

# The “3Rs” of evaluation policy

Margaret A. Johnson, PhD

Cornell University

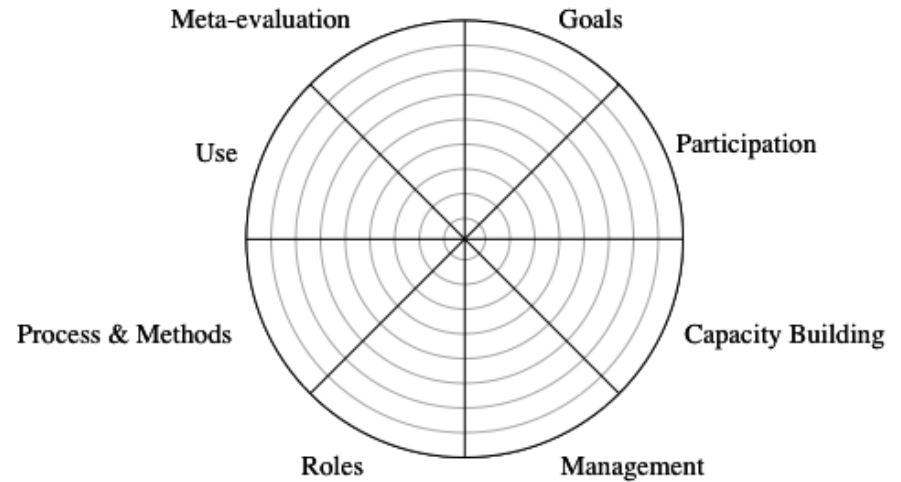
[maj35@cornell.edu](mailto:maj35@cornell.edu)

# Since 2008, we have heard a lot about **evaluation policy**



## **An Evaluation Roadmap for a More Effective Government**

September 2010



Trochim Policy Wheel

But think a minute...why would any organization (foundation, government, company) want to have evaluation policies?

You might say:

“Those are just rules.”

“Isn’t good evaluation \*tailored\* to each new situation?”

# Don't evaluation policies just...

kill intrinsic motivation?

encourage pro forma evaluation?

discourage evaluative thinking?

create backlash?

change whenever leadership changes?

# Or do they...

encourage high quality evaluation & use?

protect against political influences on the outcomes of evaluations?

help establish the conditions necessary for a healthy evaluation function in an organization?

help standardize evaluation practices across the organization & over time to allow for meaningful comparisons of results?

It depends....

But on what?

# Purpose of my research

Generate an **empirically grounded** taxonomy of evaluation policy with **workable classification criteria**

Applications:

- \*a **diagnostic** for organizations interested in building their evaluation system

- \*a **conceptual tool** for **building empirical evidence** about evaluation policies (what works when & where)

# Method for my study

Use concept mapping (Trochim, 1989)

1. respond to open-ended survey Q with ideas
2. sort ideas into conceptual “piles”
3. use MDS<sup>(1)</sup> to aggregate sorting decisions & translate to a dot map in xy space



# Next...

4. Use HCA<sup>(2)</sup> to explore different possible groupings of ideas on the map
5. choose a cluster solution
6. interpret
7. use as basis for group action

# Participant brainstorming

*“In a comprehensive set of U.S. federal evaluation policies, one policy that should be included is...”*

# Data

- In May 2009, invited a random sample of 2000 AEA members

## Response:

- 554 volunteer brainstormers
- 63 volunteer raters
- 20 volunteer sorters

# Sample

- roughly comparable to AEA membership @May 2009
- on level of education  
on primary work setting  
on main work activity  
on country of residence (US, non-US)
- but probably disproportionately includes those with  
a keen interest in U.S. federal evaluation policy

Here are some examples of evaluation  
policy ideas from the AEA members in my  
survey...

# One evaluation policy that should be included is...

“A requirement that most evaluations should use a random experimental design.”

“Redefine the "gold standard" for study design from RCT to fit between context and methods.”

“The standard for the evaluation of social policy should be accountab[ility] for advancing the goals of social justice.”

# One evaluation policy that should be included is...

“Standardize the evaluation language so that the evaluation work under one federal agency can be compared with the work under another agency”

“Rules and principles at the Federal level shall be devised in such a way as to set appropriate standards for evaluations at state and local levels.”

