


**Using the Program
Evaluation
Standards in
Metaevaluation:
Potential &
Pitfalls**


Lori A. Wingate

Annual Conference of the
American Evaluation Association

November 11, 2010 | San Antonio



**Can You
Judge an
Evaluation
by its
Report?**




Overview


1. Context
2. Study design
3. Results
4. Implications: Pitfalls & Potential



Context



Context




NSF's Advanced Technological Education program:


“With an emphasis on two-year colleges, the ATE program focuses on the education of technicians for the high-technology fields that drive our nation's economy.”

www.nsf.gov/ate

5




Context



EvaluATE's Mission:

- strengthen the ATE program's evaluation knowledge base
- expand the use of exemplary evaluation practices
- support the continuous improvement of technician education throughout the nation.

6



Context

Purposes of the Study:

1. Needs Assessment
 - Identify the strengths and weaknesses of ATE evaluations
 - 1 of 3 metaevaluation studies
2. Research
 - Investigate interrater agreement in metaevaluation

7




Study Design

The Typical Metaevaluation

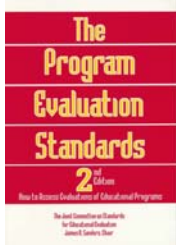
**Number of
Metaevaluators:**

1




Criteria:

Program Evaluation
Standards



Method:

Document Review



Based on my review of 54 published metaevaluations

9

Raters & Reports

Students



Practitioners



Scholars






























Eval.
Project 1

Eval.
Project 2

Eval.
Project 3

Eval.
Project 4

Eval.
Project 5

Eval.
Project 6

Eval.
Project 7

Eval.
Project 8

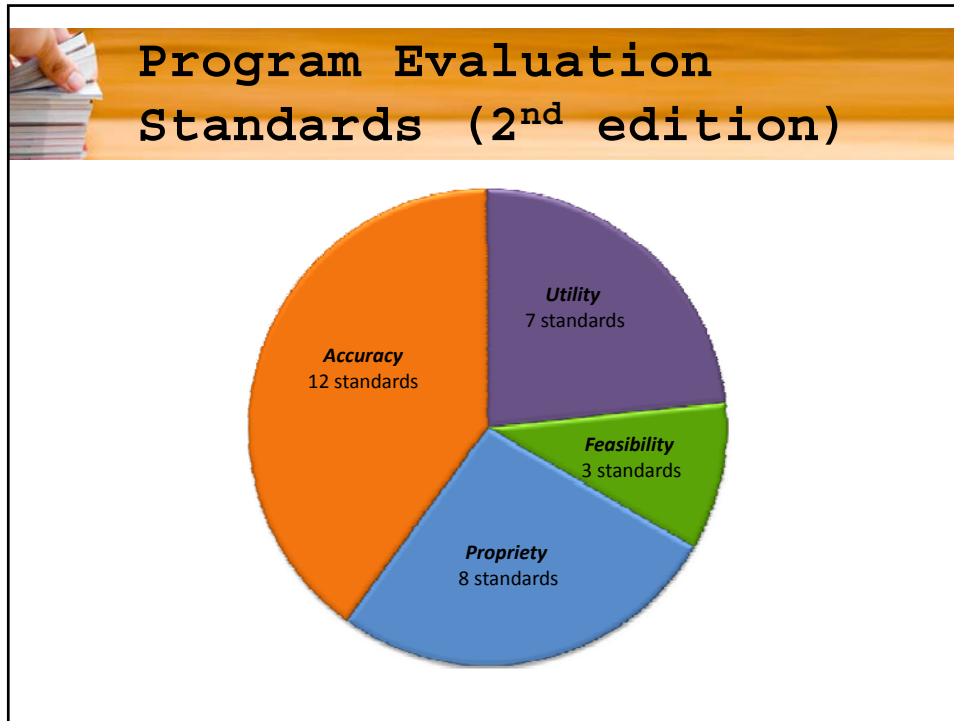
Eval.
Project 9

Eval.
Project 10

10

lori.wingate@wmich.edu

5



Metaevaluation Instrument

Standard	1 Not addressed	2 Minimally addressed	3 Partially addressed	4 Mostly addressed	5 Fully addressed	Insufficient Information	Not Applicable
F1 Practical Procedures	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F2 Political Viability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F3 Cost Effectiveness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12

Example

Standard A10 Justified Conclusions
The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.

