

# **Surveys in Test & Evaluation: The Good, The Bad, & The Ugly**

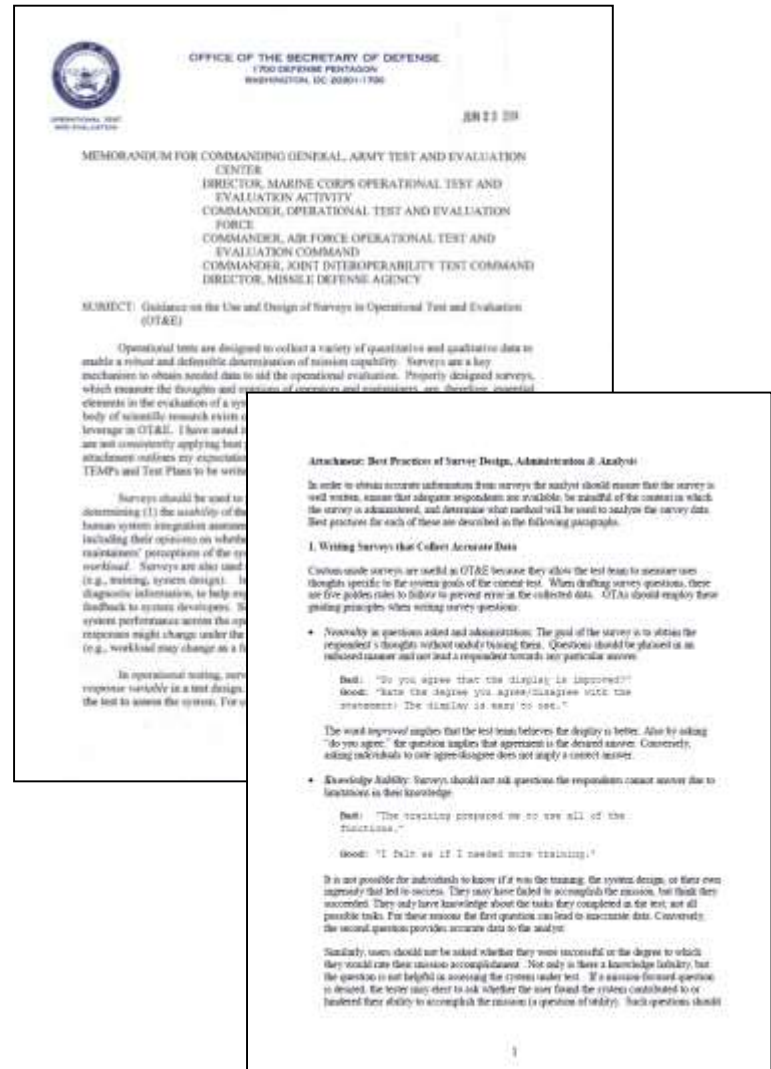
Rebecca A. Grier, Ph.D.  
Institute for Defense Analyses



# DOT&E Guidance on Surveys



- Surveys are an important aspect of DOT&E evaluation of effectiveness and suitability
- Surveys are appropriate for quantitatively measuring operator and maintainer thoughts and opinions
- Use surveys only when appropriate
- Employ best practices for writing and administering surveys
  - Memo provides a best practices guide attachment



## IDA | Survey Purpose & Goal

A systematic collection & analysis of data relating to the thoughts of a population.

- **Not Time:**

“Put your hand on a hot stove for a minute, & it seems like an hour.  
Sit with a pretty girl for an hour, & it seems like a minute.”