# One policy that should be included is...

- “Provide technical assistance to foster evaluation capacity building within organizations and agencies.”
- “A separate evaluation contract shall be included in major grants to ensure that funds available for evaluation are not reallocated to program efforts.”
- “The government shall periodically assess the adequacy of the federal evaluation workforce.”



# One policy that should be included is...

- “Involve professional evaluators in reviewing the evaluation plans submitted on federal grants.”
- “Make the position of evaluator a career civil service one, instead of politically appointed.”
- “Evaluators reporting evidence of coercion through the exercise of power, influence, or resources shall be protected under Federal Whistle Blowers Laws.”

# One policy that should be included is...

- “A ‘chief program evaluation officer’ (CPEO) shall be appointed for each federal agency, and shall serve on a council with reporting requirements to Congress and the President.”
- “A contract for federal program evaluation shall be overseen by a branch or division that is over the entity being evaluated, not by the program that is the subject of the evaluation.”
- “An independent agency shall be established to prioritize and systematically address attainment of pre-established goals and benchmarks of other federal agencies.”

So you can see from these examples  
AEA members think  
evaluation policy includes rules...  
but is more than just rules.

# Synthesis & reduction for analysis

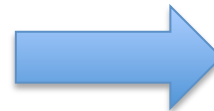
920 policy statements (too many for participant sorting!)



emergent, iterative construction of categories  
(*constant comparison*)

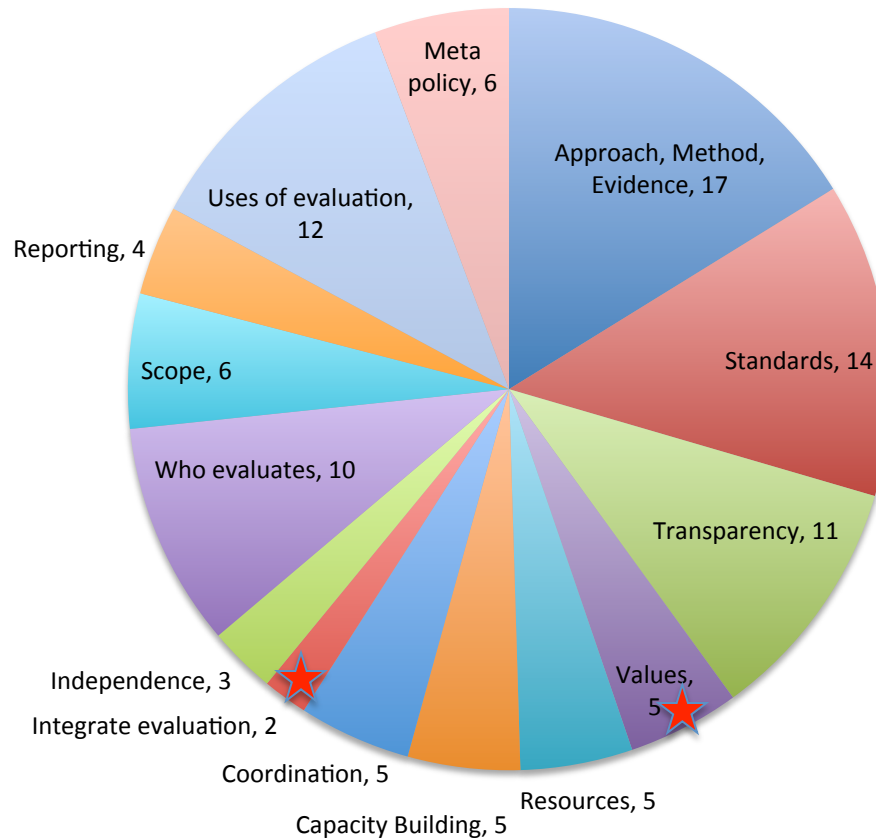


iterative reduction  
(*sampling for range*)



100 policy statements  
(*sortable*)

# Through careful coding, broad range preserved



# Participant sorting

“Please place these evaluation policy statements in categories according to meaning or theme, in a way that makes sense to YOU. Give each category a name of its own”

