



Cornell University  
Cornell Office for Research on Evaluation



MONTCLAIR STATE  
UNIVERSITY

Developmental Systems Science  
and Evaluation Research Lab

# The Development and Validation of Rubrics for Measuring Evaluation Plan, Logic Model, and Pathway Model Quality

Jennifer Brown Urban<sup>1</sup>, Marissa  
Burgermaster<sup>1</sup>, Thomas Archibald<sup>2</sup>, Monica  
Hargraves<sup>2</sup>, Jane Buckley<sup>2</sup>, Claire Hebbard<sup>2</sup>,  
Stephanie Phelan<sup>1</sup>, William Trochim<sup>2</sup>

<sup>1</sup>Montclair State University, Developmental Systems Science &  
Evaluation Research Lab

<sup>2</sup>Cornell University, Cornell Office for Research on Evaluation





# Rationale

- Encouraging high-quality evaluation work is a challenge in the field, especially in small organizations
  - Lack evaluation capacity, expertise
  - Competing demands for time and resources
- Need for concrete measures to evaluate evaluation plans and logic models for programs across various settings





# Development of Rubrics to Measure Quality

- *Systems Evaluation Protocol (SEP)*
  - A systems-based approach, which is designed to build the internal evaluation capacity, including evaluative thinking, of program staff
  - Created to be generally applicable across contexts
- Rubrics developed as measures of quality in the testing of Systems Evaluation Protocol
  - Feedback Tool
  - Quantitative Assessment





# Rubrics for Systems Evaluation and Traditional Evaluation

- What value is added by using a systems evaluation approach?
  - Goal is not to create qualitatively distinct outputs
  - Result will be higher quality models, plans, and evaluations if using systems evaluation
    - but no empirical evidence yet





# Evolution of the Rubric Measures

- Initially a feedback tool for plan & model revisions before evaluation implementation
- Revisions in language and scale made to reflect change in rubrics' purpose
- Aims
  1. Establish a reliable, valid measure applicable across evaluation contexts
  2. Have one measure that can be used for both feedback and assessment





# Evolution of the Rubric Measures

Feedback Rubric	Assessment Rubric
Items scored out of total possible points	Items scored on a Likert-type scale
Supportive wording	Neutral wording
Format spotlights comments from reviewer	Format promotes consistency in scoring





# From Feedback to Assessment

Category	Max. Score	Current Score	Comments
Evaluation Questions <ul style="list-style-type: none"> <li>Well crafted and measurable</li> <li>Feasibility</li> <li>Clearly based on activities and outcomes outlined in the Logic Model</li> </ul>	20	<div style="width: 20%; background-color: gray;"></div>	<div style="width: 10%; background-color: gray;"></div>
Sample <ul style="list-style-type: none"> <li>Specific data sources described</li> <li>Includes number of participants</li> <li>How representative is sample of program participants?</li> <li>Is sample large enough to reach conclusion?</li> </ul>	5	<div style="width: 20%; background-color: gray;"></div>	<div style="width: 10%; background-color: gray;"></div>
Measures			<div style="width: 10%; background-color: gray;"></div>



	Unacceptable	Minimally Acceptable	Adequate	Good	Excellent
Category/Criteria	0	1	2	3	4
<b>Program Mission or Purpose Statement</b>					
1. Communication of goals (i.e., statement conveys the major goals of the program)					
2. Specificity to the program being evaluated (i.e., statement is about the program and <i>not</i> just the larger organization)					
<b>Program Description</b>					
3. Description of program implementation (e.g., includes information about target audience, program scale, activities, etc.)					
4. Description of program context (e.g., includes information about the					



# Evaluation Plan Rubric Structure

Sub-sections	Number of Items
Program Mission or Purpose Statement	2
Program Description	4
Evaluation Purpose	4
Evaluation Questions	4
Sampling	6
Measurement	7
Design	4
Data Collection and Management	3
Data Analysis	2
Evaluation Reporting and Utilization	4
Evaluation Timeline	3
Overall	4







# Logic and Pathway Model Rubric Structure

Logic Model Sub-Sections	Number of Items	Pathway Model Sub-Sections	Number of Items
Inputs	2	Items	1
Activities	3	Connections & Pathways	7
Outputs	3	Overall	2
Outcomes	4		
Assumptions	2		
Context	2		
Overall	2		





# Inter-Rater Reliability

- Randomly selected 15 programs (out of a possible 28) to assess inter-rater reliability
- Randomly paired 6 raters
- Randomly assigned rater pairs to 15 programs