Students Practitioners Scholars

x 10 evaluations x 29 standards

13


Measures

Intraclass correlation coefficient (ICC)
Proportion of variance in observed scores that is due to differences in true scores

0 .70 1

No reliability Acceptable reliability Perfect reliability


14




Measures

Percent agreement

Exact agreement



Adjacent agreement



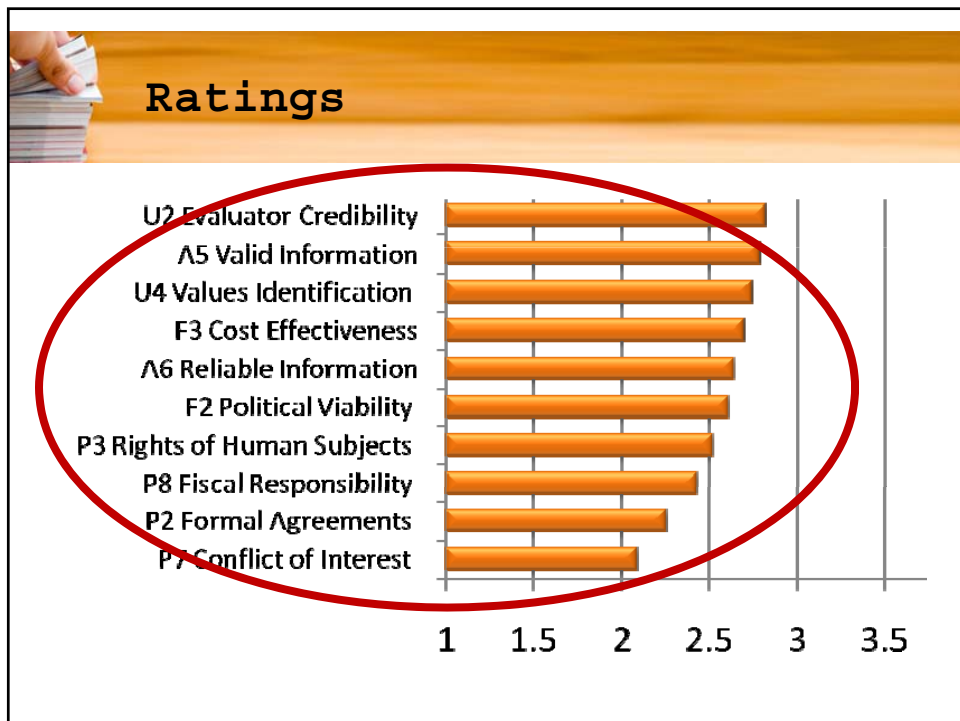
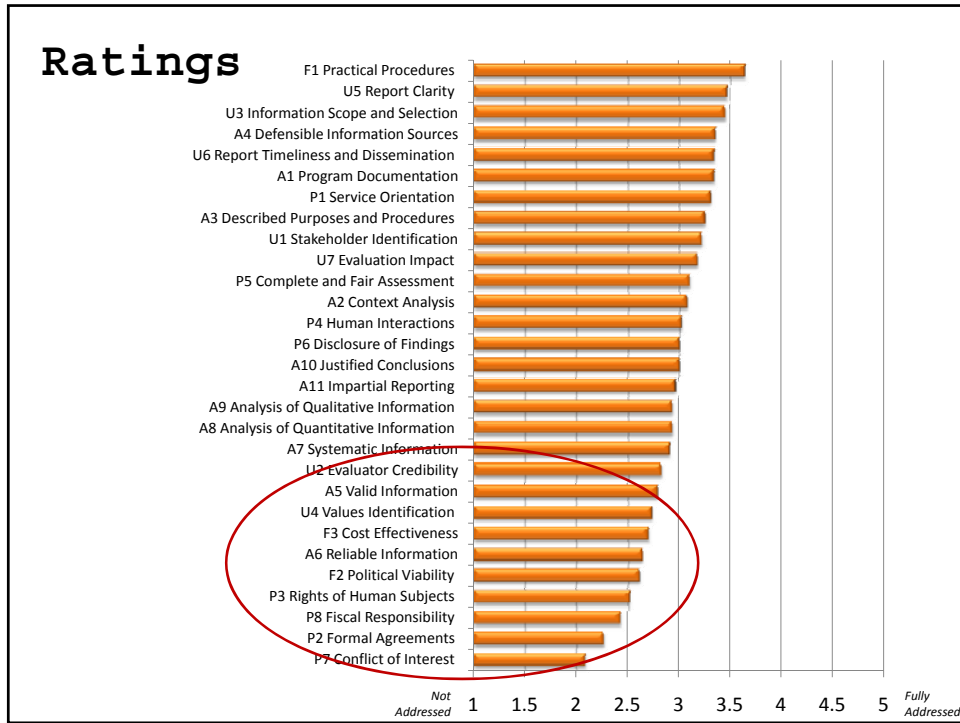
15