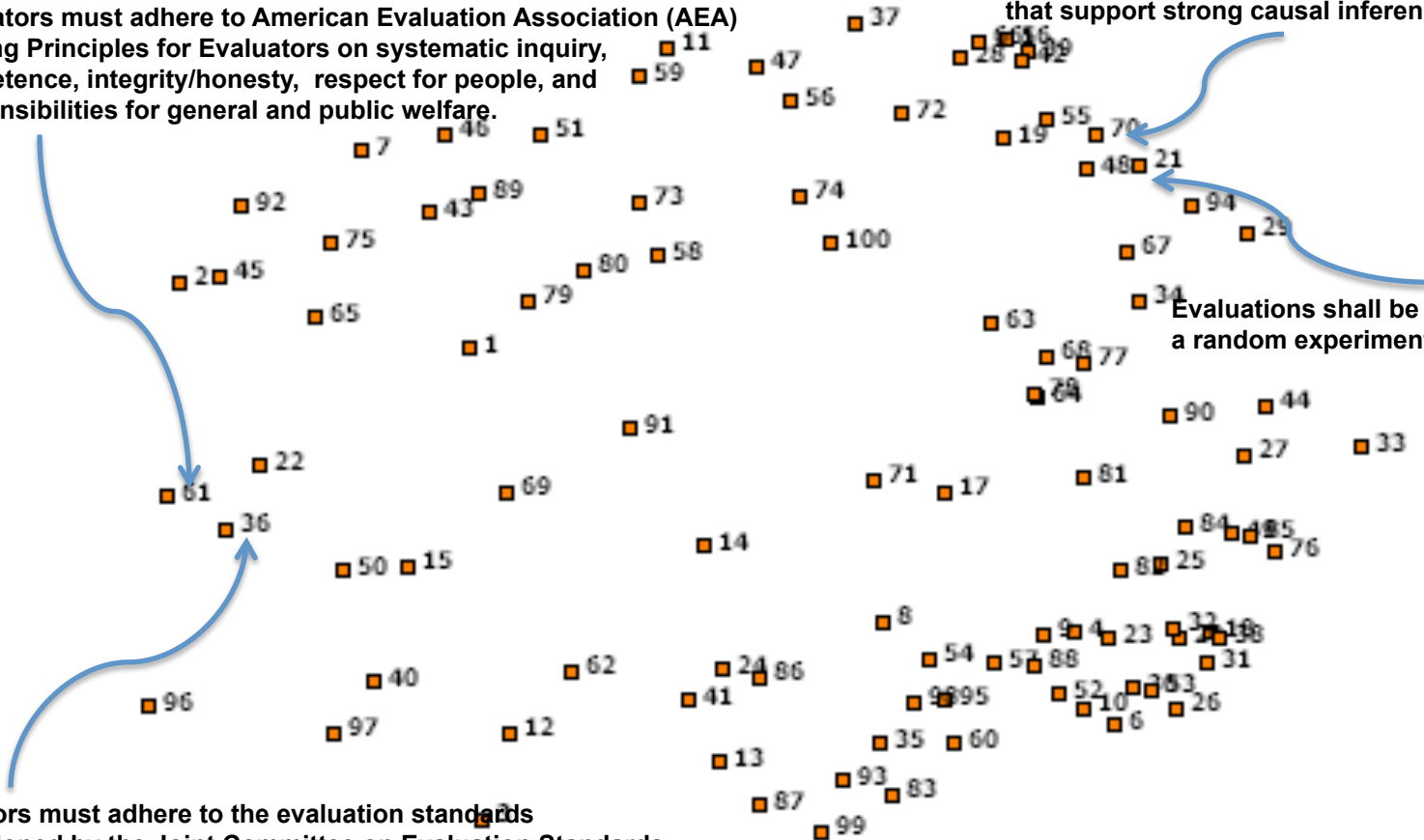
# Initial concept map results

Evaluators must adhere to American Evaluation Association (AEA) Guiding Principles for Evaluators on systematic inquiry, competence, integrity/honesty, respect for people, and responsibilities for general and public welfare.