- Albert Einstein

- **Not Accuracy:**

		Truth	
		Success	Failure
Belief	Success	☺	!!!
	Failure	☹	☹

Bad Design = Mismatch Between Truth & Belief



3 Mile Island



Vincennes Incident

- **Not Situation Awareness:**

“....There are things we do not know we don't know.” - Donald Rumsfeld

# Examples Where Measurement by Survey Is Not Appropriate

- **“The ship has protective clothing for every crew member.”**
  - Count the protective clothing & compare to number of crew.
- **“Engine exhaust levels in the mission bay do not exceed safety limits ...”**
  - Measure exhaust levels with Portable Emissions Measurement System (PEMS) & compare to safety limits
- **“Temperatures in primary work spaces were adequate.”**
  - Measure by a thermometer & compare to requirements in MIL-STD-1472G
    - e.g., 5.5.2.1.4 part g: *Limited thermal tolerance zones. Where hard physical work is to be required for more than two hours, an environment not exceeding... 25 °C (77 °F) shall be provided. ... shall be decreased 5.0 °C (9.0 °F) for complete chemical protective uniforms, 4.0 °C (7.0 °F) for intermediate clothing systems, and 3.0 °C (5.0 °F) for body armor.*

# IDA | Survey Purpose & Goal

A systematic collection & analysis of data relating to the thoughts of a population.

1. Goal Directed - Not a Fishing Expedition
2. Question Type Determines Data Analysis Possible
3. Consider Respondent Motivation.

Who	Role	Sources of Error
Commissioner	Defines Survey's Purpose & Uses Information from Survey	<ul style="list-style-type: none"><li>• Not Enough Information</li><li>• Wrong Information</li></ul>
Participant/ Respondent	Gives Data	<ul style="list-style-type: none"><li>• Answers Different Question</li><li>• Thinks Too Much</li><li>• Doesn't Think Enough</li></ul>
Analyst	Translates Data Into Information	<ul style="list-style-type: none"><li>• Unable to analyze data</li><li>• Data Aggregation</li></ul>



### Appropriate

- 1. There Isn't an Appropriate Empirical Survey**
- 2. Measure Specific User/Maintainer Thoughts**
  - Specific features/ components
  - Specific issues with regard to CONOPS
- 3. Quantify Observer Ratings**

“A good plan is like a road map: it shows the final destination and usually the best way to get there.”

*H. Stanely Judd*

### Not Appropriate

- 1. Obtain Random Thoughts of Respondents**
  - Interview
- 2. Measure Performance**
  - Time
  - Accuracy via Appropriate Physical Measure
  - Observers
- 3. Measure Requirements**
  - Appropriate Physical Measure
- 4. Measure Situation Awareness**
  - Numerous techniques in Human Factors Literature
  - Salmon et al (2006) for review

Gather Requirements

Draft Questions

Format Survey

Evaluate

Administer

Analyze

Report

Make Everything As Simple As Possible, But Not Simpler  
– Albert Einstein



Identifier	Item	Response
1.	How many parts of a question are there?	
1	2	<u>3</u> 4 5

**Knowledge Liability:** Respondents Have Enough Information to Answer the Question

**Singularity:** Only 1 Idea Per Item

**User Friendly:** Items Do Not Require a Lot of Thought or Interpretation (e.g., short, clear, specific)

**Neutrality:** Items Do Not Imply Value Judgments  
Items Are Not Emotionally Charged

**Independence:** Responses Will Not Affect Responses to Other Questions (Branching, Redundancy, etc...)  
- Especially important when aggregating responses  
- Item Response Theory & Internal Reliability

# **1. Is the training materials complete?**

# 1. Is the training materials **complete**?

Knowledge Liability

User Friendly – grammar

**1. I felt as if I needed more training.**

**2. Rate the adequacy of air-search radar & combat system to correctly decide to engage/not engage each track per Combat System Engagement Doctrine.**

2. Rate the adequacy of air-search radar & combat system to **correctly** decide to engage/not engage each track per **Combat System Engagement Doctrine**.

Knowledge Liability

2. Rate the adequacy of air-search radar & combat system to **correctly** decide to engage/not engage **each** track per **Combat System Engagement Doctrine.**

Knowledge Liability

Singularity

User Friendly – 22 words!

