

# A Trial of the Systems Evaluation Protocol (SEP) for Evaluation Planning for STEM Education

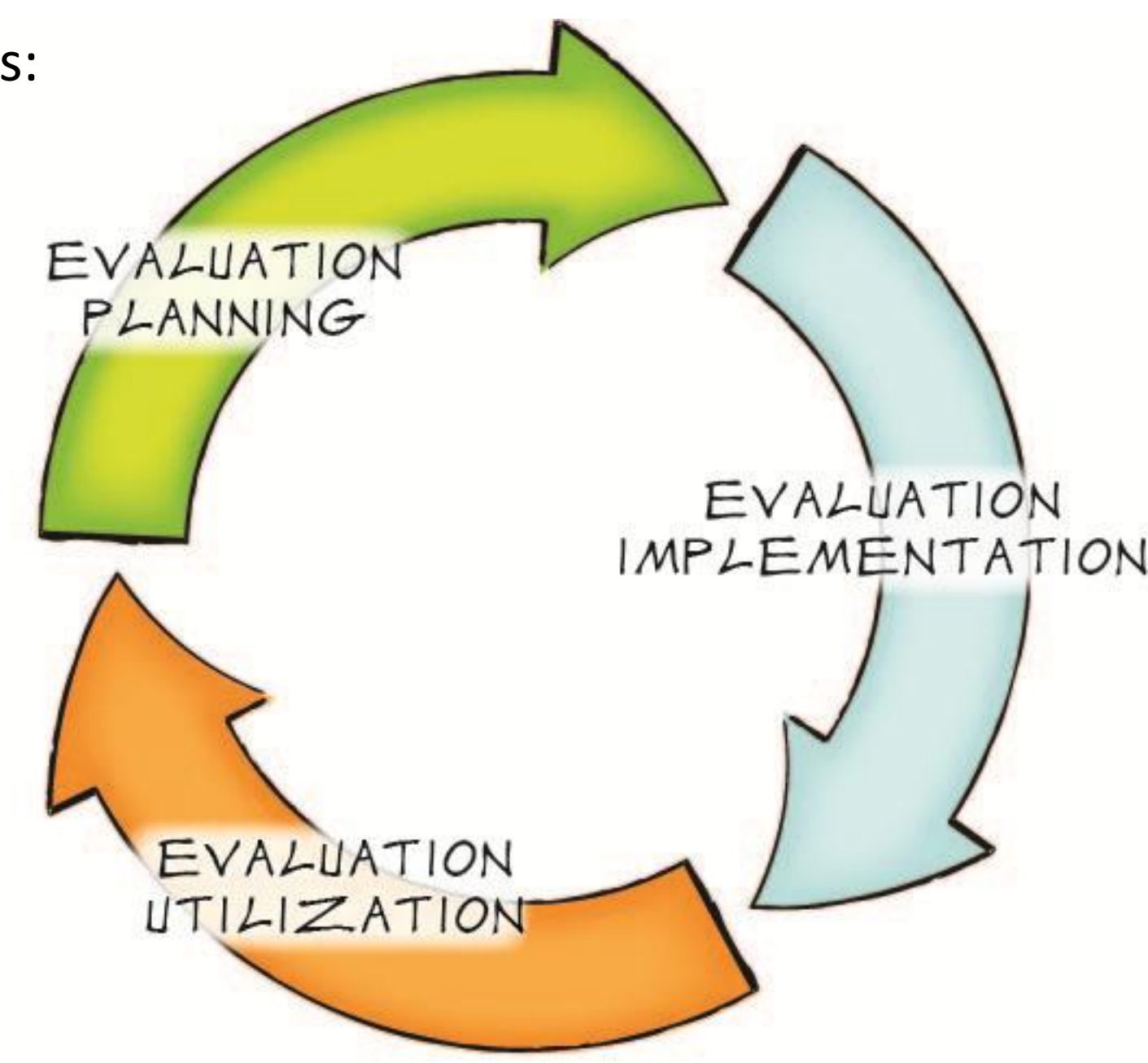
Claire Hebbard\*, Monica Hargraves\*, William Trochim\* (PI), Jennifer Brown Urban (Co-PI)\*\*, Jane Buckley\*, Tom Archibald\*\*\*, Marissa Burgermaster\*\*  
 \*Cornell University, Ithaca, NY; \*\* Montclair State University, Montclair, NJ; \*\*\*Virginia Tech, Blacksburg, VA.

