Evaluation: How Much Evaluation is Enough?
Less and More

Min Chen, University of Oxford
Evaluation: How Much Evaluation is Enough?

Less and More

Min Chen, University of Oxford
Multi-faceted evaluation

- Originality, Rigor, and Significance

Methods

- Peer judgement
  - e.g., review, publications, presentation, citation

- Relation to prior work
  - e.g., fundamental theories, algorithms, experiment results, widely-accepted models, wisdoms

- Evidence
  - e.g., mathematical reasoning, engineering and deployment, experiment design and execution and analysis

- Contemplation
  - e.g., discussions (short-term), usability studies (medium term), and comprehensive surveys (longterm).

Common interpretation:

- Usability
Evaluation

- Empirical studies for making **new discoveries** in aspects of visualization
  - Perceptual, cognitive and social aspects

Evaluation

- Common expectation:
  - a usability study
<table>
<thead>
<tr>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within a project, evaluation is mostly an <em>agile manner</em></td>
</tr>
<tr>
<td>• see: Munzner’s nested model, 2009</td>
</tr>
<tr>
<td>• <em>difficult to present in a research paper</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Common expectation:</td>
</tr>
<tr>
<td>• Waterfall model</td>
</tr>
</tbody>
</table>
What visualization is really for?
- Making observation
- Facilitating external memorization
- Stimulating hypotheses
- Evaluating hypotheses
- Disseminating knowledge
- **SAVING TIME**
- see: arXiv:1305.5670

Common expectation:
- Insight gained
Evaluation

- Evaluation is a **long-term** process
  - An independent and objective conclusion may take decades to obtain

Evaluation

- Common expectation:
  - short-term process
  - “self-certified clinical procedure”
  - positive conclusion
Evaluation ≠ Evaluation
Let us encourage:

More “Evaluation”

Let us demand:

Less “Evaluation”