QUICK REFERENCE GUIDES

Evaluators Can’t Live Without
INTRODUCTION

Kelly Robertson

Lyssa Wilson Becho

This material is based upon work supported by the National Science Foundation under grant no. 1600992 and 1841783. The content reflects the views of the authors and not necessarily those of NSF.
AGENDA

1. Quick-reference-what?
2. How we learned about QRGs in use
3. Top quick reference guides
4. Top guides you likely missed
5. Full list of evaluation QRGs
What is a Quick Reference Guide?

Practical reference or resource materials

- Checklist
- Worksheet
- Infographic
- Cheat sheet
- Table or page in a book
WHY DO QUICK REFERENCE GUIDES MATTER?

- Learn new tools
- Solve problems in action
- Increase evaluation quality
- Open access
WHAT DID WE DO?
Do you use quick reference guides? How? Which ones? (n = 79)
WHAT DID WE DO?

2

- Do you use quick reference guides? How? Which ones? (n = 79)
- How useful are these quick reference guides (from list)? (n = 45)
WHAT DID WE LEARN?

78% (n = 45)
WHAT DID WE LEARN?

53% to teach others
51% to guide work
29% to enhance knowledge
27% to help clients

(n = 45)
TOP GUIDES

Overall by Categories
TOP GUIDES

Evaluation Planning
Evaluation Design
Data Collection
Data Analysis
Evaluation Reporting
EVALUATION FLASH CARDS

Michael Patton (2017)
**Evaluation Flash Cards**

**Purpose**
Introduce core concepts of evaluation in an easily accessible and retrievable format.

**Intended Use**
Information refresher

**Topics Covered**

- Evaluative Thinking
- Evaluation Questions
- Logic Models
- Theory of Change
- Evaluation vs Research
- Dosage
- Disaggregation
- Changing Denominators, Changing Rates
- SMART Goals
- Distinguishing Outcomes From Indicators
- Performance Targets
- Qualitative Evaluation
- Triangulation Through Mixed Methods
- Important and Rigorous Claims of Effectiveness
- Accountability Evaluation
- Formative Evaluation
- Summative Evaluation
- Development Evaluation
- The IT Questions
- Fidelity or Adaptation
- High-Quality Lessons Learned
- Evaluation Quality Standards
- Complete Evaluation Reporting
- Unitization-Focused Evaluation
- Distinguish Different Kinds of Evidence

Michael Patton (2017)
EVALUATION FLASHCARDS

Michael Patton (2017)

About Concept

Example

Review

Bottom Line
EVALUATION DATA MATRIX TEMPLATE

EvaluATE (2017)

Evaluation Data Matrix Template

This material was prepared with support by the National Science Foundation under grant number 1823992. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of NSF.

An evaluation plan should include a clear description of what data will be collected, from what sources and how, by whom, and when, as well as how the data will be analyzed. Placing this information in a matrix helps ensure that there is a viable plan for collecting all the data necessary to answer each evaluation question and that all collected data will serve a specific, intended purpose. The table below may be copied into another document, such as a grant proposal, and edited/expanded as needed. An example is provided on the next page.

**Evaluation Question**

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Data Source and Methods</th>
<th>Responsible Party</th>
<th>Timing</th>
<th>Analysis Plan</th>
<th>Interpretation</th>
</tr>
</thead>
</table>

**DEFINITIONS**

Evaluation Questions are overarching questions about a project's quality or impact. The number of evaluation questions depends on the scope and purpose of the evaluation. A typical evaluation may include:

- **Who:** who will benefit from the project? How many people will participate?
- **What:** what will be measured? What are the targets for improvement?
- **How:** how will the evaluation be conducted? What methods will be used to collect data?
- **Why:** why is this evaluation important? What are the implications of the results?

Data Sources are the entities from which data will be collected. Typical data sources for NTC evaluations include project partners, stakeholders, funders, project managers, and community organizations.

**Data Collection Methods** are the means by which information will be garnered. Typical methods include surveys, focus groups, interviews, observations, and institutional data sources.

**Responsible Parties** are the individuals or organizations tasked with collecting the needed information. In many cases, data collection requires cooperation among multiple entities. For example, an external evaluator may be responsible for administering a survey, but a member of the project staff may need to supply the contact information.