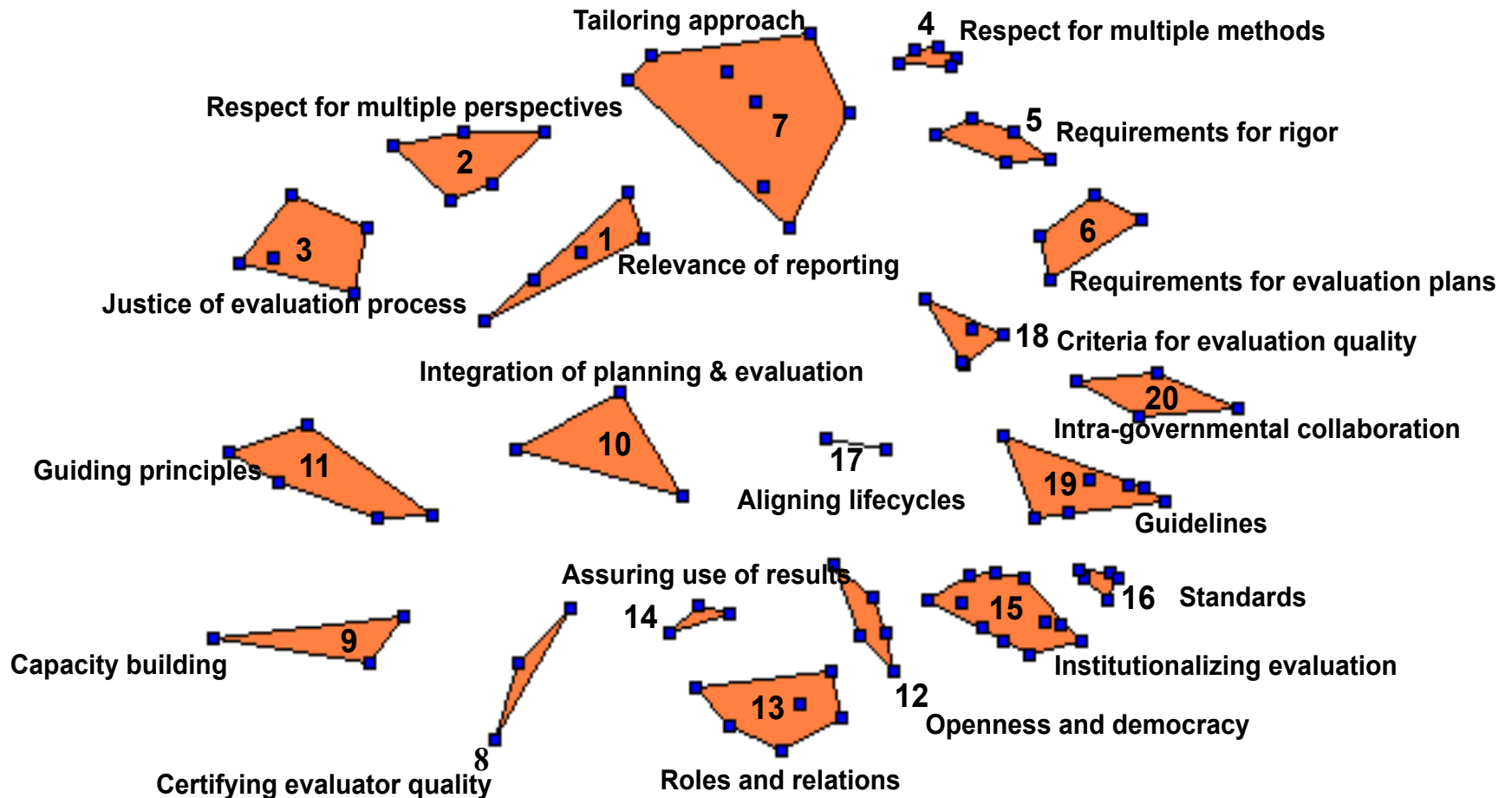
Impact evaluations must utilize research designs that support strong causal inference.

Evaluations shall be required to use a random experimental design.

Evaluators must adhere to the evaluation standards as developed by the Joint Committee on Evaluation Standards at the Evaluation Center (JSEC), Western Michigan University, on utility, feasibility, propriety, and accuracy.

