?...