2. I trusted the system's engagement decisions.

### **3. Rate the overall usefulness of the report: Accuracy**

Completely unacceptable      Largely unacceptable      Somewhat unacceptable      Somewhat acceptable      Largely acceptable      Completely acceptable

### 3. Rate the overall usefulness of the report: **Accuracy**

Completely      Largely      Somewhat      Somewhat      Largely      Completely  
unacceptable    unacceptable    unacceptable    acceptable    acceptable    acceptable

**Knowledge Liability**



### 3. Rate the overall **usefulness** of the report: **Accuracy**

Completely unacceptable    Largely unacceptable    Somewhat unacceptable    Somewhat acceptable    Largely acceptable    Completely acceptable

**Knowledge Liability**

User Friendly

What is being rated: report acceptability, usefulness, or accuracy?

**Strongly Agree**

**Somewhat Agree**

**Slightly Agree**

**Slightly Disagree**

**Somewhat Disagree**

**Strongly Disagree**

**4. The SSO functionality increases my productive time within the clinic by reducing the amount of time I spend logging into different applications when documenting the healthcare provided.**

**Strongly  
Agree**

**Somewhat  
Agree**

**Slightly  
Agree**

**Slightly  
Disagree**

**Somewhat  
Disagree**

**Strongly  
Disagree**

4. The SSO functionality increases my productive time within the clinic **by reducing the amount of time I spend logging into different applications** when documenting the healthcare provided.

Neutrality

User Friendly – 27 words!

4. The SSO function is useful.

Strongly  
Agree

Somewhat  
Agree

Slightly  
Agree

Slightly  
Disagree

Somewhat  
Disagree

Strongly  
Disagree

**5. Based on your responses above,  
rate the acceptability of the system.**

**5. Based on your responses above,  
rate the acceptability of the system.**

Independence

User Friendly

(acceptability in terms of what)

**5. I would like to use this system to  
accomplish the mission.**

Knowledge Liability Singularity User Friendly Neutrality Independence

- Accurate
- Timely
- Situation Awareness
- Effective
- Efficient
- n/a
- And
- Each
- All
- Never
- None
- Better
- Easier
- Improved
- Based on
- If
- Considering

### Is your workspace acceptable?

1. It is perfect, I wouldn't change anything.
2. It is the best space available, but I would change things if I could.
3. I am able to function in my workspace, but better spaces are available.
4. I am unable to function in my workspace.

### My workspace is acceptable.

Strongly Disagree	Somewhat Disagree	Slightly Disagree	Slightly Agree	Somewhat Agree	Strongly Agree
-------------------	-------------------	-------------------	----------------	----------------	----------------

### Rate your workspace.

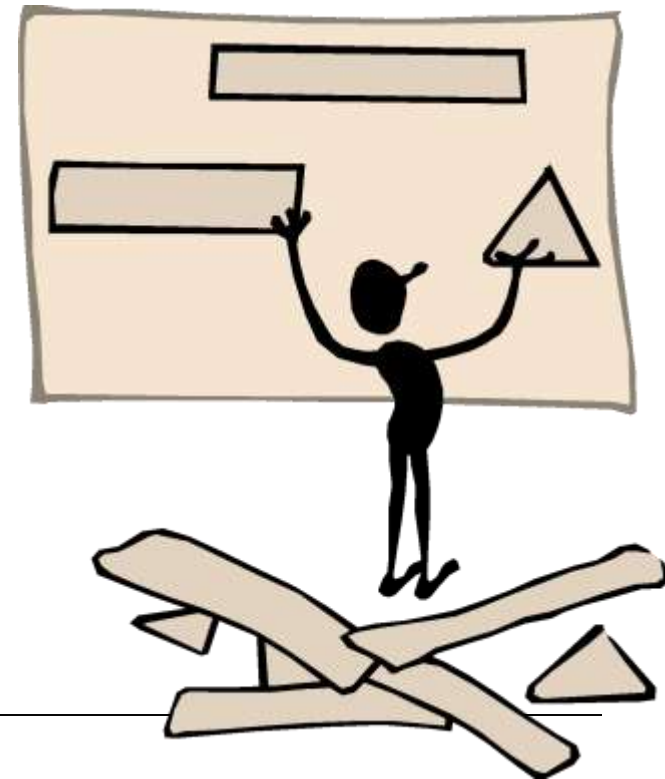
Completely Acceptable	Reasonably Acceptable	Barely Acceptable	Barely Unacceptable	Reasonably Unacceptable	Completely Unacceptable
-----------------------	-----------------------	-------------------	---------------------	-------------------------	-------------------------