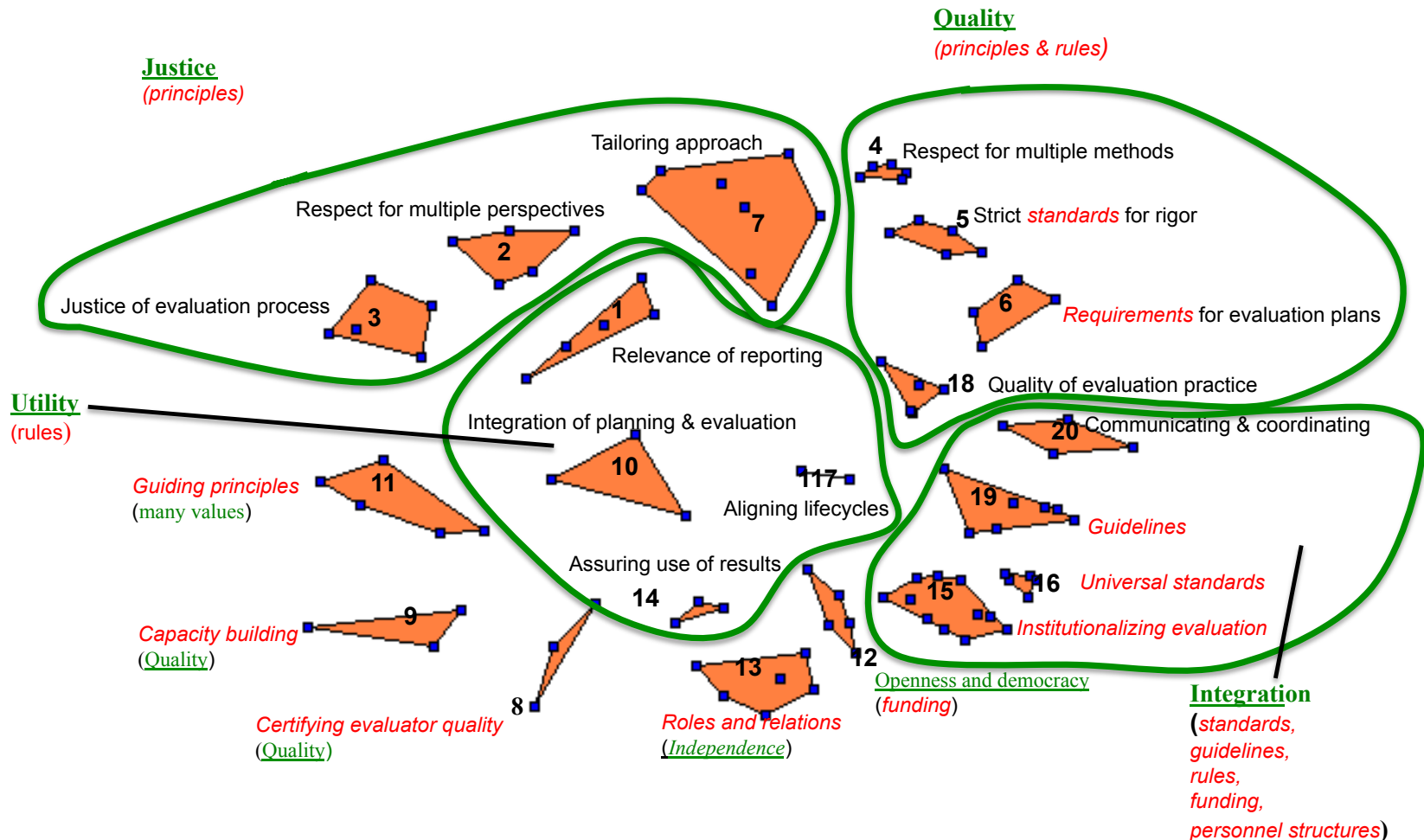


# Cluster solution





# 2 dimensions: **values** and **policy mechanisms**



That was pretty messy.

How would you use it?

How about something we can use?

# “3 Rs”

## evaluation policy framework

	Justice	Utility	Quality	Integration
<b>Rules</b> (principles, standards or guidelines for practice)				
<b>Resourcing</b> (time, money, capacity building)				
<b>Role structures</b> (independence, oversight)				

# “3Rs” definition of evaluation policy

A rule, resourcing strategy or role structure used by an organization to support evaluation, including evaluation planning, implementation, reporting or use.

