System Usability Scale: Application and Interpretation in Military OT&E

Keith Kidder
AFOTEC Det 4
Release Date: 27 Mar 2017
Introduction

• What is the SUS?
• How SUS is administered and scored
• What SUS does and does not tell us
• Practical use of SUS with space program
• Suggested alternatives
• SUS has proven itself to be useful from a consumer software perspective
• Currently there are no existing benchmarks for the interpretation of the SUS numerical score with military systems
• It does not adequately scope the usability attributes of military systems under test
• Let’s consider an alternative for creating meaningful usability ratings for military systems
SUS Background

- Developed in 1986 by John Brooke while at the Digital Equipment Corporation
- "Quick and dirty usability scale"
- Designed for DEC’s ALL-IN-1 software
- Generic usability scale for consumer products
- Meant as a high level assessment of usability
- Not intended to provide diagnostic information
- SUS scores can range from 0-100 (not a percentage)
System Usability Scale


1. I think that I would like to use this system frequently
   1 2 3 4 5

2. I found the system unnecessarily complex
   1 2 3 4 5

3. I thought the system was easy to use
   1 2 3 4 5

4. I think that I would need the support of a technical person to be able to use this system
   1 2 3 4 5

5. I found the various functions in this system were well integrated
   1 2 3 4 5

6. I thought there was too much inconsistency in this system
   1 2 3 4 5

7. I would imagine that most people would learn to use this system very quickly
   1 2 3 4 5

8. I found the system very cumbersome to use
   1 2 3 4 5

9. I felt very confident using the system
   1 2 3 4 5

10. I needed to learn a lot of things before I could get going with this system
    1 2 3 4 5
<table>
<thead>
<tr>
<th>Number</th>
<th>Statement</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I think that I would like to use this system frequently</td>
<td>Like to use</td>
</tr>
<tr>
<td>2</td>
<td>I found the system unnecessarily complex</td>
<td>Complex</td>
</tr>
<tr>
<td>3</td>
<td>I thought the system was easy to use</td>
<td>Easy</td>
</tr>
<tr>
<td>4</td>
<td>I think that I would need the support of a technical person to be able to use this system</td>
<td>Need tech support</td>
</tr>
<tr>
<td>5</td>
<td>I found the various functions in this system were well integrated</td>
<td>Well integrated</td>
</tr>
<tr>
<td>6</td>
<td>I thought there was too much inconsistency in this system</td>
<td>Inconsistent</td>
</tr>
<tr>
<td>7</td>
<td>I would imagine that most people would learn to use this system very quickly</td>
<td>Quick to learn</td>
</tr>
<tr>
<td>8</td>
<td>I found the system very cumbersome to use</td>
<td>Cumbersome</td>
</tr>
<tr>
<td>9</td>
<td>I felt very confident using the system</td>
<td>Confident</td>
</tr>
<tr>
<td>10</td>
<td>I needed to learn a lot of things before I could get going with this system</td>
<td>Too much to learn</td>
</tr>
</tbody>
</table>
SUS Analysis

**System Usability Scale**

1. I think that I would like to use this system frequently
   - [ ] 1 2 3 4 5
   - [x] 5

2. I found the system unnecessarily complex
   - [ ] 1 2 3 4 5
   - [x] 3

3. I thought the system was easy to use
   - [ ] 1 2 3 4 5
   - [x] 3

4. I think that I would need the support of a technical person to be able to use this system
   - [ ] 1 2 3 4 5
   - [x] 4

5. I found the various functions in this system were well integrated
   - [ ] 1 2 3 4 5
   - [x] 4

6. I thought there was too much inconsistency in this system
   - [ ] 1 2 3 4 5
   - [x] 4

7. I would imagine that most people would learn to use this system very quickly
   - [ ] 1 2 3 4 5
   - [x] 4

8. I found the system very cumbersome to use
   - [ ] 1 2 3 4 5
   - [x] 4

9. I felt very confident using the system
   - [ ] 1 2 3 4 5
   - [x] 5

10. I needed to learn a lot of things before I could get going with this system
    - [ ] 1 2 3 4 5
    - [x] 5

**SUS Score:** 70
What SUS does and does not tell us

• SUS has demonstrated good reliability and validity for consumer systems
• SUS measures system usability
• SUS measures system learnability
• SUS does NOT indicate why a system is or isn’t usable
• SUS is NOT a diagnostic tool so it will not identify what needs to be fixed
• A high SUS score does NOT indicate successful mission/task accomplishment
SUS with a SUT

• Real-world test of multi-billion dollar program
• Individual respondents: 111
• Average SUS score: 55 (F)
• Post SUS interviews and test team observations indicated the system user interface did have some room for improvement but was not entirely lacking and allowed for mission accomplishment
• Interviews/observations indicate SUT usability = B-
• Currently there are no existing benchmarks for SUS interpretation with military systems
• The SUS data had limited value for this program
Suggestion 1

- Limit SUS to EOAs/OAs
  - High level view of a growing system’s usability
  - Use SUS as “A quick and dirty usability scale”
  - Concurrently test team will continue detailed scope of specific system usability attributes
Suggestion 2

• Add a 7-pt rating scale as an 11\textsuperscript{th} item to the SUS
  – Use of an adjective rating scale will greatly aid interpreting SUS scores for military systems
  – Similar to work done by A. Bangor, P.T. Kortum, J.T. Miller (2009)

11. Overall, I would rate the usability of this system as:

<table>
<thead>
<tr>
<th>Completely Unacceptable</th>
<th>Moderately Unacceptable</th>
<th>Slightly Unacceptable</th>
<th>Equally Acceptable &amp; Unacceptable</th>
<th>Slightly Acceptable</th>
<th>Moderately Acceptable</th>
<th>Completely Acceptable</th>
</tr>
</thead>
</table>
Suggestion 3

• Create a list of usability attributes specifically tailored to the system under test
  – Gather observations/ratings for each usability attribute
  – Identify problem areas and provide recommendations to system engineers and operational users
  – Identify and preserve well-designed areas of user interface
Conclusion

• Every test is different
  – There is no one-size-fits-all measurement solution
• Rating instruments should be customized for each test
• SUS is one tool in the toolbox
• Are there other suggestions or methods to assess software usability with military systems?

Software Usability Working Group?