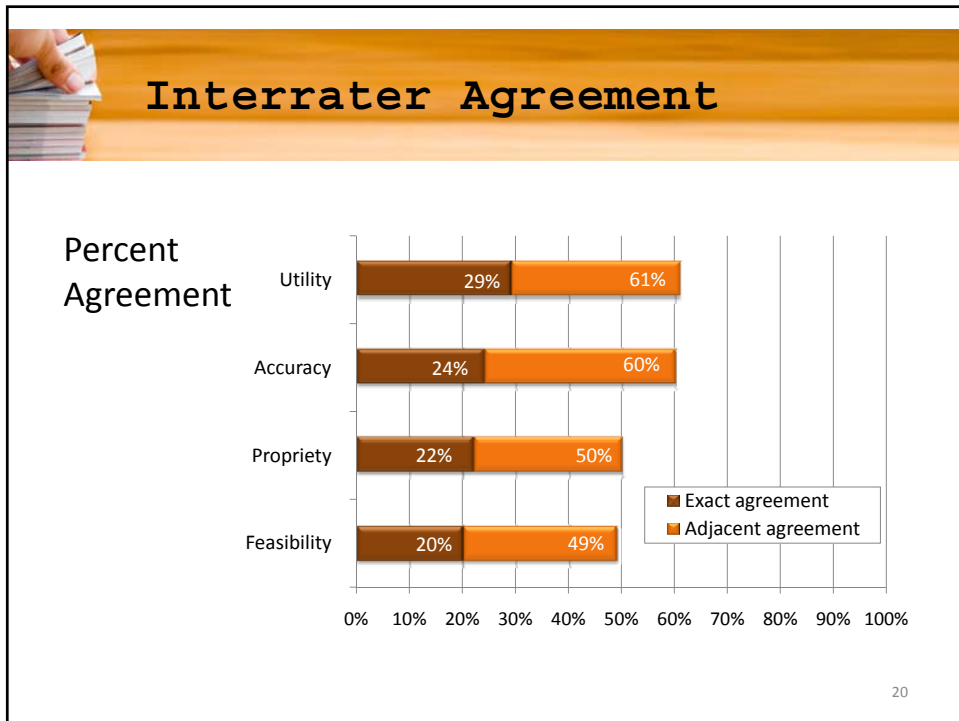
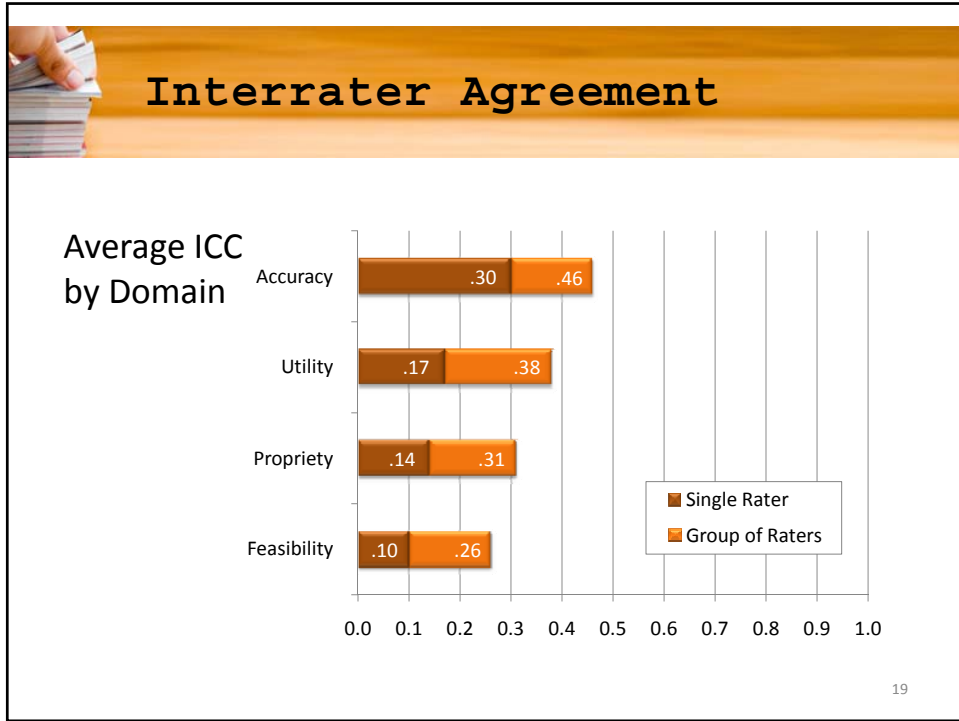
The slide illustrates two types of percent agreement. Under 'Exact agreement', two black stick figures are shown, each holding a white sign with a red border and the number '4'. Under 'Adjacent agreement', two black stick figures are shown; the one on the left holds a sign with the number '4', and the one on the right holds a sign with the number '3'. The background is a light orange gradient.