## Introduction

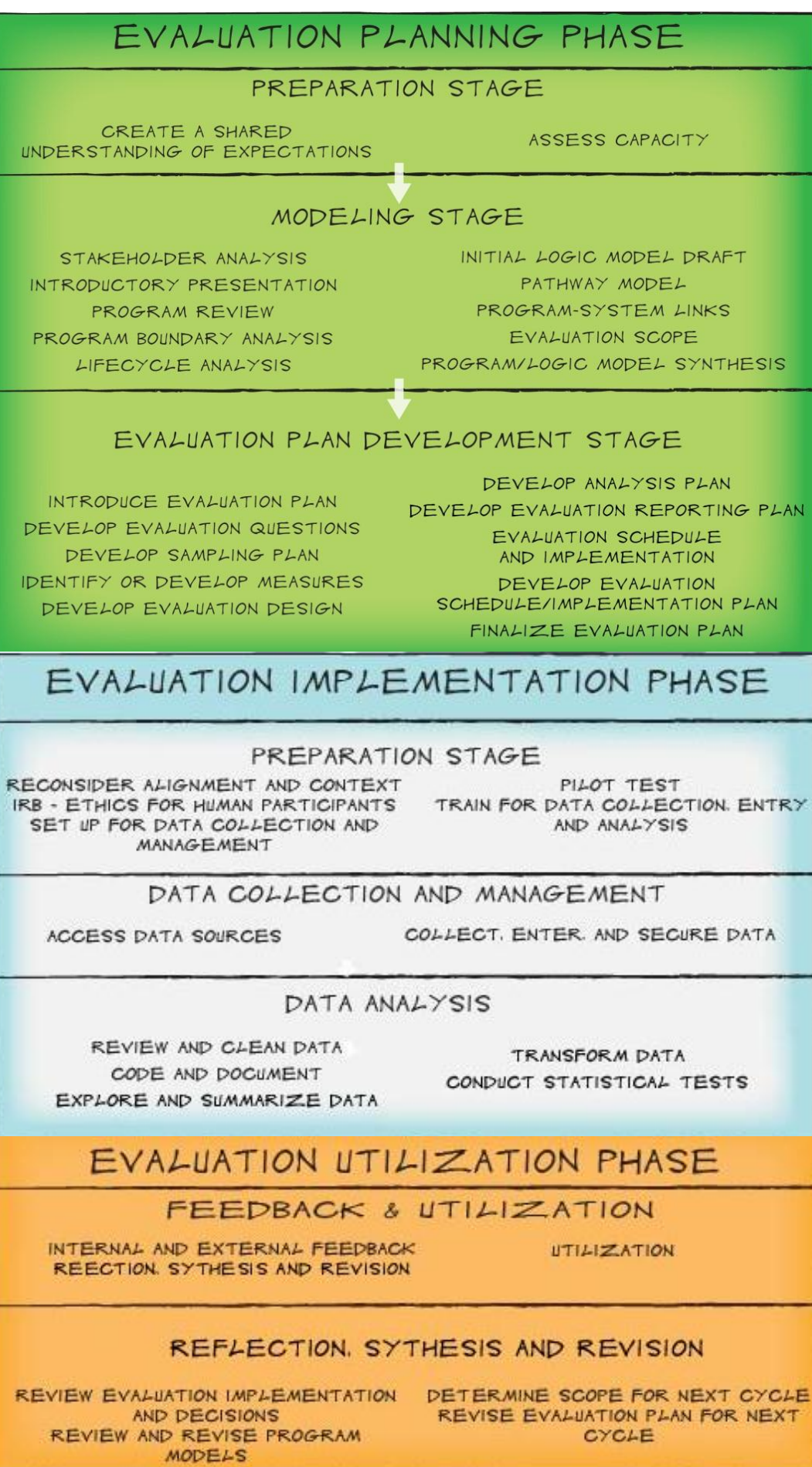
What is the Systems Evaluation Protocol?

An approach to program evaluation that incorporates:

- Program modeling and causal pathways
- Stakeholder mapping
- Local and global contexts
- Links to published research\*\*
- Program evolution\*@
- Practitioner knowledge



## Protocol Steps



The steps of the Protocol\*<sup>1</sup> provide a framework for building a high quality evaluation plan and supporting its implementation and utilization. In practice, and consistent with a systems perspective, the steps do not need to be followed exactly in the order presented at left, but can be conducted in a different order if that suits the needs of the organization and stakeholders involved. Steps should be revisited throughout the planning process.

