Focusing Sessions

A valuable stakeholder engagement and evaluation design tool

Veronica S. Smith
National Science Foundation Research Traineeship Annual Meeting
September 2019
1. Purpose
Donaldson’s theory-driven evaluation science

• Contributes to team learning and program improvement
• Increases stakeholder engagement
• Reduces evaluation anxiety
• Hypothesis driven

Focusing session goals

• Create common understanding
• Make program theory explicit and testable
• Frame evaluation questions and identify key measures
• Provide framework for data collection and analysis
Are you ready?
3. Be gameful!
Theory of Action

**GUIDING PRINCIPLES**
- Systematic Inquiry
- Competence
- Integrity
- Respect
- Common Good

1. **Engage stakeholders**
2. **Map program/desired impact**
3. **Focus measurement + evaluation**
4. **Gather credible, timely + relevant evidence**
5. **Ensure use + share lessons learned**
4. Focus on learning together
Theory of Change

- Informed measures + questions
- Increased communication + trust
- More systematic measurement + evaluation
- Evaluative thinking
- Actionable information/feedback
- Improved decision making
- More effective action
- Organizational learning

Increased Impact

Timeline:
- Focusing Session
- 6-12 Months
- 1-3 Years
- 3+ Years
Community of Learning
Where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole (reality) together.

-Peter Senge
What are the 3 attributes you would like to contribute to our learning community?

Veronica Smith

1. Diversity, equity, inclusion best practices
2. Facilitation skills
3. Data science
How do we want to conduct the evaluation as co-creators of knowledge?

What kind of learning community do we want to create?
Norms

• Experience Discomfort
• Take Risk
• Stay Engaged
• Listen for Understanding
• Speak Your Truth
• Expect/Accept Non-Closure
• No Fixing
5. Focus on change together
Program Logic

ACTION \[\rightarrow\] CHANGE

- Input
- Activity
- Output
- Short-term Outcome
- Intermediate Outcome
- Long-term Outcome

Process evaluation
Impact evaluation
Focus on **CHANGE**

- Increase
- Broaden
- Diversify
- Accelerate
- Decrease
- Mediate
- Compress
- Decelerate
- Deepen
- Moderate
- Temper
- Temper
6. Map together
Need impact map to move beyond waving of hands

YOUR NRT PROGRAM

STEM Diversity, Equity & Inclusion

Trainee Knowledge & Skill

Interdisciplinarity & Convergence

Timeline

5+ years

Anchor Changes
Impact Map

- Desired Change A
- Desired Change B
- Desired Change C
- Desired Change D
- Desired Change E

Timeline:
- 1-2 months
- 2-4 years
- 5+ years

Anchor change leads to Desired Change D, which leads to Desired Change E, which leads to Desired Change A.
Theory of Change

Short-term Outcomes

- Increase number of departments offering transcriptable options
- Increase alignment of the DIRECT program courses with student learning needs
- Increase number of prospective employers in capstone project
- Increase trainee professional development opportunities (e.g., communication)

Intermediate Outcomes

- Increase the number and recognition of data scientists with advanced materials knowledge and skills - Energy Data Scientists

Long-term Outcomes

- Workforce improve employability and options for DIRECT program participants
- Improve visibility and recognition of DIRECT program value by prospective employers
7. Focus on knowledge needs together
What if you could get a sound answer to a question about your program that would help improve outcomes?
What are the measures you need to track progress against goals?
Example questions

• How confident are trainees about their career prospects and competitiveness in the clean energy industry?
• How effective are the NRT courses at preparing students for the capstone project?
• How capable are students at applying and communicating about data science across domains?
“I felt like my opinion as a student and my contributions were really valued.”

“Enjoyed being able to share as a stakeholder.”

“Healthy discussion and reframing of comments.”

“Having input into the program’s focus.”

“The collaborative process allowed everyone to participate.”

“Concrete development of questions and group discussion.”

“It was great to learn the goals and general direction of the organization.”
Data2insight founder Veronica Smith and strategic partner Antonella Guidoccio are presenting at the 2018 European Evaluation Society conference in Thessaloniki, Greece from Oct 1-5, 2018. We provide here information and resources to download from their round table sessions.

**Focusing sessions: a stakeholder engagement and evaluation design method**

When working with diverse communities, allowing time for stakeholder engagement is foundational to knowledge co-creation. A focusing session is a best practice that comes from theory-driven evaluation that contributes to better informed evaluation questions, which leads to more robust evaluation design, and findings that stakeholders value.

