



*'Landscapes'*

# The Organizational and Political Landscape for Evaluation: The Mantra of “Evidence-based Policy Making”

**Trachtenberg School  
of Public Policy  
& Public Administration**

**THE GEORGE WASHINGTON UNIVERSITY**

Kathryn Newcomer  
September, 2016

# Thank you for This Valuable Learning Experience!

- ▶ Hospitality
- ▶ Inclusiveness
- ▶ Energy
- ▶ Warmth
- ▶ Intellectual Sparks
- ▶ Optimism



# “Evidence-based Policy,” “Data-Driven Decision-making” – the New Normal?



# Questions to Address Today

- ▶ Where are we in the pursuit of using evidence available about when and how government programs works to inform decision-making?
- ▶ What have we learned about how and when decision-makers in government learn from evidence?
- ▶ How can evaluators help government decision-makers use evidence to inform decision-making?



# “Evidence-based Policy”

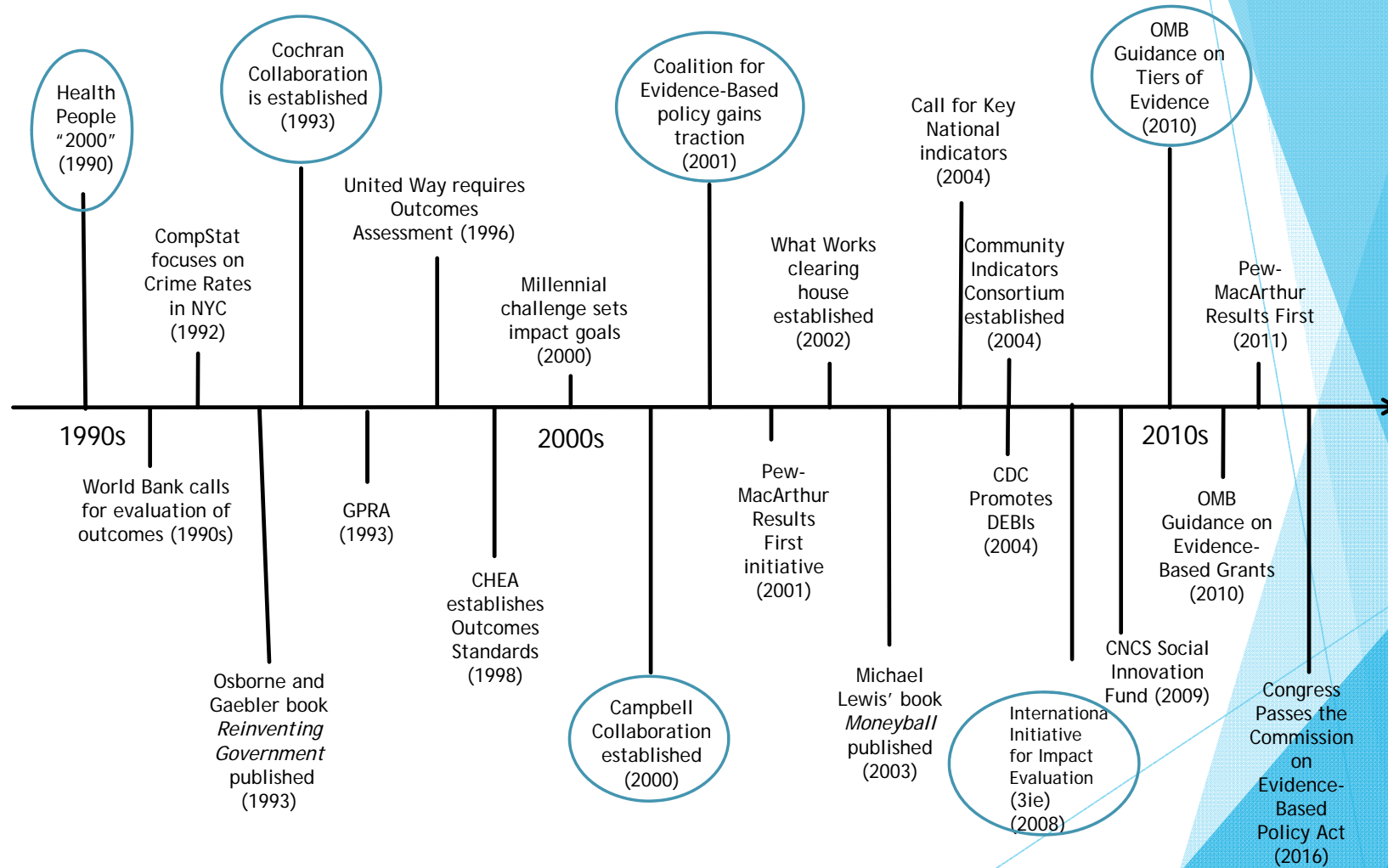
- ▶ The Mantra affecting governmental decision-makers, foundations, nonprofit boards, intermediaries and --- evaluation practice!
- ▶ Myth or reality?
- ▶ Advantages and disadvantages for evaluators?