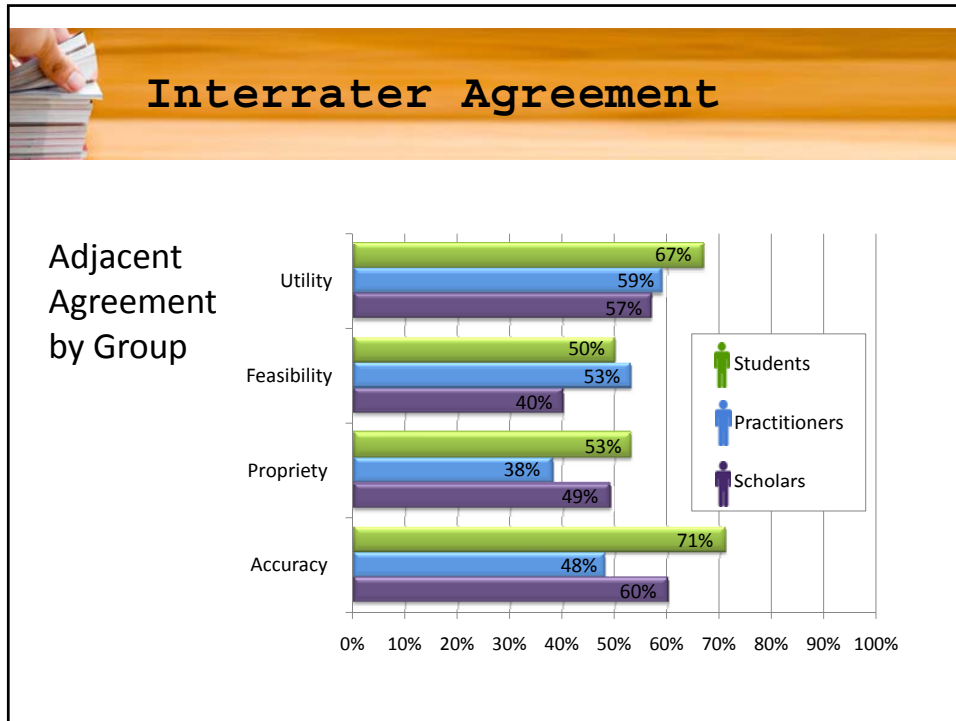


Results

The slide features a close-up photograph of a hand holding a thick stack of papers, with the pages slightly fanned out. The background is a blurred light orange color. The word 'Results' is printed in a black, typewriter-style font on the right side of the image.