**Meaningful M&E systems through evidence-based leadership and evolutionary purpose**

In this age of increasing complexity, rapid change, and demand for greater transparency and accountability from leaders the most successful teams of the future will excel at practical performance measurement and strategic use of evaluation.
Questions?

veronicasmith@data2insight.com
Measurement Invariance Across Time

Cheryl Schwab PhD
Internal Evaluator
Overview of the DS421 Program
Purpose

Test the assumption of measurement invariance when assessing change in trainee ability (latent variable) using the same instrument at several time points.
Assumption

Responses across time points for the same person are conditionally independent given the latent variable (i.e., the correlations between the responses at each time point are smaller than can be attributed to trainee ability).
Item Response Theory

Underlying function called an item characteristic curve which specifies that as the level of a trait increases the probability of a correct response to an item increases.

(see Hambleton, Swaminathan & Rogers, 1991)
Outcome Survey

Rate your ability to do the following in a collaboration (i.e., project, research, work).

I am a(n)...novice, advanced beginner, proficient, expert.

---contribute environmental science disciplinary knowledge.
---conduct collaborative research.
---write for audiences inside primary discipline.
---present ideas to audiences outside primary discipline.
Differential Item Functioning (DIF)

Test if the difficulty of an item is different across groups or across time.
Data

Collected between November 2015 and May 2018
Administrations 4 times across two year program
Cohorts = 3
Departments = 11
Cases n= 78
# Modeling Items

<table>
<thead>
<tr>
<th>MODEL</th>
<th>DEVIANCE</th>
<th>#PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating Scale Model</td>
<td>4314</td>
<td>33</td>
</tr>
<tr>
<td>item + step</td>
<td>4230</td>
<td>90</td>
</tr>
<tr>
<td>Partial Credit Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>item + item*step</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Map of Latent Distribution and Thresholds
Latent Regression

REGRESSION COEFFICIENTS

CONSTANT                      -0.226 ( 0.236)
c2                         0.007 ( 0.383)
c3                         0.414 ( 0.395)

Variance                   2.002
Unconditional Variance    2.123  ~ 3%
Cohort by EAP

EAP (logit)

Cohort

1

2

3
<table>
<thead>
<tr>
<th>MODEL</th>
<th>DEVIANCE</th>
<th>#PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>item + step</td>
<td>4319</td>
<td>35</td>
</tr>
<tr>
<td>item + step + time</td>
<td>4310</td>
<td>38*</td>
</tr>
</tbody>
</table>

* Significant change in deviance 0.05
## Main Effect Time

<table>
<thead>
<tr>
<th>TIME</th>
<th>ESTIMATE</th>
<th>ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.523</td>
<td>0.108</td>
</tr>
<tr>
<td>2</td>
<td>0.225</td>
<td>0.165</td>
</tr>
<tr>
<td>3</td>
<td>-0.483</td>
<td>0.175</td>
</tr>
<tr>
<td>4</td>
<td>-0.265*</td>
<td></td>
</tr>
</tbody>
</table>

*estimate constraint

Chi-square test of parameter equality = 32.77, df = 3, Sig Level = 0.000
Map of Latent Distribution and Parameters
Results

- Distribution of item difficulties shifts as the students move through the program.
Next Steps

• Modeling differences between cases
• Case dependence
• Including more data per case
Faculty and Staff

David Ackerly (Co-PI), Integrative Biology (L&S)
Max Auffhammer (PI and Director), Agricultural and Resource Economics (CNR)
Maggi Kelly (Co-PI), Environmental Science, Policy and Management (CNR)
Philip Stark (Co-PI), Statistics (L&S)
David Anthoff, Energy and Resources Group (ERG)
David Culler, Electrical Engineering and Computer Sciences (CE)
Kristina Hill, Landscape Architecture and Environmental Planning (CED)
Solomon Hsiang, Goldman School of Public Policy (GSPP)
Laurel Larsen, Geography (L&S)
Heather Constable (BiGCB Berkeley Initiative on Global Change Biology & DS421 NSF Research Traineeship: Data Science for the 21st Century & UC Berkeley Field Stations Campus Coordinator)
A Collaborative Approach to Program Evaluation for an NRT

Valerie Decker
Assistant Director, Center for Evaluation and Assessment
University of Iowa
Sustainable Water Development Program

Poster Session 2
Number 211

waterhawks.uiowa.edu
Program Evaluation Overview

• Systematic data collection
• Informing program improvements
• Determining the success of a program
• Driven by the information needs of stakeholders
The Program Evaluation Standards