**Timing** identifies when and how frequently data will be collected (e.g., at events, quarterly, annually). It is essential to clearly identify the timeline for data collection so that all stakeholders are aware of the timeline for reporting purposes and decision making.

**Analysis Plan** is how the quantitative and qualitative data will be summarized and interpreted. It is crucial that the analysis plan is developed early in the planning process.

**Interpretation** is how the analyzed data will be used to reach conclusions related to the evaluation questions.
EVALUATION DATA MATRIX TEMPLATE

EvaluATE (2017)

Purpose
Organize evaluation indicators, collection source, analysis, and interpretation.

Intended Use
Provide structure for evaluation design

Topics Covered
• Indicators
• Data sources and methods
• Responsible party
• Timing
• Analysis plan
• Interpretation
Outline for matrix structure

Explanation of each column

EVALUATION DATA MATRIX TEMPLATE
EvaluATE (2017)
# Evaluation Data Matrix Template

**EvaluATE (2017)**

## Example

<table>
<thead>
<tr>
<th>Evaluation Question</th>
<th>Indicator</th>
<th>Data Source and Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Definitions**

- **Evaluation Questions**: overarching questions on the scope and purpose of implementation and outcomes.
- **Indicators**: specify, in plain language, the evaluation questions, outcomes, and outputs.
- **Data Sources**: the entities being monitored, students, participants, websites, usage statistics, and the like.
- **Data Collection Methods**: surveys, interviews, observation.
- **Responsible Parties**: the entities responsible for selecting and collecting data.
- **Timing**: identifies when and how the data will be collected.
- **Analysis Plan**: how the data will be analyzed.

---

**Table**

<table>
<thead>
<tr>
<th>Evaluation Question</th>
<th>Indicator</th>
<th>Data Source and Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Matrix**

<table>
<thead>
<tr>
<th>Evaluation Question</th>
<th>Indicator</th>
<th>Data Source and Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure**

![Figure Description](image-url)
### Likert-Type Scale Response Anchors

**Vegias (2006)**

<table>
<thead>
<tr>
<th>Level of Accountability</th>
<th>Level of Support/Opposition</th>
<th>Level of Probability</th>
<th>Level of Agreement</th>
<th>Level of Uncertainty</th>
<th>Level of Participation</th>
<th>Level of Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>My beliefs</td>
<td>My beliefs</td>
<td>My beliefs</td>
<td>My beliefs</td>
<td>My beliefs</td>
<td>My beliefs</td>
<td>My beliefs</td>
</tr>
</tbody>
</table>

**Level of Appropriateness**

- 1. Absolutely inappropriate
- 2. Inappropriate
- 3. Slightly inappropriate
- 4. Neutral
- 5. Slightly appropriate
- 6. Appropriate
- 7. Absolutely appropriate

**Level of Importance**

- 1. Not at all important
- 2. Little important
- 3. Slightly important
- 4. Important
- 5. Moderately important
- 6. Very important
- 7. Extremely important

**Level of Agreement**

- 1. Strongly disagree
- 2. Somewhat disagree
- 3. Neither agree nor disagree
- 4. Somewhat agree
- 5. Strongly agree

**Knowledge of Actor**

- 1. Novice
- 2. Beginner
- 3. Advanced but inexperienced
- 4. Neutral
- 5. Experienced
- 6. Usually true
- 7. Always true

**Reflect on Me?**

- 1. Very unsure of me
- 2. Somewhat unsure of me
- 3. Neutral
- 4. Slightly sure of me
- 5. True of me
- 6. Usually true of me
- 7. Always true of me

**Likert-Type Scale Response Anchors Citation**

Likert-Type Scale Response Anchors

Vegias (2006)

Purpose
Examples of likert scale labels for a variety of constructs.