# Inter-Rater Reliability: Evaluation Plan Rubric

Program	Intra-Class Correlation	Program	Intra-Class Correlation
001	.889***	009	.566*
002	.856***	010	.735**
003	.706**	011	.557*
004	.847***	012	.865***
005	.557*	013	.822***
006	.946***	014	.645**
007	.936***	015	.608**
008	.870***	<b>Average</b>	<b>.760</b>

\*\*\*p < .001, \*\*p < .01, \*p < .05, +p < .10





# Inter-Rater Reliability: Logic Model Rubric

Program	Intra-Class Correlation	Program	Intra-Class Correlation
001	.958***	009	.748*
002	.966***	010	.891***
003	.812**	011	.947***
004	.749*	012	1.00
005	.605 <sup>+</sup>	013	.168
006	.815**	014	.510 <sup>+</sup>
007	.631*	015	.620*
008	.727*	<b>Average</b>	<b>.743</b>

\*\*\*p < .001, \*\*p < .01, \*p < .05, +p < .10





# Inter-Reliability: Pathway Model Rubric

Program	Intra-Class Correlation	Program	Intra-Class Correlation
001	.948*	009	.974**
002	.995**	010	.444
003	.959*	011	.995**
004	.940*	012	.999***
005	.924*	013	.962*
006	.992**	014	.948*
007	.988**	015	.993**
008	.970*	<b>Average</b>	<b>.935</b>

\*\*\*p < .001, \*\*p < .01, \*p < .05, +p < .10





# Inter-Rater Reliability Summary

Rubric Type	Average Intra-Class Correlation
Evaluation Plan	.760
Logic Model	.743
Pathway Model	.935





# Internal Consistency: Evaluation Plan Rubric

Sub-Section	Cronbach's Alpha	Average Corrected Inter-Item Correlation	Average Item-Total Correlation
Program or Mission Statement	.945	.897	.897
Program Description	.809	.537	.644
Evaluation Purpose	.737	.393	.527
Evaluation Questions	.745	.424	.543
Sampling	.844	.483	.631
Measurement	.759	.337	.501
Design	.771	.493	.597
Data Collection & Management	.886	.725	.782
Data Analysis	.762	.643	.643
Evaluation Reporting & Utilization	.780	.521	.616
Evaluation Timeline	.809	.607	.672
Overall	.865	.617	.717



# Internal Consistency: Logic Model Rubric

Sub-Section	Cronbach's Alpha	Average Corrected Inter-Item Correlation	Average Item-Total Correlation
Inputs	.634	.491	.491
Activities	.600	.351	.437
Outputs	.749	.494	.584
Outcomes	.710	.358	.510
Assumptions	.893	.844	.844
Context	.733	.617	.617
Overall	.870	.792	.792
All Items	.913	.382	.592







# Internal Consistency: Pathway Model Rubric

Sub-Section	Cronbach's Alpha	Average Corrected Inter-Item Correlation	Average Item-Total Correlation
Items	n/a	n/a	n/a
Connections & Pathways	.823	.391	.566
Overall	.601	.440	.440
All Items	.829	.322	.520





## Next Steps

- Factor Analysis
- Assess quality of plans and models for programs in remaining cohorts
- External raters





# Applications & Potential Uses

- Compare quality to:
  - program/organizational capacity
  - attitudes toward evaluation
  - use of cyber-infrastructure
- Compare quality between programs that use systems evaluation vs. traditional evaluation approaches
- Decision making
  - Funders can measure and compare quality of submitted applications





Cornell University  
Cornell Office for Research on Evaluation



MONTCLAIR STATE  
UNIVERSITY

Developmental Systems Science  
and Evaluation Research Lab

# Contact Information

Jennifer Brown Urban

Developmental Systems Science & Evaluation Research Lab

Montclair State University

[urbanj@mail.montclair.edu](mailto:urbanj@mail.montclair.edu)

Claire Hebbard

Cornell Office for Research on Evaluation

Cornell University

[cer17@cornell.edu](mailto:cer17@cornell.edu)





Cornell University  
Cornell Office for Research on Evaluation



MONTCLAIR STATE  
UNIVERSITY

Developmental Systems Science  
and Evaluation Research Lab

This material is based upon work supported by the  
National Science Foundation under Grant No. 0814364.

Any opinions, findings, and conclusions or  
recommendations expressed in this material are those of  
the author(s) and do not necessarily reflect the views of  
the National Science Foundation





# Pathway Model