Dimensions of quality:
• Utility
• Feasibility
• Propriety
• Accuracy
• Evaluation Accountability

Outcome of a Collaborative Approach

Strong foundational plan with opportunities to adapt built in

Necessary Components

Evaluators as collaborators
Budget for evaluation
Questions
Assessing Interdisciplinary Competency in the Disaster Resilience and Risk Management (DRRM) Graduate Program using Concept Maps

Marie C. Paretti, PhD
Education Director, DRRMVT
Professor, Engineering Education
mparetti@vt.edu

Image courtesy of NASA Goddard Space Flight Center

This material is based upon work supported by the National Science Foundation under Grant Number 1735139. 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.
DRRM Program: Purpose

- Promote collaboration through a transdisciplinary educational experience
- Train business analysts, educators, engineers, scientists, and urban planners to better prepare for, respond to, and recover from natural disasters
Concept Maps

- Measures domain understanding
- Captures depth (levels of hierarchy) & interconnections (number of links)
- Supports multiple scoring methods
  - Quantitative
  - Holistic
## Concept Map Scoring: Holistic Criteria

<table>
<thead>
<tr>
<th></th>
<th>Complexity</th>
<th>Density</th>
<th>Organizational Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construct</strong></td>
<td>Depth of demonstrated knowledge</td>
<td>Breadth of demonstrated knowledge</td>
<td>Sophistication of organization</td>
</tr>
<tr>
<td><strong>Measurement</strong></td>
<td>Number of hierarchies included in map</td>
<td>Number of concepts included in map</td>
<td>Organizational strategy</td>
</tr>
</tbody>
</table>

Holistic Scoring: DRRM Scholars* Only

Organizational Strategy

Topical
Structured Topical
Functional Topical
Functional

Density: Number and range of topics and subtopics
Complexity: Connections across topics
Organizational Strategy: Topical, Structured Topical, Functional Topical, Functional

*Scholars are those students participating in the full doctoral program
Strengths of Concept Maps as Assessment Tool in this Context

1. Provides a representation of student learning
2. Supports analysis of depth of thinking and interactions among concepts
Student G

Pre-Map

- Mitigation, Assessments, Plans, Outreach, Analysis
- Relocation, New Policies, Focused Efforts on Mitigation, Laws, Recovery
- FEMA, Red Cross, City Government, NGO
- Impact of Disaster
- Organizations

Post-Map

- Resilience and Risk Management
- Stakeholders
- Researchers
- Models
- Decision Support Systems

Scalability: Return to a state of being that is better than pre-disaster conditions.
Challenges of Concept Maps as Assessment Tool in this Context

- DRRM concepts span, bridge, and intersect multiple disciplinary boundaries.
  1. Interdisciplinarity creates a high level of complexity.
  2. 2D Concept map may be insufficient to map knowledge.
  3. Quantitative scoring complicated by dimensional limits, scope, and “between” spaces.
Suggestions for using Concept Maps as Assessment Tool in Highly Interdisciplinary Contexts

1. Written explanations can complement visual representations and expand our ability to assess learning.
Suggestions for using Concept Maps as Assessment Tool in Highly Interdisciplinary Contexts

1. Written explanations can complement visual representations and expand our ability to assess learning.

2. Global (holistic) scoring approaches may more readily capture nuances in the concept maps than quantitative (# of nodes, # of connections, # of levels) methods.
Questions?

Jessica Deters
✉️ jdeters@vt.edu
🐦 @jess_deets

Dr. Marie C. Paretti
✉️ mparetti@vt.edu
🐦 @mparetti2

DRRM VT
https://www.drrm.fralin.vt.edu

@DRRM_VT
Using Formative Evaluation Practices to Promote Continuous Improvement in two NRT Projects

Gili Marbach-Ad
College of Computer, Mathematical, and Natural Sciences, University of Maryland
Mission:
To enhance graduate students’ following skills:
- Interdisciplinary Research
- Communication
- Teamwork/Leadership
- Mentorship

- Cohort 1: Spr2017 - Spr2019 (N=11)
- Cohort 2: Fall2017 - Fall2019 (N=12)
- Cohort 3: Fall2018 - Fall2020 (N=15)
- Cohort 4: Fall2019 - Fall2021 (N=9)
- Cohort 5: Fall2020 - Fall2022

Poster #212 Session 2
Mission:

A program that trains students to effectively work and communicate across disciplines to address food/energy/water nexus issues.