Intended Use
Inspiration during survey item development

Topics Covered
- Acceptability
- Appropriateness
- Importance
- Agreement
- Knowledge
- Beliefs
- Priority
- Concern
- Frequency
- Frequency of Use
- Level of Problem
- Affect
- Level of Consideration
- Level of Support
- Agreement
- Desirability
- Participation
- Reflection
- Frequency of Use
- Frequency of Consideration
- Frequency of Interaction
- Frequency of Use
- Frequency of Support
- Frequency of Agreement
- Frequency of Knowledge
- Frequency of Beliefs
- Frequency of Priority
- Frequency of Concern
## Likert-Type Scale Response Anchors

**Vegias (2006)**

<table>
<thead>
<tr>
<th>Measurement construct</th>
<th>Example anchor labels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Likert-Type Scale</strong></td>
<td><strong>Response Anchors</strong></td>
</tr>
<tr>
<td>Level of Accountability</td>
<td>My beliefs</td>
</tr>
<tr>
<td>1. Totally unacceptable</td>
<td>1. Very untrue of what I believe</td>
</tr>
<tr>
<td>2. Unacceptable</td>
<td>2. Untrue of what I believe</td>
</tr>
<tr>
<td>5. Slightly acceptable</td>
<td>5. Somewhat true of what I believe</td>
</tr>
<tr>
<td>6. Acceptable</td>
<td>6. True</td>
</tr>
<tr>
<td>Level of Appropriateness</td>
<td>Priority</td>
</tr>
<tr>
<td>1. Absolutely inappropriate</td>
<td>1. Not a priority</td>
</tr>
<tr>
<td>2. Inappropriate</td>
<td>2. Low priority</td>
</tr>
<tr>
<td>3. Slightly inappropriate</td>
<td>3. Somewhat important</td>
</tr>
<tr>
<td>5. Slightly appropriate</td>
<td>5. Moderate Priority</td>
</tr>
<tr>
<td>6. Appropriate</td>
<td>6. High priority</td>
</tr>
<tr>
<td>7. Absolutely appropriate</td>
<td>7. Essential priority</td>
</tr>
<tr>
<td>Level of Importance</td>
<td>Level of Concerns</td>
</tr>
<tr>
<td>1. Not at all important</td>
<td>1. Not at all concerned</td>
</tr>
<tr>
<td>2. Less important</td>
<td>2. Slightly concerned</td>
</tr>
<tr>
<td>3. Slightly important</td>
<td>3. Somewhat concerned</td>
</tr>
<tr>
<td>5. Slightly more important</td>
<td>5. Extremely concerned</td>
</tr>
<tr>
<td>6. Very important</td>
<td>6. Extremely important</td>
</tr>
<tr>
<td>7. Extremely important</td>
<td>7. Essential importance</td>
</tr>
<tr>
<td>Level of Agreement</td>
<td>Priority Level</td>
</tr>
<tr>
<td>1. Strongly disagree</td>
<td>1. Not a priority</td>
</tr>
<tr>
<td>2. Disagree</td>
<td>2. Low priority</td>
</tr>
<tr>
<td>3. Neither agree nor disagree</td>
<td>3. Moderate priority</td>
</tr>
<tr>
<td>5. Slightly agree</td>
<td>5. Essential</td>
</tr>
<tr>
<td>6. Agree</td>
<td>6. Very important</td>
</tr>
<tr>
<td>7. Strongly agree</td>
<td>7. Essential importance</td>
</tr>
<tr>
<td>Knowledge of Actor</td>
<td>Level of Problems</td>
</tr>
<tr>
<td>1. Not be</td>
<td>1. Not at all a problem</td>
</tr>
<tr>
<td>2. Unclear</td>
<td>2. Minor problem</td>
</tr>
<tr>
<td>5. Somewhat sure</td>
<td>5. Major problem</td>
</tr>
<tr>
<td>6. Usually true</td>
<td>6. Absolutely true</td>
</tr>
<tr>
<td>7. Always true</td>
<td>7. Very true of the</td>
</tr>
<tr>
<td>Reflect on X</td>
<td>Level of Consideration</td>
</tr>
<tr>
<td>1. No effect</td>
<td>1. Not important</td>
</tr>
<tr>
<td>2. Minor effect</td>
<td>2. Somewhat important</td>
</tr>
<tr>
<td>3. Moderate effect</td>
<td>3. Most important</td>
</tr>
<tr>
<td>4. Major effect</td>
<td>4. Absolutely important</td>
</tr>
<tr>
<td>Frequency of the</td>
<td>Frequency of the</td>
</tr>
<tr>
<td>1. Never</td>
<td>1. Very high</td>
</tr>
<tr>
<td>2. Occasionally</td>
<td>2. High</td>
</tr>
<tr>
<td>3. Sometimes</td>
<td>3. Moderate</td>
</tr>
<tr>
<td>4. Always</td>
<td>4. Low</td>
</tr>
<tr>
<td>5. No problem</td>
<td>5. Very low</td>
</tr>
</tbody>
</table>