# Efforts in Several Arenas Have Moved the Dialogue to Embrace “Evidence-Based Policy”

- ▶ Since the 1960s in the U.S. dialogue about the target for government’s efforts has changed from a focus on effectiveness to outcomes to results to evidence ---- Why?
  - ▶ Tracking of diseases in Public Health, e.g., “Healthy People 2000,” and The Cochran Collaboration
  - ▶ North American and European Social Scientists Established the Campbell Collaboration to mimic The Cochran Collaboration
  - ▶ The U.S. Office of Management and Budget’s efforts to Assess “evidence of program effectiveness” in both the George W. Bush and Barack Obama Administrations
  - ▶ Leading Foundations have invested resources to encourage evidence-based decision-making, e.g., Pew, MacArthur, Arnold, and Grant
  - ▶ Evaluation and Monitoring of International Development Efforts



# Embracing Evidence-Based Policy: Influential Events Across the Years

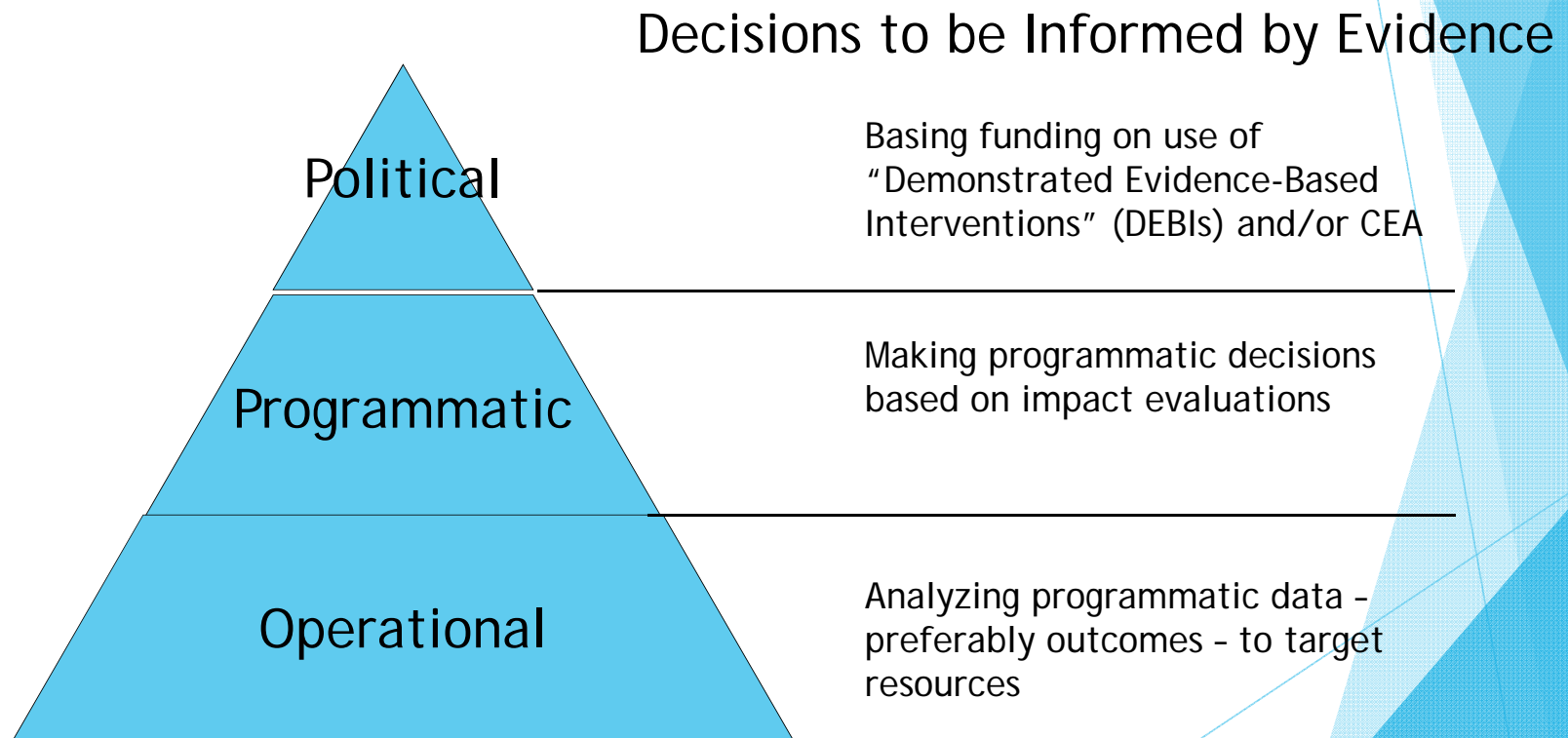


# What is the Underlying Theory about Evidence-Based Policy?

- ▶ In an idealized world, the process by which research and evaluation informs policy would be fairly straightforward:
  - ▶ Researchers and evaluators would conduct methodologically rigorous and relevant policy-oriented studies.
  - ▶ These researchers/evaluators would produce study deliverables that clearly explain the research methods, key findings, and policy implications using translational research principles.
  - ▶ Policymakers would be educated in assessing the quality of research and understanding the policy implications of the research.
  - ▶ These educated policymakers would review policy-oriented research and evaluations, determine how best to interpret the findings and recommendations, and enact policies or programs based on the recommendations/findings.
  - ▶ **However**, in the real world, there are numerous mitigating factors at each stage of this idealized process, as well as factors that make it difficult for researchers/evaluators to study the issues when the policymakers are interested in the issues.