May be explicit (formal, written)

May be implicit (“rule of thumb”)

# Using the “3Rs” framework

1. Gather the organization’s evaluation policies & examine them together

Assign each policy to a cell on the “3Rs” framework.

- Having to choose a column forces you to think about what values or goals a particular policy is really promoting
- Having to choose a row can show imbalances (e.g. lots of rules but no funding structures or roles to support their implementation)

# Using the “3Rs” framework

2. Assess needs & feasibility

3. ID explicit evaluation policies

4. ID implicit evaluation policies

# Using the “3Rs” framework

5. Evaluate –  
complete, harmonized & aligned?
6. Make a plan for change & implement!

# 1: Form a working group

...to take stock of the organization's evaluation policy.

Include key staff.

Include key stakeholders.

Identify organization's core mission & values.

→ These values could substitute for the column headings in the "3Rs" matrix.



## 2: Assess needs & feasibility

Does the working group:

- see problems in the organization's approach to evaluation?

- have specific goals for the organization related to evaluation?

- (e.g. improving performance by raising evaluation quality)

- see the organization as able to change key evaluation policies affecting it?

# 3: ID explicit evaluation policies

- in organizational charters, strategic plans
- in staff job descriptions
- in written guidance
- in incentives
- in budget lines
- in funding criteria

Make sure to also include evaluation policies from outside the organization that influence its evaluation practice & use.

## 4: ID \*implicit\* evaluation policies

- in past evaluation reports
- in interviews with staff
- in public rhetoric of leaders

# 5: Evaluate

Are evaluation policies simply missing?

Are eval policies out of sync with eval practices?

Are explicit eval policies out of sync with implicit eval policies?

Which values does the current set of evaluation policies reflect?

Do these values align with the organization's overall values & mission or clash?

Which of the “3 Rs” policy mechanisms are in use?

Have these been effective tools in this organization?

# 6: Make a plan for change

Compare initial working group needs & goals to inventory findings and amend if needed.

For each issue, note barriers to change and possible solutions.

Identify priorities and make an action plan to develop or change evaluation policies.

# Using the “3Rs” framework to build evidence about what works

Test “3Rs” framework to see if inclusion rules for each cell can reliably be used by many different people to classify evaluation policies; refine framework.

Conduct inventories of real-world organizational evaluation policies; refine step-by-step instructions.

Compile a clearinghouse of real-world evaluation policies.

Test hypotheses about which evaluation policies work best in different situations. Find out what “it depends” on!

# References

- W.M.K. Trochim, M. M. Mark, & L. J. Cooksey (Eds.), Evaluation policy and evaluation practice. New Directions for Evaluation, vol. 123 (Fall 2009)
- Kane, Mary and William M.K. Trochim (2007) Concept Mapping for Planning and Evaluation, in Applied Social Research Methods Series, eds. Leonard Bickman and Debra J. Rog
- Datta, Lois-ellen (2011) Politics and evaluation: more than methodology *American Journal of Evaluation* 32(2) pp. 273-294
- Mitchell, Douglas and Marshall, C., Wirt, F. (1985) Building a taxonomy of state education policies, *Peabody Journal of Education*, 62(4) p. 7-47

# (1) Multidimensional scaling

MDS attempts to locate  $n$  observations in a reduced dimensional space while minimizing the discrepancy of the (quantitative) differences between pairs of points in this reduced space from the true (quantitative) differences of the observations

MDS with non-metric scaling in 2 dimensions:

item-to-item similarity matrix  $\rightarrow$  MDS  $\rightarrow$  coordinate matrix  $(x,y)$



## (2) Hierarchical cluster analysis

Hierarchical cluster analysis using Ward's algorithm takes observational data in coordinate form and starts with one cluster per observation, generating a successively smaller number of larger clusters, while minimizing the total distance between statements inside each cluster.

The initial cluster distances in Ward's minimum variance method are defined to be the squared Euclidean distance between points.

coordinate matrix  $\rightarrow$  HCA (Ward's algorithm)  $\rightarrow$ , given  $n$  sets, produces a series of cluster solutions, successively considering the union of all possible  $n(n-1)/2$  sets