**Likert-Type Scale Response Anchors**

Citation

YOU’RE INVITED TO A DATA PARTY

Hutchinson (2016)
YOU’RE INVITED TO A DATA PARTY

Hutchinson (2016)

Purpose
Summarize the Who, What, and When of a Data Party

Intended Use
Plan a data interpretation session

Topics Covered
• Who should be at your Data Party
• What is a Data Party
• When you should hold it
• Why you should hold a Data Party
• Where to hold it
• How to ask reflective questions
You’re invited to a Data Party!

Answer common questions

Example of reflective questions to ask

Hutchinson (2016)
### Data Visualization Checklist by Stephanie Evergreen & Ann K. Emery

May 2016

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling the most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. n/a should not be used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. If the guidelines has been broken intentionally to make a point, rate it n/a and deduct those points from the total possible. Refer to the Data Visualization Anatomy Chart on the last page for guidance on vocabulary and the Resources at the end for more details.

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
<td></td>
</tr>
<tr>
<td>Graphs don’t contain much text, so existing text must encapsulate your message and pack a punch.</td>
<td></td>
</tr>
<tr>
<td>6-12 word descriptive title is left-justified in upper left corner</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. Rather than a generic phrase, use a descriptive sentence that encapsulates the graph’s finding or “so what?” Western cultures start reading in the upper left, so locate the title there.</td>
<td></td>
</tr>
<tr>
<td>Subtitle and/or annotations provide additional information</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Subtitles and annotations (call-out text within the graph) can add explanatory and interpretive power to a graph. Use them to answer questions a viewer might have or to highlight specific data points.</td>
<td></td>
</tr>
<tr>
<td>Text size is hierarchical and readable</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than axis labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen.</td>
<td></td>
</tr>
<tr>
<td>Text is horizontal</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points. Consider switching graph orientation (e.g., from column to bar chart) to make text horizontal.</td>
<td></td>
</tr>
<tr>
<td>Data are labeled directly</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars and next to lines). Eliminate/embel legends when possible because eye movement back and forth between the legend and the data can interrupt the brain’s attempts to interpret the graph.</td>
<td></td>
</tr>
<tr>
<td>Labels are used sparingly</td>
<td>2 1 0  n/a</td>
</tr>
<tr>
<td>Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis. Do not add numeric labels <em>and</em> use a y-axis scale, since this is redundant.</td>
<td></td>
</tr>
</tbody>
</table>
DATA VISUALIZATION CHECKLIST

Evergreen & Emery (2016)

Purpose
Guide for the development of high impact data visualizations.

Intended Use
Rating graphs and charts

Topics Covered
- Text
- Arrangement
- Color
- Lines
- Overall

### Data Visualization Checklist

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling the most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. n/a should not be used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. If the guidelines have been broken intentionally to make a point, rate it n/a and deduct those points from the total possible. Refer to the Data Visualization Anatomy Chart on the last page for guidance on vocabulary and the resources at the end for more details.

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-12 word descriptive title is left-justified in upper left corner</td>
<td>2/1</td>
</tr>
<tr>
<td>Short titles enable readers to comprehend takeaway messages even while quickly skimming the graph. Rather than a generic phrase, use a descriptive sentence that encapsulates the graph's finding or &quot;so what?&quot; Western cultures start reading in the upper left, so locate the title here.</td>
<td>n/a</td>
</tr>
<tr>
<td>Subtitle and/or annotations provide additional information</td>
<td>2/1</td>
</tr>
<tr>
<td>Subtitles and annotations (call-out text within the graph) can add explanatory and interpretive power to a graph. Use them to answer questions a viewer might have or to highlight specific data points.</td>
<td>n/a</td>
</tr>
<tr>
<td>Text size is hierarchical and readable</td>
<td>2/1</td>
</tr>
<tr>
<td>Titles are in a larger size than subtitles or annotations, which are larger than labels, which are larger than source information. The smallest text - axis labels - are at least 9 point font size on paper, at least 20 on screen.</td>
<td>n/a</td>
</tr>
<tr>
<td>Text is horizontal</td>
<td>2/1</td>
</tr>
<tr>
<td>Titles, subtitles, annotations, and data labels are horizontal (not vertical or diagonal). Line labels and axis labels can deviate from this rule and still receive full points. Consider switching graph orientation (e.g., from column to bar chart) to make text horizontal.</td>
<td>n/a</td>
</tr>
<tr>
<td>Data are labeled directly</td>
<td>2/1</td>
</tr>
<tr>
<td>Position data labels near the data rather than in a separate legend (e.g., on top of or next to bars and next to lines). Eliminate embedded legends when possible because eye movement back and forth between the legend and the data can interrupt the brain's attempts to interpret the graph.</td>
<td>n/a</td>
</tr>
<tr>
<td>Labels are used sparingly</td>
<td>2/1</td>
</tr>
<tr>
<td>Focus attention by removing the redundancy. For example, in line charts, label every other year on an axis. Do not add numeric labels <em>and</em> use a y-axis scale, since this is redundant.</td>
<td>n/a</td>
</tr>
</tbody>
</table>
### Data Visualization Checklist