# Evidence-Based Policy – Made by Whom?

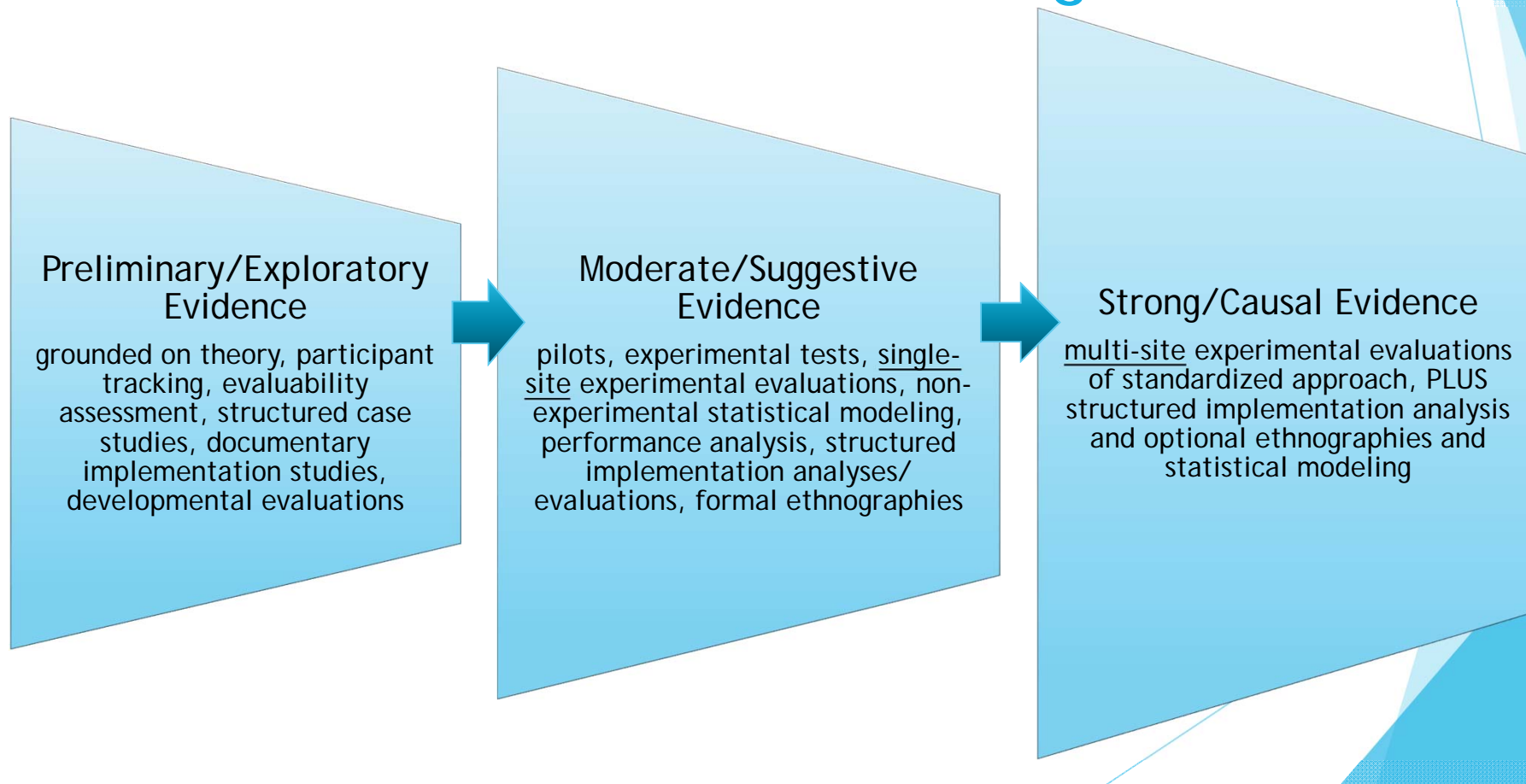


# Obama Administration: Explicit Emphasis on Producing and Acting on Evidence

- ▶ A series of office Memoranda from OMB between 2009 and 2013 signaled that performance measurement and evaluation were to be used to produce "evidence on what works" (OMB, 2010; OMB, 2011a; OMB 2011b; OMB, 2012a; OMB, 2012b; and OMB, 2013)
- ▶ OMB Circular A-11 defines evidence for the federal government:
  - ▶ "For purpose of A-11 Part 6, evidence is the available body of facts or information indicating whether a belief or proposition is true or valid. Evidence can be quantitative or qualitative and may come from a variety of sources, including performance measurement, evaluations, statistical series, retrospective reviews and other data analytics and research. Evidence has varying degrees of credibility, and the strongest evidence generally comes from a portfolio of high-quality evidence rather than a single study."



# The Obama Administration Recognizes Tiers of Evidence Based on Perceived Rigour



# What are the Opportunities for Evidence to Inform Policymaking?

- ▶ Analyses of “performance” data collected by agencies (or delegated service delivery agents such as grantees)
- ▶ Implementation, Outcome and Impact evaluations typically performed by other agents for government
- ▶ Manipulations of services in experiments by agencies – “behavioral economics”
- ▶ Syntheses or systematic reviews of impact evaluations by external agents, e.g. websites like “What Works”
- ▶ Other?