Cohort 1: Spring 2019

Poster #110 Session 1
COMBINE Change Model

Mission
To enhance graduate students in the following skills:
- Interdisciplinary Research
- Communication
- Teamwork/Leadership
- Mentorship

Rationale: Traditional Ph.D. programs fail to develop interdisciplinary knowledge, skills, and research approaches that are necessary to current workplaces.
Reasons: Programs often include insufficient mentoring, too much emphasis or siloed research, failure to acknowledge useful methodologies from outside academic fields.
Recommendations: Develop a model for interdisciplinary graduate research programs that emphasize the above desired skills.

Strategies
- Build in targeted education and training into traditional Ph.D. programs, including:
  - Discipline-bridging coursework
  - Advanced interdisciplinary coursework such as a team-based module course and an intensive data project course
  - Weekly NRT seminars
  - Peer to Peer tutorial week
  - Internships with NRT partners
  - Mentoring
  - Annual Career Development Workshop
  - Annual NRT Symposium
  - Outreach via UMD channels like CompSci Connect and GEMS
  - Undergraduate research mentoring
- Work with advisors and PIs
- Include COMBINE faculty member from outside of core discipline on thesis committee

Target Groups
Ph.D. students in Mathematical/Physical, Computer Engineering/Computational, and Life Sciences at the University of Maryland.

Overall Outputs:
- 35 Scholars with stipend (more without stipend)
- Increased reported confidence in targeted skills
- Increased number of publications and presentations of interdisciplinary research
- Increased advisor awareness and appreciation for interdisciplinary research
- Increased number of internships
- Develop a model for other future interdisciplinary programs

First Year Outcomes:
- Develop program components

First Year Reflections
- Assessing progress of course components and their potential strengths and weaknesses (Feedback from students and instructors)
- Organizing future schedule of COMBINE program and identify chances for further investigation
- Adapting tools to better reflect experience in COMBINE program
Framework for Assessment: Measuring Five Levels of Program Evaluation

- Participation
- Satisfaction
- Learning
- Application
- Impact


Mixed Model Research Design

COMBINE: Cohort 1 (Jan 2017 - Jan 2019)

Data Collection:
- Observations
- Artifacts
- Surveys
- Interviews

Cohort 1 (N=13-1=12)
Mixed Model Research Design

COMBINE: Cohort 2 (May 2017- May 2019)

Surveys

- Pre-Survey
- Post-Survey

Data Practicum

Research & Progress Seminar (n=12), Network Analysis (n=21)

Lit. Survey Seminar & Data Practicum

Network Analysis & Research & Progress Seminar

Lit. Survey Seminar

Interviews

- Int. with instructors (n=2)
- Int. with students (n=4)
- Focus group with students (~15)
- Follow-up interview with students (n=3)
- Interview with Advisors (n=3)
- Int. with students (n=4)
- Follow-up interview with students (n=4)

Cohort I

Cohort II
Mixed Model Research Design

COMBINE: Cohort 3 (May 2018- May 2020)

Data Practicum & Research & Progress Seminar
Symposium

May/Aug 2017
Jan 2017

Network Analysis & Research & Progress Seminar

Pre-Survey
Post-Survey

May/Aug 2018
Dec/Jan 2018

Interview with instructors (n=2)
Interview with students (n=4)

Focus group with students (~15)

Follow-up interview with students (n=3)

Interview with Advisors (n=3)

Int. with students (n=4)

Research & Progress Seminar

May/Aug 2019
Dec/Jan 2019

Follow-up interview with students (n=4)

Lit. Survey Seminar & Data Practicum

May/Aug 2020
Dec/Jan 2020

Post-Program Survey (+control group)

Post-Program Survey

Follow-up interview with students (n=4)

Interview with students (n=4)

Interview with instructors (n=2)

Int. with students (n=4)

Follow-up interview with students (n=4)

Focus group with students (~15)

Interview with Advisors (n=3)

Int. with students (n=4)

Follow-up interview with students (n=4)

Interview with instructors (n=2)

Int. with students (n=4)

Follow-up interview with students (n=4)

Int. with instructor

Int. with instructor

Mixed Model Research Design
Mixed Model Research Design

UMD Global STEWARDS
Feedback from First Cohort

For the COMBINE components that you participated in, check those that helped you with each of the following objectives (check all that apply)