For purposes of this trial, the Protocol was used as part of an Evaluation Capacity Building effort with Cornell Cooperative Extension (CCE) in New York State, and with Materials Research, Science and Engineering Centers (MRSECs), nationally.

Previous to this study, the Protocol and materials were developed (in working with 46 educational outreach programs) through support from NSF (Grant # 0535492), and Cornell Cooperative Extension. The Protocol and materials were further developed through the current project.

## Methods & Design

Cohort	Programs	Y1	Y2	Y3	Y4
C 2	N=22	Evaluation Planning	Implementation & Utilization		
C 3	N=12		Evaluation Planning	Implementation & Utilization	
C 4	N=11			Evaluation Planning	Implementation & Utilization
C 5	N=17				MySEP Evaluation Planning

Evaluation Planning	Programs		
	Method	Total	Incomplete
Facilitated	45	2	43
Self-directed	17	3	14
<b>Grand Total</b>	<b>62</b>	<b>5</b>	<b>57</b>

Evaluation Implementation	Programs		
	Method	Total	Incomplete
Facilitated	29	6	23

The focus of this project was originally on the evaluation planning phase. Cohorts 2-4 worked with CORE staff to facilitate the work of evaluation planning, and cohort 5 independently accessed the same materials online (non-facilitated). We found that it was also important to provide support for the implementation phase, so we provided additional support. This poster includes a summary of results from their implementation and utilization efforts.

## Key Publications:

\*Trochim, W., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., Johnson, M., and Burgermaster, M. (2012). The Guide to the Systems Evaluation Protocol (V2.2). Ithaca, NY.

\*\*Urban, J. B., & Trochim, W. (2009). The Role of Evaluation in Research-Practice Integration: Working Toward the "Golden Spike". *American Journal of Evaluation*, 30(4), 538-553.

\*@Urban, J. B., Hargraves, M. and Trochim, W. M. (2014). "Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate." *Evaluation and Program Planning*, 45, 127-139.

\*%Urban, J. B., Burgermaster, M., Archibald, T., Byrne, A. (In Press.) Relationships Between Quantitative Measures of Evaluation Plan and Program Model Quality and a Qualitative Measure of Participant Perceptions of an Evaluation Capacity Building Approach. *Journal of Mixed Methods Research*

\*^Buckley, J., Archibald, T., Hargraves, M., Trochim, W., (TBD) Defining and Teaching Evaluative Thinking: Insights from Research on Critical Thinking. (planned for AJE late 2014)

## Evaluation Planning Selected Outcomes & Analysis

Program Models	Average # Activities	Average # Outputs	Average # Short-term Outcomes	Average # Medium-Term Outcomes	Average # Long-Term Outcomes	Average # Pathways
Facilitated	5.7	4.8	7.6	7.3	5.4	51.4
Self-directed	6.9	5.5	5.6	5.2	3.9	24.3
<b>Overall</b>	<b>6.0</b>	<b>5.0</b>	<b>7.2</b>	<b>6.8</b>	<b>5.1</b>	<b>45.0</b>