# What are Challenges for Evidence to Inform Policymaking?

## ► Expectations

- What constitutes evidence?
- How transferable is evidence?
- When and where do we underestimate the role played by the “impactees?”
- Where is the capacity to support both the demand and supply of evidence?

# Why isn't There Agreement About the Quality of Evidence?

- ▶ Differing professional standards and “rules” or criteria for evidence, e.g., lawyers, accountants, engineers, economists
- ▶ Disagreements about methodologies within professional groups, e.g., RCTs
- ▶ The constancy of change in problems and the characteristics of the targeted impacttees



# We Underestimate the Evolving Sources of Complexity Affecting the Production of Relevant Evidence

- ▶ Change in the nature of problems to be addressed by government, e.g., the nature of natural security threats, the use of the internet in crime
- ▶ Change in the context in which programs and policies are implemented, e.g., increasingly complicated service delivery networks, PPPs
- ▶ Changing priorities of political leaders (and would-be leaders)

# We Overstate the Certainty of the Evidence we Can Collect

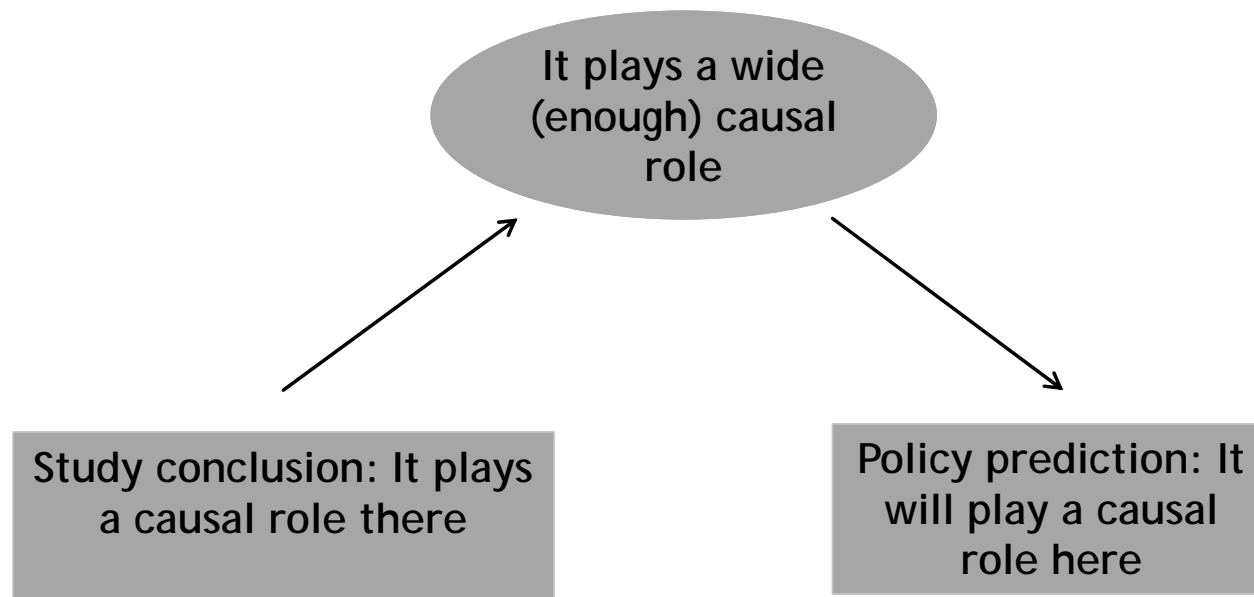


- Perceptions of the certainty of "evidence" have changed.



