
Erika S. Mesh\textsuperscript{1}, Scott Hawker\textsuperscript{2}

\textsuperscript{1}Golisano College of Computing and Information Sciences
\textsuperscript{2}Software Engineering Department
Rochester Institute of Technology
Rochester, NY

2013 International Workshop on Software Engineering for Computational Science and Engineering
In Conjunction with ICSE 2013
May 18, 2013
Project Context
  Software Engineering
  SPI for Scientific Software

Research Overview
  Goals and Proposed Model
  Research Questions

Methodology
  Overview
  Data Collection and Analysis
  Preliminary Results

Discussion

References
What is “software” and what is a “software project”? 

- Software projects consist of people, the product they build, and the process used to tie it all together. ¹

SE processes evolve iteratively over time:

- to meet project/product goals
- via a cycle of assessment, planning and enactment.²

---

¹ Jacobson et al. (1999)  
² Humphrey (1984); Pettersson et al. (2008)
Software Process Improvement (SPI) Frameworks

Prescriptive

+ Standardized guidelines for assessment and planning
  - Requires assumptions about the project priorities
Software Process Improvement (SPI) Frameworks

**Prescriptive**

- Standardized guidelines for assessment and planning
  - Requires assumptions about the project priorities

**Supportive**

- Tailored SPI cycle and task plans based on project characteristics
  - Depends on SE expertise
## Software Engineering for Scientists

<table>
<thead>
<tr>
<th>SPI Activity</th>
<th>Traditional SE</th>
<th>Scientific SE</th>
</tr>
</thead>
</table>
| **Assessment** | Technical vs. Management Complexity  
B&T Model  
CMMI KPAs | Product quality  
General process characteristics |
| **Prescriptive Planning** | CMMI Staged Model  
CMMI Template Roadmaps  
Established process models | Checklists  
Case studies and practice summaries  
Generalized models |
| **Supportive Planning** | CMMI Continuous Model  
Agile disciplines SPI approach | |
| **Enactment** | CMMI KPAs  
Miscellaneous texts, tutorials, and classes to teach software development and management skills. | Checklists  
software-carpentry boot camps |

---

3Boehm and Turner (2004); Cannegieter et al. (2008); carpentry (2013); Carver et al. (2007); Heroux and Willenbring (2009); Kelly et al. (2009); Kelly (2008, 2011); Kroll and Maclsaac (2006); Segal (2008); Sletholt and Hannay (2012); Software Engineering Institute (2010)
Can we have consistency and flexibility in scientific SPI?

Scientific SPI Framework (SciSPIF)

*Leverage domain expertise while supplementing software development knowledge*
What do we need to implement SciSPIF?

1. Which traditional SPI planning criteria are applicable to scientific software development projects?

2. What other factors influence priorities and decisions during scientific SPI planning activities?

3. Which scientific SPI planning criteria are applied in decisions about specific SE practices?
Constraints

Answering these questions can be completed via:

1. Experimental studies of controlled scenarios
2. Empirical study of real-world projects

Above all else, we need consistent, logical results clearly connected to the data to ensure:

- Acceptance and use in the scientific community
- Validity of future work to leverage the model
Using Grounded Theory (GT)

Systematic qualitative study of the world

- Hypotheses evolve over time via concurrent, rigorous data collection and analysis.

Data collection and analysis are applied iteratively

4 Corbin and Strauss (1990)
Open Encoding

FALCON is used by an external team of highly knowledgeable and experienced product engineers to assess the potential behavior of new and existing product designs. Validation by the users is done by comparing the output of the software to data from past experiments and a few new experiments. The engineers’ level of experience and expertise is sufficiently high that they can not only identify defects and model deficiencies, but can often identify the source of the defects and needed model improvements. In a very real sense, the users participate constructively and effectively in the development, verification and validation of the code.

FALCON is extensively documented on an internal web-site (approximately 400 Mybytes of HTML files). The documentation consists of descriptions of the physics; the algorithms and models; the input and output; and instructions for executing the code. This documentation has proved to be highly useful.