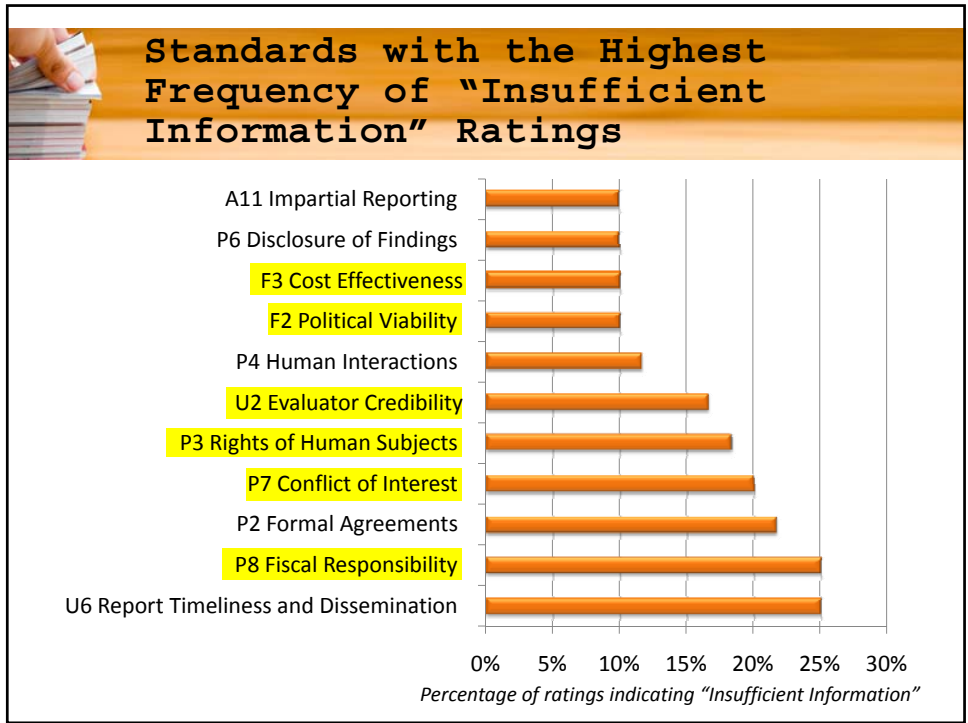



Metaevaluation Instrument

Standard	1 Not addressed	2 Minimally addressed	3 Partially addressed	4 Mostly addressed	5 Fully addressed	Insufficient information	Not Applicable
F1 Practical Procedures	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F2 Political Viability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F3 Cost Effectiveness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

“It is impossible to render a reasonably well-informed judgment based on available information.”

22






Pitfalls


When the Standards are applied to reports:

- Poor interrater agreement
- Uncertainty about how adherence to the Standards should manifest
- Ambiguity of the rating term “Addressed”
- High degree of inference required



Potential


- The Program Evaluation Standards are widely known and accepted
- Metaevaluation *can* serve research and quality assurance purposes



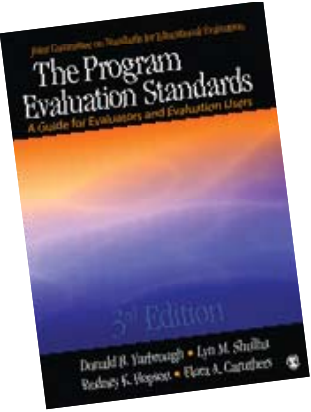
Potential

Recommendations:

- Use the Standards as a springboard to developing specific metaevaluation tools and criteria
- Tailor criteria to context
- Establish reliability
- Avoid single-rater designs




Potential



Evaluator Accountability Standards (3)

- **E1 Evaluation Documentation**
Evaluations should fully document their negotiated purposes and implemented designs, procedures, data, and outcomes.


www.jcsee.org



Thank You!

Marvin Alkin	Tala Davidson	Mhora Newsom-Stewart
Mohammed Alyami	Jonathan Engelman	Lindsay Noakes
Ezechukwu Awgu	Jan Fields	Michael Patton
Lee Balcom	Jody Fitzpatrick	Jon Pedersen
Stacy Berkshire	Amy Germuth	Peter Saflund
Valerie Caracelli	Neal Grandgenett	James Sanders
Christina Christie	Carl Hanssen	Jean Sando
Jan Corbin	John Hattie	Fayez Shafloot
Elaine Craft	Rodney Hopson	Sanjeev Sridharan
Lois-ellin Datta	Robbie McCarty	Carl Westine

Drs. Chris Coryn, Leslie Cooksy, & Arlen Gullickson



This material is based upon work supported by the National Science Foundation under Grant No. 0802245. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

29