<table>
<thead>
<tr>
<th>COMBINE Components</th>
<th>Interdisciplinary Research</th>
<th>Communication</th>
<th>Career Preparation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learn about important research results in network science</td>
<td>Learn about the theory and methods of network analysis</td>
<td>Develop an appreciation and understanding of the questions and methods of other fields</td>
<td>Develop methods of network analysis to biological data</td>
</tr>
<tr>
<td>Learn about the theory and methods of network analysis</td>
<td>Develop an appreciation and understanding of the questions and methods of other fields</td>
<td>Develop methods of network analysis to biological data</td>
<td>Develop skills for oral presentations</td>
</tr>
<tr>
<td>Develop an appreciation and understanding of the questions and methods of other fields</td>
<td>Develop methods of network analysis to biological data</td>
<td>Develop skills for oral presentations</td>
<td>Learn to build effective visualizations</td>
</tr>
<tr>
<td>Apply methods of network analysis to biological data</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Learn to build effective visualizations</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
<td>Gain experience mentoring others</td>
</tr>
<tr>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
<td>Develop leadership skills</td>
</tr>
<tr>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
<td>Develop leadership skills</td>
<td>Gain exposure to diverse career options</td>
</tr>
<tr>
<td>Build teamwork skills</td>
<td>Develop leadership skills</td>
<td>Gain exposure to diverse career options</td>
<td></td>
</tr>
<tr>
<td>Gain exposure to diverse career options</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMBINE Components</th>
<th>Interdisciplinary Research</th>
<th>Communication</th>
<th>Career Preparation</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMBINE Courses Data Practicum with Communication Training (PHYS798N)</td>
<td>Learn about important research results in network science</td>
<td>Learn about the theory and methods of network analysis</td>
<td>Develop an appreciation and understanding of the questions and methods of other fields</td>
</tr>
<tr>
<td>Computational and Mathematical Analysis of Networks (CMSC8280)</td>
<td>Learn about the theory and methods of network analysis</td>
<td>Develop an appreciation and understanding of the questions and methods of other fields</td>
<td></td>
</tr>
<tr>
<td>Seminar Courses Network Science</td>
<td>Develop methods of network analysis to biological data</td>
<td>Develop written communication skills</td>
<td></td>
</tr>
<tr>
<td>Literature Survey Seminar</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td></td>
</tr>
<tr>
<td>Seminar Courses Network Science Literature Survey Seminar</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Network Biology Research-in-Progress Seminar</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Elective Courses Discipline-bridging 3-4-credit elective course (chosen from list of options)</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Disciplined-bridging elective seminar (chosen from list of options)</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Other Activities Peer-to-peer tutorial</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>COMBINE Annual Symposium</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Career Development Workshop</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>COMBINE Outreach</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Outside research project/Internship</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
<tr>
<td>Other COMBINE-related element</td>
<td>Develop written communication skills</td>
<td>Communicate to a diverse audience</td>
<td>Build teamwork skills</td>
</tr>
</tbody>
</table>
Feedback from First Cohort

- Not enough opportunities for fellows to connect with each other outside of their scheduled courses or the required COMBINE activities.

  ✓ Leadership team began hosting a series of “coffee talks” wherein fellows were invited to socialize and collaborate with each other in a more social, non-academic setting.
Acknowledgements

• COMBINE and UMD Global STEWARDS Fellows, their advisors, and the leadership team

• The Evaluation Teams:
  • Jack Marr
  • Frances Turner
  • Megan Gerdes
  • Hannoori Jeong
Questions

• COMBINE: Poster #212 Session 2
• UMD Global STEWARDS: Poster #110 Session 1
Evaluation of graduate program courses: going beyond the typical end of course survey

2019 NRT Annual Meeting

Evaluators for RIPTIDES program at San Francisco State University

Poster # 124 (session 1)
Challenge

- Y1 evaluation focused on newly developed courses
  - Useful & actionable feedback for the program
Challenge

- **Y1 evaluation focused on newly developed courses**
  - Useful & actionable feedback for the program
- **Some issues related to the existing courses**
  - Not enough information
  - Disagreed with university course evaluation
Challenge

- Y1 evaluation focused on newly developed courses
  - Useful & actionable feedback for the program
- Some issues related to the existing courses
  - Not enough information
  - Disagreed with university course evaluation
- For Y2 look at **ALL** the core courses
Methodology:

• Student focus groups
• Student surveys
• Limited observations
• Internal memo of findings
• Faculty interviews
Key Findings:

• Improve interdisciplinarity
• Expectations didn’t always match
• More coordination across the courses
• More statistics training
• Less than ideal meeting day and time
Program Action

- Setting clearer expectations
- Writing course now specific for program students
- One course changed day & time
- One course redesigned to be a statistics course
- Faculty planning meeting for team taught course
Questions???

agrillo@wested.org