## Rubrics for Quality of Models and Evaluation Plan\*%

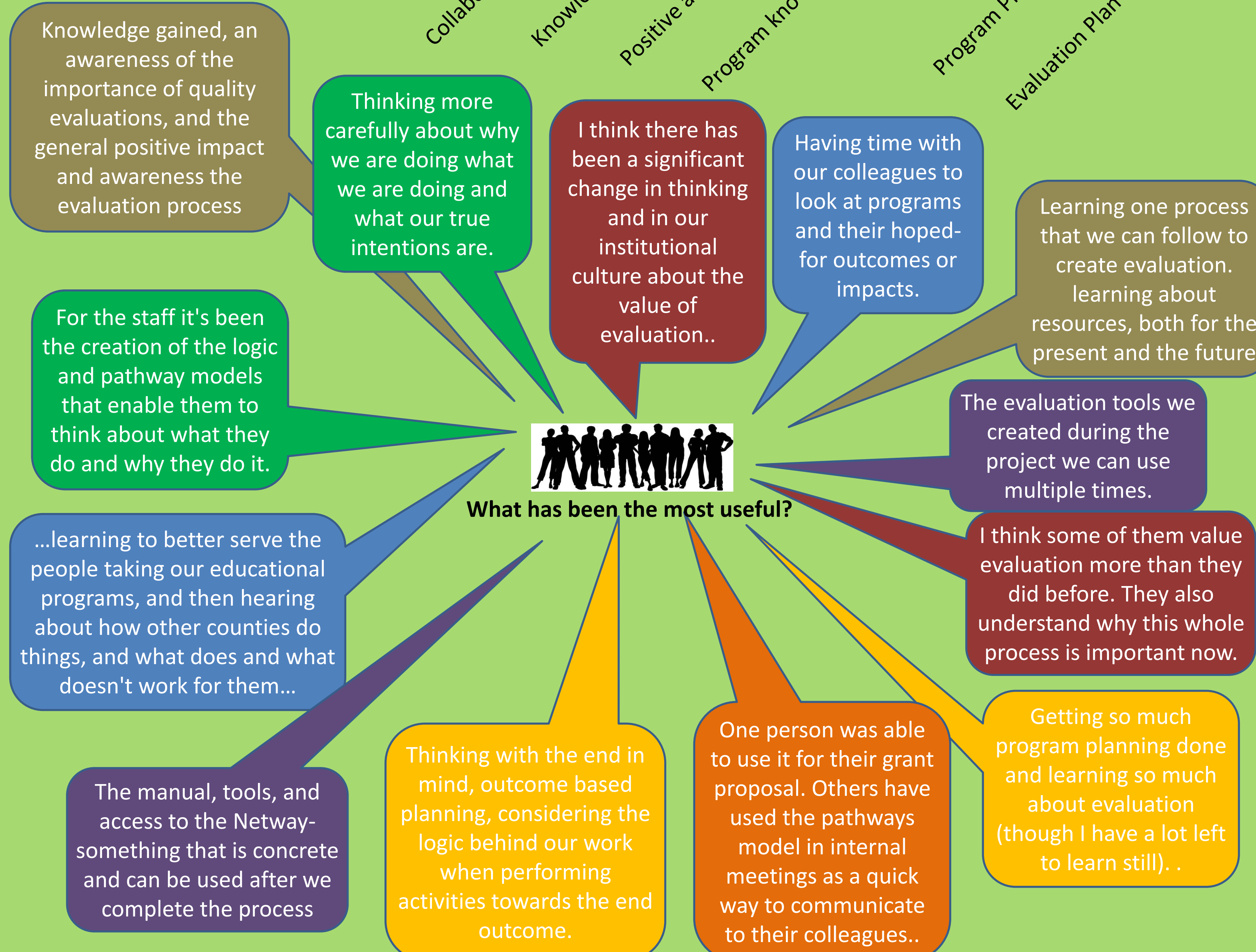
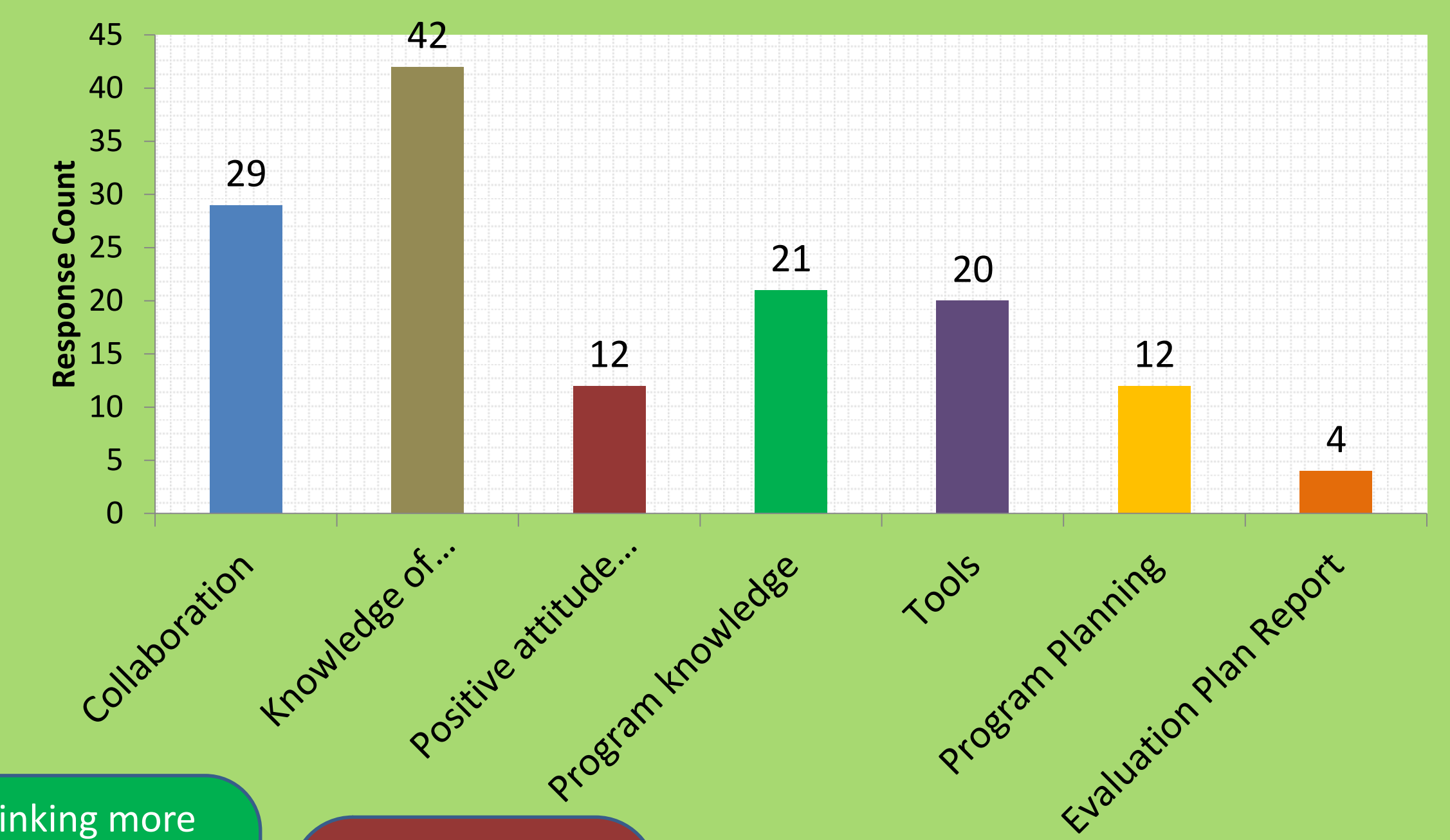
Logic Model	Min	Max	Average (out of 72)	% of Possible
Facilitated	16	71	52.2	73%
2	28	69		54.1
3	47	65		60.0
4	16	71		43.6
Self-directed	20	67	42.8	59%
5	20	67		42.8
<b>Grand Total</b>	<b>16</b>	<b>71</b>	<b>50.0</b>	<b>69%</b>