4.2 HAWK Project

The purpose of the HAWK project was to develop a program to support. Work on

4.3 CONDOR

The purpose of the CONDOR project was to develop a simulator of a family of stresses. The CONDOR project had the earliest antecedents in the entities with stress analysis, time-consuming physical testing. CONDOR is a 1) displacement, 2) momentum, and

5 Carver et al. (2007)
Axial/Selective Encoding

<table>
<thead>
<tr>
<th>People</th>
<th>Sources</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collective Ownership</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Communication</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Coordination</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Culture</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Customer support</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Developer Type</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>Location</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>Process Experience</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Stability</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Stakeholder involvement</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Team Organization</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Team Performance</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Balance</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Trust and confidence</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Team Size</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Large</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Small</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Users</td>
<td>3</td>
<td>26</td>
</tr>
</tbody>
</table>
Number of References in Source Literature

<table>
<thead>
<tr>
<th>Area</th>
<th>Number of References by Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>People</td>
<td>103</td>
</tr>
<tr>
<td>Process</td>
<td>42</td>
</tr>
<tr>
<td>Product</td>
<td>76</td>
</tr>
<tr>
<td>Project</td>
<td>42</td>
</tr>
</tbody>
</table>

---

Carver et al. (2007); Easterbrook and Johns (2009); Kane (2003)
Initial Conclusions

Supports observations from previous work

- Science/software accuracy and the people involved are considered more often than project/process concerns

Reaffirms a need for continuing the project to address research questions

1. Applicability of traditional SPI planning criteria?
2. Other criteria?
3. Correlations?

---

Carver et al. (2007)
Future Work

Refinement via Semi-Structured Interviews

▶ Leverage candidate criteria in semi-structured interviews

Dissemination and Concurrent Data Collection

▶ On-line, interactive version of SciSPIF disseminated to a set of representative research projects
▶ Web-based data collection, journaling and ongoing analysis done by the subjects themselves

Theory Validation

▶ Framework validation via surveys and experiments
Risks

Methodology

- GT based methodologies are useful in empirical SE research, but require discipline \(^8\)
- Highly dependent on the quantity and quality of interviews performed

“So what?”

- Existing work demonstrates a need for further understanding/research of scientific SPI.
- Will it prove to have a significant ROI long term? What if it doesn’t?

\(^8\)Adolph et al. (2008)
Special Thanks To...

Questions? Comments? Want to get involved?

Erika S. Mesh
esm1884@rit.edu
http://people.rit.edu/esm1884
References I


References II


Backup Slides
Sample Construct
Candidate SPI Decision Criteria

Encoded references from published case studies
Construct Classification
Subject Interviews

<table>
<thead>
<tr>
<th>General Background:</th>
<th>“Please tell me a bit about your background.”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Background:</td>
<td>“Describe your general research area, motivations, and goals.”</td>
</tr>
<tr>
<td>Current Research:</td>
<td>“Describe a current or recent research project.”</td>
</tr>
<tr>
<td>Research Software Project:</td>
<td>“Discuss the role of custom computer software in this research project.”</td>
</tr>
<tr>
<td>Software Development Personnel:</td>
<td>“Describe your and your team’s general computing and software development experience.”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Software Product</th>
<th>Definition:</th>
<th>“Describe how you determine what you need the software to do?”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Design:</td>
<td>“What are the primary components of your software?” “How was this sub-division determined?”</td>
</tr>
<tr>
<td></td>
<td>Development:</td>
<td>“Are there specific strategies or skills that you adopted that made development easier? How? How did you learn about this in order to apply it?”</td>
</tr>
<tr>
<td></td>
<td>Quality:</td>
<td>“How do you ensure that the software meets your needs?” “How do you know you can trust your software?”</td>
</tr>
<tr>
<td></td>
<td>Tools and Methods:</td>
<td>“Describe any specific tools or techniques that aid in your software related work. How did you learn about them and why did you choose to use them?”</td>
</tr>
</tbody>
</table>