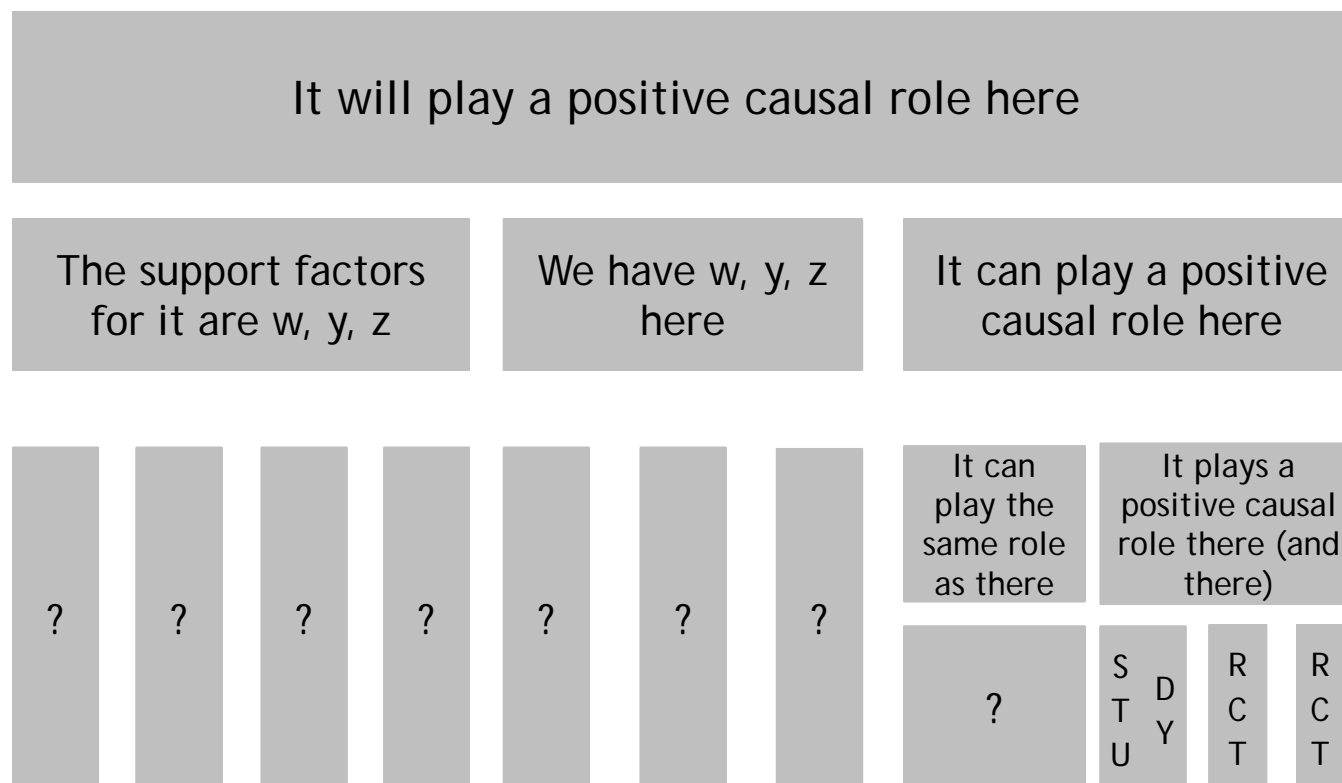


# We Overstate the Ease of Flow of Evidence



Source: Cartwright, N. (2013). Knowing what we are talking about: why evidence doesn't always travel. *Evidence & Policy: A Journal of Research, Debate and Practice*, 9(1), 97-112.