Pathway Model	Min	Max	Average (Out of 40)	% of Possible
Facilitated	18	40	30.8	77%
2	20.5	37		29.5
3	18	40		31.3
4	28	40		32.8
Self-directed	0	34	14.8	37%
5	0	34		14.8
<b>Grand Total</b>	<b>0</b>	<b>40</b>	<b>27.0</b>	<b>68%</b>

Eval Plan	Min	Max	Average (Out of 192)	% of Possible
Facilitated	58	170	124.1	65%
2	59.5	158		116.5
3	79	155		129.6
4	58	170		135.0
Self-directed	24	178	88.3	46%
5	24	178		88.3
<b>Grand Total</b>	<b>24</b>	<b>178</b>	<b>115.6</b>	<b>60%</b>

Overall	Min	Max	Average (out of 304)	% of Possible
Facilitated	104	270	207.1	68%
2	127	259.5		200.1
3	144	260		220.9
4	104	270		211.5
Self-directed	53	240	145.9	48%
5	53	240		145.9
<b>Grand Total</b>	<b>53</b>	<b>270</b>	<b>192.7</b>	<b>63%</b>

## Interview Responses: What has been the most useful? n=140



## Conclusions

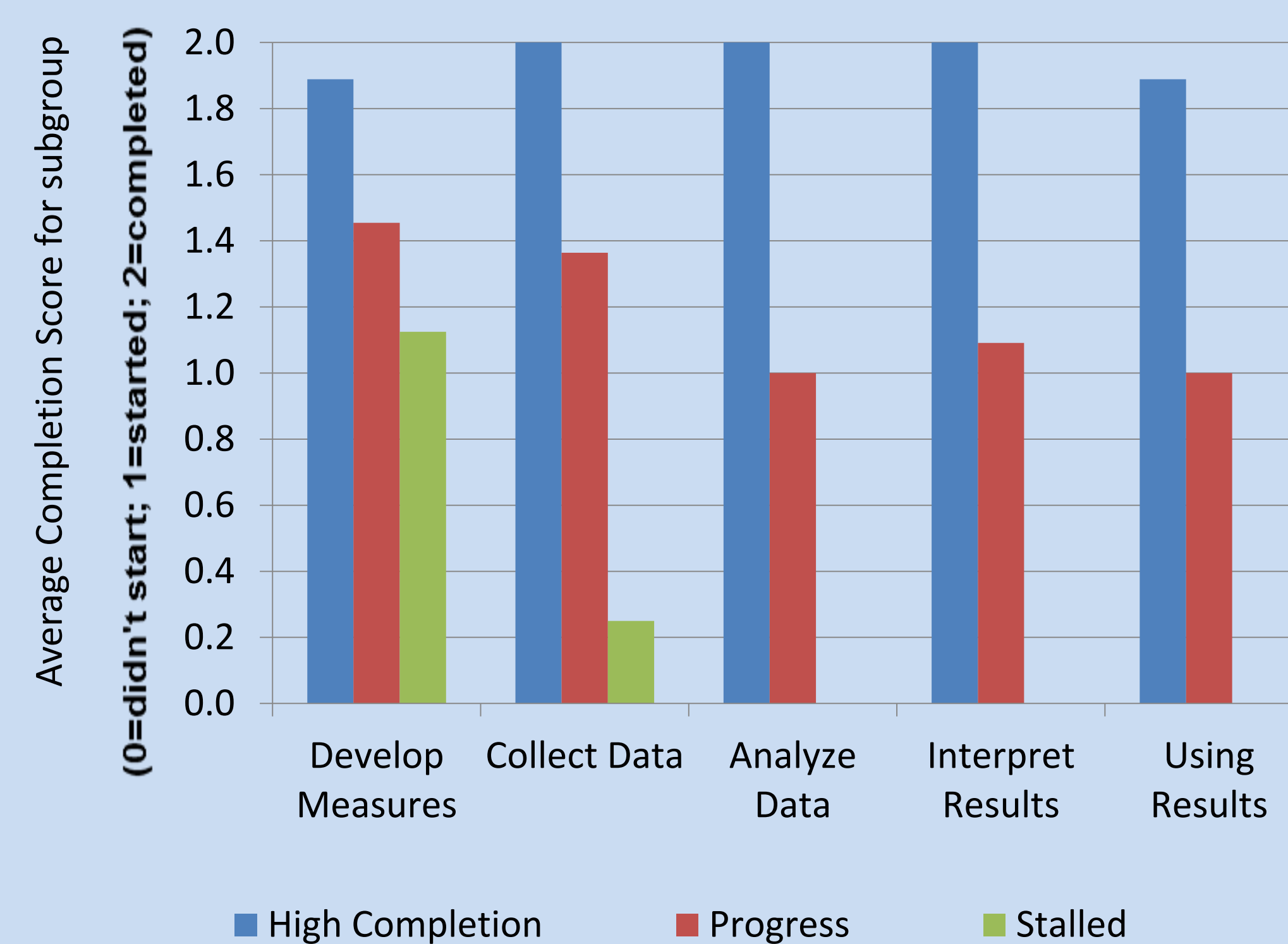
- Active engagement with the process changes the way people think<sup>\*\*</sup> – about their programs, about evaluation, about the value of data
- Doing the program modeling alone yields significant "Aha's!" about programs
- Those who really "get it" strengthen their subsequent program development and evaluation work on many programs beyond their EP program, and share tools and insights with others
- Preliminary analysis of completion of evaluation implementation reflects that everyone faces barriers to conducting evaluation, but the reported magnitude of barriers doesn't seem to be predictive of who will or will not complete evaluation.

