
Presented by: Jeffrey Carver
Research Questions

• RQ1: What are the characteristics of the artifacts and evaluations contained in the papers?
  • RQ1.1: What Artifacts are evaluated?
  • RQ1.2: What Evaluation Methods are used?
  • RQ1.3: Do papers build on prior work?
  • RQ1.4: What is the relationships between Artifact and Method?

• RQ2: Do the papers contain all the information necessary to support the science of security?
Study Overview

Evaluation Subject Types

Evaluation Subject Source

Evaluation Approach

Completion Rubric
Study Overview

Evaluation Subject Types

Model
Language
Protocol
Process
Theory
Tool

Evaluation Approach

Completion Rubric
Study Overview

Evaluation Subject Types

Evaluation Subject Source

Evaluation Approach

Completion Rubric
Study Overview

Evaluation Subject Types

Empirical Study

Evaluation Approach

Evaluation Subject Source

Proof

Discussion

Completion Rubric
### Study Overview: Empirical Study

<table>
<thead>
<tr>
<th>Participant Type</th>
<th>Study Type</th>
<th>Data Type</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation</td>
<td>Observational</td>
<td>Historical</td>
<td>None</td>
</tr>
<tr>
<td>Humans</td>
<td>Interventional</td>
<td>New</td>
<td>None</td>
</tr>
<tr>
<td>Systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automated</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Study Overview

Evaluation Subject

Objectives

Case Selection

Data Collection

Data Analysis

Threats to Validity

Evaluation Approach

Completion Rubric

Is New?
<table>
<thead>
<tr>
<th>Rubric Item</th>
<th>Yes</th>
<th>Partial</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EM1</strong>: Are the research objectives of the study described? (e.g., goals, questions, hypotheses)?</td>
<td>Clearly defined early in the paper (i.e. not in the results or discussion) and labeled (e.g. in bold, italics, underlined or set apart from the text with labels like Research Question, RQ, Objective)</td>
<td>Included in the text but either in the wrong location or not clearly labeled (see Yes above)</td>
<td>Not present</td>
</tr>
<tr>
<td><strong>EM2</strong>: Is the context of the study described? Does the paper offer details on what is being tried to solve</td>
<td>The paper explicitly defines the context of the study (i.e. the problem background or why it is important to study)</td>
<td>The paper defines some, but not all, of the above</td>
<td>The paper defines none of the above</td>
</tr>
<tr>
<td><strong>EM3</strong>: Are the methods for subject sampling described? (e.g., recruitment/selection process, inclusion/exclusion criteria)?</td>
<td>The paper explicitly describes how the cases were selected, including the rationale for selecting the particular case(s)</td>
<td>N/A</td>
<td>The paper does not explicitly describe how the cases were selected</td>
</tr>
<tr>
<td><strong>EM4</strong>: Are the exclusion criteria?</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>EM5</strong>: Are the analysis procedures (by name) or other analysis method described?</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>EM6</strong>: Are the methods for subject sampling described? (e.g., demographics, specification)?</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>EM7</strong>: Are the characteristics of the sample described? (e.g., mean, std dev, charts or tables to describe accuracy, False positive, False negative etc.)?</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>EM8</strong>: Are the analysis procedures (by name) or other analysis method described?</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>EM9</strong>: Do they discuss and provide reasoning for &quot;why&quot; the results had the given outcome?</td>
<td>There is a discussion of why a particular outcome occurred in the study. Rather than presenting only the results, the authors explain &quot;why&quot; such results were obtained.</td>
<td>N/A</td>
<td>No reasoning for the outcome of the study is given.</td>
</tr>
<tr>
<td><strong>EM10</strong>: Is there a dedicated discussion of the threats to validity to the experiment (i.e., limitations or mitigations)?</td>
<td>There is a separate Threats to Validity Section</td>
<td>Threats to validity are discussed, but not in a separate section</td>
<td>Threats to validity are not discussed</td>
</tr>
</tbody>
</table>
Methodology

128 Papers (CCS)
55 Papers (S&P)

7 Reviewers

Author
Results
RQ1: Paper Characteristics
RQ1.1: Type of Artifact

![Bar chart showing the distribution of types of artifacts in papers. The x-axis represents percentages, ranging from 0 to 100. The y-axis lists the types of artifacts: Tool, Protocol, Model, Process, Theory, and Language. The chart compares two categories: S&P (represented by yellow bars) and CCS (represented by blue bars). The Tool category has the highest percentage for both S&P and CCS, followed by Protocol and Model, with Process, Theory, and Language having lower percentages.](chart.png)
R1: Paper Characteristics

RQ1.3: Building on Prior Work

Authors Elsewhere

Others Modified

Others Not Modified

Authors Here

S&P

CCS
### RQ1: Paper Characteristics

#### RQ1.4: Subject Type x Approach Type

<table>
<thead>
<tr>
<th></th>
<th>Empirical</th>
<th>Proof</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>27</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Tool</td>
<td>104</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Model</td>
<td>23</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Protocol</td>
<td>21</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Theory</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
RQ2: Paper Completeness
Author Feedback

Changed average
Of 2 items/paper

Response Rate
Agreement

- CCS
- S&P
Author Feedback: Changes

No
Partial
Yes

Threats to Validity

Analysis Procedures

Research Objectives

S&P
CCS
Comparison with Prior Results

• Most common Evaluation Subject - Tools

• Most common Evaluation Approach – Empirical Studies

• Little evidence of Replication

• Most studies missing Threats to Validity
The Team

- Morgan Burcham
- Mahran Al-Zyoud
- Jeffrey Carver

- Hongying Du
- Akond Rahman
- Ozgur Kafali
- Laurie Williams

- Mohammed Alsaleh
- Fida Gilani
- Ehab Al-Shaer

- Jun Jiang
Characterizing Scientific Reporting in Security Literature:
An analysis of ACM CCS and IEEE S&P Papers

Jeffrey Carver
carver@cs.ua.edu