This checklist is meant to be used as a guide for the development of high impact data visualizations. Rate each aspect of the data visualization by circling the most appropriate number, where 2 points means the guideline was fully met, 1 means it was partially met, and 0 means it was not met at all. n/a should not be used frequently, but reserved for when the guideline truly does not apply. For example, a pie chart has no axes lines or tick marks to rate. If the guidelines have been broken intentionally to make a point, rate it n/a and deduct those points from the total possible. Refer to the Data Visualization Anatomy Chart on the last page for guidance on vocabulary and the resources at the end for more details.

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong> Way to reduce visual clutter.</td>
<td></td>
</tr>
<tr>
<td>- Use short descriptive titles, giving your message and pack a punch.</td>
<td></td>
</tr>
<tr>
<td>- 6-12 word descriptive title is left justified in upper left corner.</td>
<td>2</td>
</tr>
<tr>
<td>- Subtitle and additional information provide additional information.</td>
<td>2</td>
</tr>
<tr>
<td>- Text size is hierarchical and readable.</td>
<td>2</td>
</tr>
<tr>
<td>- Titles are in a larger size than subtitles or annotations, which are larger than axis labels, which are larger than source information.</td>
<td>2</td>
</tr>
<tr>
<td>- Text is horizontal.</td>
<td>2</td>
</tr>
<tr>
<td>- Data are labeled directly.</td>
<td>2</td>
</tr>
<tr>
<td>- Labels are used sparingly.</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Explanation of aspects

- **Data visualization guidelines with explanations**
- **Ratings**
Data Visualization Checklist
Evergreen & Emery (2016)

Visual comparison
GUIDES YOU LIKELY MISSED

- Evaluation Planning
- Evaluation Design
- Data Collection
- Data Analysis
- Evaluation Reporting
CHECKLIST OF KEY CONSIDERATIONS FOR DEVELOPMENT OF PROGRAM LOGIC MODELS

MacDonald (2018)
CHECKLIST OF KEY CONSIDERATIONS FOR DEVELOPMENT OF PROGRAM LOGIC MODELS

MacDonald (2018)

Purpose
Support the development of logic models.

Intended Use
Provide guidance or refresh knowledge on creating logic models

Topics Covered
• Scope
• Basic components
• Underlying logic of program
• Key items to include
• Accompanying narrative
Checklist of Key Considerations for Development of Program Logic Models

MacDonald (2018)

Steps for development

Practical guidance

1. Determine the Scope of the Logic Model

Steps and practical guidance for developing a program logic model. The checklist includes key considerations for program development.

2. Include the Main Components of a Logic Model

Additional steps and considerations for developing a logic model.

Checklist of Key Considerations for Development of Program Logic Models
Goldie MacDonald

The checklist provides a framework for developing a program logic model, including key considerations for program development. The checklist is designed to help developers ensure that their logic models are comprehensive and effective.
GET TO KNOW YOUR EVALUATION STAKEHOLDERS

Hutchison
GET TO KNOW YOUR EVALUATION STAKEHOLDERS

Hutchison

Purpose
Provide ideas of questions to ask in meeting with evaluation stakeholders.

