The Effect of Team Exploratory Testing – Experience Report from F-Secure

Paula Raappana, F-Secure, M.Sc. student
Soili Saukkoriipi, F-Secure, PhD student
Ilkka Tervonen, University of Oulu, prof. emeritus
Mika Mäntylä, University of Oulu, prof.
What is Exploratory Testing?
Is it popular?
What about the future?
Exploration is not ON/OFF variable

ISTQB survey shows exploratory testing is popular

Different testing types [1] – What is the future

<table>
<thead>
<tr>
<th>Exploratory Testing</th>
<th>Confirmatory Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performed by human testers</strong></td>
<td><strong>Automated</strong></td>
</tr>
<tr>
<td>Testing philosophy</td>
<td>Testing is a knowledge intensive and creative activity requiring skills.</td>
</tr>
<tr>
<td></td>
<td>Testing is a mechanic and repetitive activity that can be described in explicit instructions.</td>
</tr>
<tr>
<td>Difficult to automate or outsource due to knowledge and skill needs, see [2]</td>
<td>Testing is automated and repeatable to provide fast feedback to development.</td>
</tr>
<tr>
<td>To be automated or performed by low wage workers</td>
<td>Rapid releases -&gt; increased regression testing [3]</td>
</tr>
</tbody>
</table>

---

Study
RQs

• *RQ1*: How does using the team exploratory testing approach affect the testing results?

• *RQ2*: How are TET-sessions experienced by the participants?
What is Team Exploratory Testing?

- Different skills and backgrounds
- Roles
- Nominated leader
- Common goal
- Common working method, ET
- Common outputs

Team

- Uninterrupted
- Reviewable
- Chartered
- Focused
- Has facilitator and leader
- Arranged regularly

TET session

- Framework for managing and controlling TET sessions
- Phases: preparation, session, and completion
- Controlled by the team

TET session approach

What is Team Exploratory Testing?

RQ1: ...testing results? Defects found & Efficiency

<table>
<thead>
<tr>
<th>Source</th>
<th>F-Secure TET sessions</th>
<th>F-Secure Other testing</th>
<th>ET test sessions Case A/B [14]</th>
<th>Sub-system and System testing in Testing phase T1/T2/T3/T4 [28]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defects found</td>
<td>115</td>
<td>427</td>
<td>169 / 34</td>
<td>20/12/24/0</td>
</tr>
<tr>
<td>Effective testing hours</td>
<td>116.5</td>
<td>1761.5</td>
<td>36 / 4</td>
<td>32/570/3150/160</td>
</tr>
<tr>
<td>Efficiency (defects per hour)</td>
<td>0.99</td>
<td>0.24</td>
<td>4.8 / 8.7</td>
<td>0.63/0.021/0.0076/0.0</td>
</tr>
</tbody>
</table>
RQ1: ...testing results? Defect severity

<table>
<thead>
<tr>
<th>Defect type</th>
<th>TET session</th>
<th>Other testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show stopper</td>
<td>0.9%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Urgent</td>
<td>2.6%</td>
<td>4.4%</td>
</tr>
<tr>
<td>High</td>
<td>20.9%</td>
<td>31.9%</td>
</tr>
<tr>
<td>Medium</td>
<td>53.9%</td>
<td>42.9%</td>
</tr>
<tr>
<td>Low</td>
<td>20.9%</td>
<td>18.3%</td>
</tr>
<tr>
<td>Enhancement</td>
<td>0.0%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Undefined</td>
<td>0.9%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>
RQ1: ...testing results? Defect type

<table>
<thead>
<tr>
<th>Defect categories</th>
<th>TET</th>
<th>non-TET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhancement</td>
<td>4%</td>
<td>1%</td>
</tr>
<tr>
<td>Inconsistency</td>
<td>9%</td>
<td>2%</td>
</tr>
<tr>
<td>Localization</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>Documentation and guidance</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Usability</td>
<td>24%</td>
<td>14%</td>
</tr>
<tr>
<td>UI</td>
<td>28%</td>
<td>18%</td>
</tr>
<tr>
<td>Functionality</td>
<td>22%</td>
<td>40%</td>
</tr>
<tr>
<td>Performance and reliability</td>
<td>3%</td>
<td>8%</td>
</tr>
<tr>
<td>Technical</td>
<td>1%</td>
<td>6%</td>
</tr>
</tbody>
</table>
RQ2 … participant experiences? Benefits vs Drawbacks

Benefits

• “immediate discussions”, “instant feedback”, “I can ask”,
• “colleagues collaborating in the same room” and “communality”
• “insight from non-testers on how applications are used” “broader look at the application”

Drawbacks

• vague defect descriptions: "Home view does not look good, it should look better".
• Test-sessions require arranging and after work
RQ2 ... participant experiences?

Suitable for testing
• For full features
• Regression related defects
• The exploratory testing flow is more random in the nature so it is not uncommon to make a discovery that would not have been made in actual structured regression testing.

Non-suitable for testing
• functions that require a lot of steps
• long running functions or
• testing that requires backend, device side or remote service configurations
Paper Summary

• Quantitative data of TET
  • High efficiency
  • More usage and usability related problems

• Qualitative findings pros
  • Feedback & Discussions
  • Many defects found
  • Insight from non-testers & Broader view
  • For testing full features

• Qualitative findings cons
  • Requires effort in set-up and after-work
  • Vague defect descriptions
  • Not for tests that take long time to execute
The Effect of Team Exploratory Testing—Experience Report from F-Secure

Paula Raappana, Soili Saukkoriipi
F-Secure Oyj
Helsinki, Finland
{paula.raappana, soili.saukkoriipi}@f-secure.com

Ilkka Tervonen, Mika V. Mäntylä
Department of Information Processing Science
University of Oulu
{ilkka.tervonen, mika.mantyla}@oulu.fi