## Analysis of the Implementation Phase of Evaluation

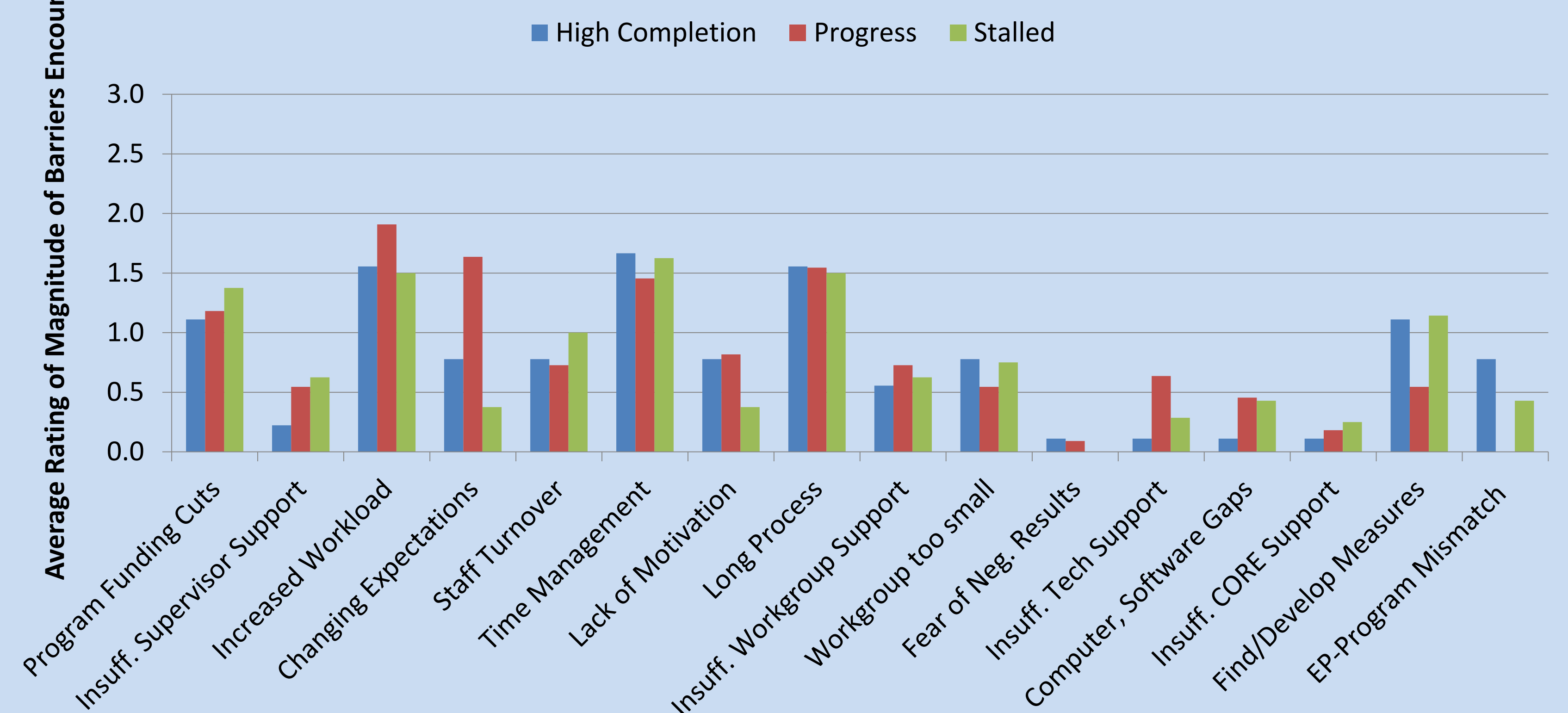
N=28 from CCE, Cohorts 1-5

Subgroups:  
 "High Completion" About one third (n= 9) of the programs fully completed their planned evaluations  
 "Progress" Slightly more than a third (n=11) made substantial progress, and obtained usable results  
 "Stalled" Slightly less than a third (n=8) stalled at some point in the implementation process

### Patterns of Progress for subgroups



## Magnitude of Barriers Encountered (Rating Scale: 0 = Not a Problem, 1 = Small Problem, 2 = Problem, 3 = Big Problem)



Challenges don't explain patterns of progress: Contrary to expectations, in general there was little difference in ratings between the three subgroups. Future research will explore resilience factors.

This material is based upon work supported by the National Science Foundation (Grant No. 0814364) and Cornell Cooperative Extension.

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, Cornell University, or Cornell Cooperative Extension.