Intended Use
Facilitate conversation with stakeholders

Topics Covered
• Questions for evaluation stakeholders
GET TO KNOW YOUR EVALUATION STAKEHOLDERS

Example questions

You’ve identified your potential stakeholders. Now what? Get their input by asking some of the following questions...

- What do they perceive the purpose of your program is?
- What is their current opinion of the program?
- What concerns, if any, do they have about the program?
- What influences their opinion of the program, whom their opinions generally?

What have they heard about the proposed program evaluation?

What advice do they think are important to address for the evaluation?

What do they say they will learn from the evaluation?

How much progress do they think is reasonable to expect for this program in the timeline?

What conditions, if any, do they have with the program evaluation?

What financial or emotional interest do they have on the outcome of the evaluation?

Is it meeting an agenda?

If they are not likely to be positive what will win them around or if you are unlikely to win them around, how will you manage their opposition?

How available are they to participate in the evaluation process?

What resources (e.g., time, funds, evaluation expertise, access to respondents, access to policymakers, etc.) do they currently have?

What support do you want from them?

What are the political implications of their involvement in the evaluation?

How will they use the results of this evaluation?

What decisions are going to be made by whom, and why?

How can you best meet their communication needs?

What information do they want from you?

How do they want to receive this information? What is the best way of communicating with them?
Collecting Race and Ethnicity Data from Students and Staff Using the New Categories

Institutions MUST give students and staff the opportunity to self-report their race and ethnicity. Students and staff do NOT have to respond. Institutions MUST use a 2-part question to collect these data. The first part of the question collects ethnicity, and the second part of the question collects race. The questions must be presented in this order:

1. Are you Hispanic or Latino?
2. Select one or more of the following races:
   - American Indian or Alaska Native
   - Asian
   - Black or African American
   - Native Hawaiian or Other Pacific Islander
   - White

Students and staff MUST always be shown both parts of the question, regardless of their answer to either part.

In the second part of the question, the wording MUST read “...one or more...” instead of alternatives such as “all that apply.”

Institutions may NOT present the following choices (or any variations thereof) to students and staff:
- Unknown
- Refuse or decline to respond
- None of the above
- Other
- Nonresident alien

Institutions MAY collect subcategories of the 6 race and ethnicity categories presented in the example 2-part question above.
COLLECTING RACE AND ETHNICITY DATA FROM STUDENTS AND STAFF USING THE NEW CATEGORIES

Purpose
Guide for the development of questions to collect demographic data.

Intended Use
For programs and institutions that need to align data collection with NCES

Topics Covered
• Guidelines for questions
• Reporting missing or messy data to NCES
Collecting Race and Ethnicity Data from Students and Staff Using the New Categories

Separate questions for race and ethnicity

How not to ask questions
### CHOOSING THE CORRECT STATISTICAL TEST IN SAS, STATA, SPSS, AND R

<table>
<thead>
<tr>
<th>Number of</th>
<th>Nature of Independent Variables</th>
<th>Nature of</th>
<th>Test(s)</th>
<th>How to SAS</th>
<th>How to Stata</th>
<th>How to SPSS</th>
<th>How to R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent</td>
<td>Variables 1</td>
<td>IV(s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 IVs (1 population)</td>
<td></td>
<td>interval &amp; normal</td>
<td>one-sample t-test</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td>ordinal or interval</td>
<td></td>
<td>one-sample median</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>categorical (2 categories)</td>
<td></td>
<td>binomial test</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>categorical</td>
<td></td>
<td>Chi-square goodness-of-fit</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>1 IV with 2 levels (Independent groups)</td>
<td></td>
<td>interval &amp; normal</td>
<td>2 Independent sample t-test</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td>ordinal or interval</td>
<td></td>
<td>Wilcoxon-Mann Whitney test</td>
<td>SAS</td>
<td>Stat</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
</tbody>
</table>
## Purpose
Increase appropriateness of statistical analyses.