**Improved Consistency Between Respondents = Higher Reliability**  
**More Sensitivity & Specificity = Higher Validity**

- **Provide Brief Clear Introduction**
- **Keep the Survey as Brief as Possible to Obtain the Necessary Information**
- **Group Questions by Topic & Response Type**
- **Logically Ordered Questions**
  - Begin with interesting items that are clearly connected to the goals of the survey
  - Follow order of events
  - Within Group: start generally and get more specific
- **Minimize Open Responses**
  - 1 open ended at end of each group/section
  - No more than 4 open ended at end of survey

*1. Minimize the Respondents' Burden  
(Perceived & Real)*

*2. Maximize the Importance of the  
Respondents' Contribution*





## More Than 3 Pages

FOR OFFICIAL USE ONLY (when filled in)

Survey Sheet ES-2  
Page 2 of 10

2. The wearing of PPE (e.g., gloves, boots, etc.) does not preclude operation of control devices.

STRONGLY AGREE	AGREE	SOMEWHAT AGREE	SOMEWHAT DISAGREE	DISAGREE	STRONGLY DISAGREE
1	2	3	4	5	6

If you rated 3 through 6, explain:

---



---

3. The size and setup of the JHSV bridge is adequate for all operations.

STRONGLY AGREE	AGREE	SOMEWHAT AGREE	SOMEWHAT DISAGREE	DISAGREE	STRONGLY DISAGREE
1	2	3	4	5	6

If you rated 3 through 6, explain:

---



---

4. The external communication (i.e., VHF bridge-to-bridge, etc.) and information systems (i.e., NIPRNET, SIPRNET, etc.) provide the tools for effective conduct of the mission.

STRONGLY AGREE	AGREE	SOMEWHAT AGREE	SOMEWHAT DISAGREE	DISAGREE	STRONGLY DISAGREE
1	2	3	4	5	6

If you rated 3 through 6, explain:

---



---

FOR OFFICIAL USE ONLY

## 1 Page

FOR OFFICIAL USE ONLY (when filled in)

	Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1. The JHSV bridge design provides adequate situational awareness for safe navigation.	1	2	3	4	5	6
2. The wearing of PPE (e.g., gloves, boots, etc.) does not preclude operation of control devices.	1	2	3	4	5	6
3. The size and setup of the JHSV bridge is adequate for all operations.	1	2	3	4	5	6
4. The external communication (i.e., VHF bridge-to-bridge, etc.) and information systems (i.e., NIPRNET, SIPRNET, etc.) provide the tools for effective conduct of the mission.	1	2	3	4	5	6
5. The data transfer rate supporting LOS and beyond LOS communications was effective in all operations.	1	2	3	4	5	6
6. JHSV navigation systems (GPS, ECDIS-M, etc.) provide adequate support for voyage and mission planning.	1	2	3	4	5	6
7. JHSV Communication, Electronics, and Navigation systems provide safe and accurate navigation of the vessel.	1	2	3	4	5	6
8. JHSV engineering propulsion systems (e.g., engine, water jet, etc.) adequately support ship's operations during transit.	1	2	3	4	5	6
9. JHSV engineering power systems (e.g., generators, switchboards) adequately support ship's operations during transit.	1	2	3	4	5	6
10. The JHSV can moor to the pier/quay wall without tug support.	1	2	3	4	5	6
11. The rub rail interfaces adequately with piers.	1	2	3	4	5	6
Comments:						

FOR OFFICIAL USE ONLY

## Performance Data

What: time & accuracy

## Subject Matter Expert Observation

How: actions taken, moments of frustration, etc.

## User Surveys

Why: usability, workload, thoughts about specific design features, etc.

- Questions known ahead to be appropriate for test
- Finite set of concise responses possible

## User Interviews

Why: non-specific thoughts

- Questions in response to rare or unexpected test events
- Infinite number of possible responses
- Possible responses are long

Effectiveness & Suitability



# Questions?