## Intended Use
Identifying and coding analysis

## Topics Covered
- Type of inferential test
- Code for conducting test in SAS, Stata, SPSS, and R

<table>
<thead>
<tr>
<th>Number of Dependent Variables</th>
<th>Nature of Independent Variables</th>
<th>Nature of Dependent Variable(s)</th>
<th>Test(s)</th>
<th>How to SAS</th>
<th>How to Stata</th>
<th>How to SPSS</th>
<th>How to R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 IVs (1 population)</td>
<td>interval &amp; normal</td>
<td>one-sample t-test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ordinal or interval</td>
<td>one-sample median</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>categorical (2 categories)</td>
<td>binomial test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>categorical</td>
<td>Chi-square</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td>1 IV with 2 levels (Independent groups)</td>
<td>interval &amp; normal</td>
<td>2 independent sample t-test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ordinal or interval</td>
<td>Wilcoxon-Mann Whitney test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
</tbody>
</table>
Choosing the Correct Statistical Test in SAS, Stata, SPSS, and R

Leeper

Identifies correct test for data

Provides example code

<table>
<thead>
<tr>
<th>Number of Dependent Variables</th>
<th>Nature of Independent Variables</th>
<th>Nature of Dependent Variable(s)</th>
<th>Test(s)</th>
<th>How to SAS</th>
<th>How to Stata</th>
<th>How to SPSS</th>
<th>How to R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 IVs (1 population)</td>
<td></td>
<td>interval &amp; normal</td>
<td>one-sample t-test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ordinal or interval</td>
<td>one-sample median</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td>Categorical (2 categories)</td>
<td></td>
<td>binomial test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>Categorical</td>
<td></td>
<td>Chi-square goodness-of-fit</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>1 IV with 2 levels</td>
<td>(Independent groups)</td>
<td>Interval &amp; normal</td>
<td>2 Independent sample t-test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ordinal or interval</td>
<td>Wilcoxon-Mann Whitney test</td>
<td>SAS</td>
<td>Stata</td>
<td>SPSS</td>
<td>R</td>
</tr>
</tbody>
</table>
RECOMMENDATIONS THAT ROCK

Hutchinson (2015)
RECOMMENDATIONS THAT ROCK
Hutchinson (2015)

Purpose
Support the development of evaluation recommendations.

Intended Use
Writing evaluation reports that include recommendations

Topics Covered
• Criteria for high quality recommendations
• Guidance for organization
• Practical strategies
RECOMMENDATIONS THAT ROCK

Hutchinson (2015)

Criteria for quality recommendations
Practical strategies for writing
Organizational tips
## Quick Reference Guides

### Periodic Table of Evaluation

<table>
<thead>
<tr>
<th>Paradigms</th>
<th>Onological and epistemological stances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>Evaluation outcomes</td>
</tr>
<tr>
<td>Objectives</td>
<td>Overarching evaluation questions</td>
</tr>
<tr>
<td>Criteria</td>
<td>Logics to claim causality</td>
</tr>
<tr>
<td>Approaches</td>
<td>Tools to collect data</td>
</tr>
<tr>
<td>Designs</td>
<td>Tools to convince</td>
</tr>
<tr>
<td>Methods</td>
<td>Tools to implement</td>
</tr>
<tr>
<td>Framework</td>
<td>Tools inherent to evaluation</td>
</tr>
</tbody>
</table>

### Evaluation Design

- **Paradigms**: Realist, Pragmatist, Constructivist
- **Purpose**: Policy, Program, Practice
- **Objectives**: Outcome, Impact, Process
- **Criteria**: Rigor, Reliability, Validity
- **Approaches**: Quantitative, Qualitative
- **Designs**: Ex Post Facto, Ex Ante
- **Methods**: Surveys, Interviews, Focus Groups
- **Framework**: Iterative, Cycles

---

SanzVacca.com (@VisualBrains)
40 Purposeful Sampling Strategies

By Minh Q. Tran
Interactive Chart Chooser

Wondering which type of graph is the best fit for your data? Explore our interactive chart chooser using the filters. For example, if you click on 3+ Points in Time, you'll see familiar faces like line graphs and meet new friends like multimedia timelines.

- All
- 1 Point in Time
- 2 Points in Time
- 3+ Points in Time
- Comparisons
- Correlation
- Distribution
- Exploratory
- Part to Whole
- Progress Towards Goals
- Relationships

Chart Types:
- Fan Chart
- Estimations
- Target Lines on Line Charts
WHERE TO FIND MORE?

BetterEvaluation
Racial Equity Tools
USAID Performance Monitoring and Evaluation TIPS
Community Tool Box
Evaluation Checklist Project
EvaluATE
QUESTIONS?
THANK YOU

Kelly Robertson &
Lyssa Wilson Becho