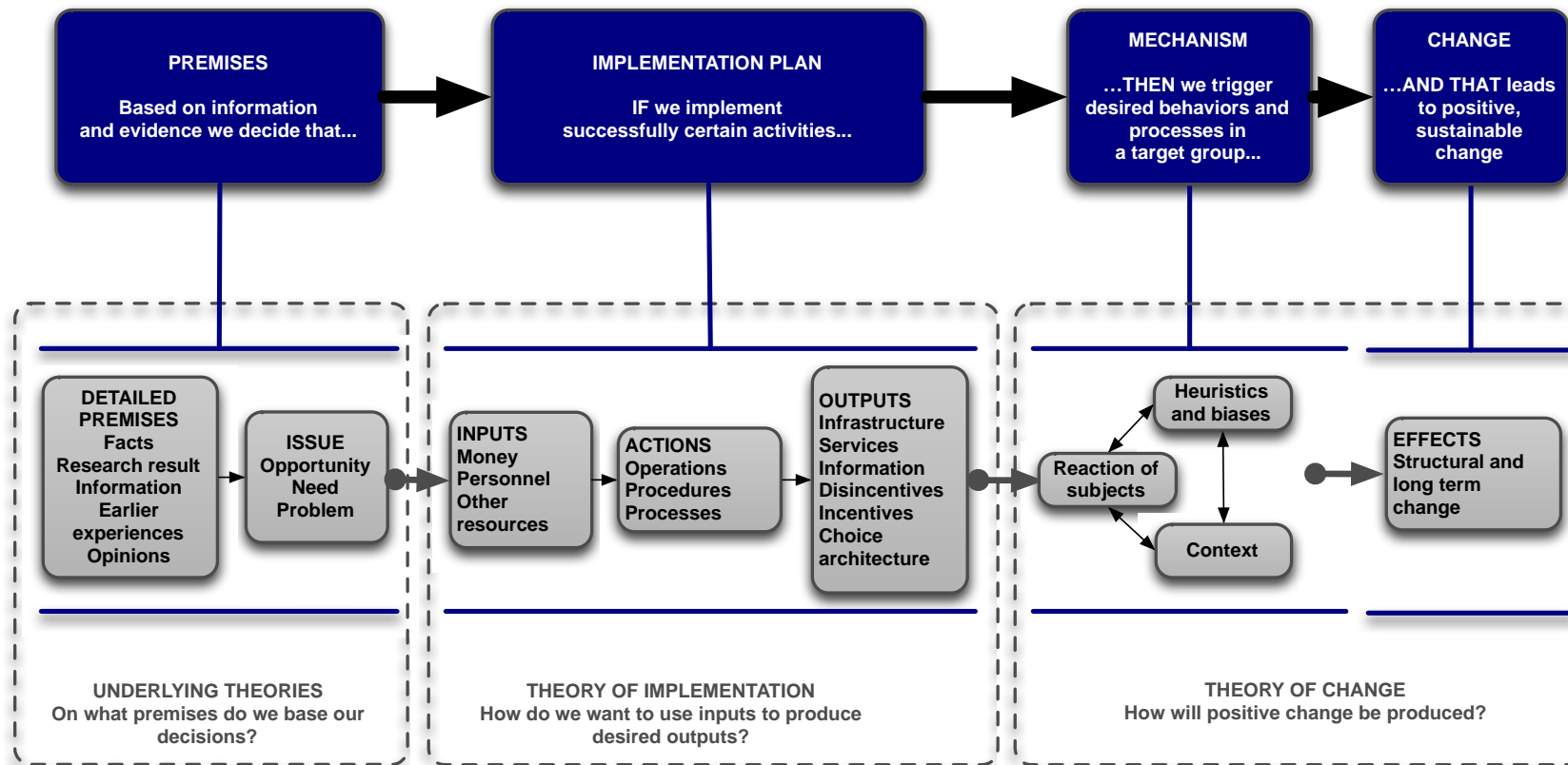
# What is needed for a well-supported effectiveness (impact) prediction?



Source: Cartwright, N. (2013). Knowing what we are talking about: why evidence doesn't always travel. *Evidence & Policy: A Journal of Research, Debate and Practice*, 9(1), 97-112.



# We Underestimate the Role of Volition Among Impactees and their Own Heuristics



Olejniczak, K. & Newcomer, K. (2014). "Moving towards accountability for learning";  
 in: Olejniczak, K. & Mazur, S. (eds.) Organizational Learning. A Framework for Public Administration, p.81-99. Warsaw: Scholar Publishing House.

# We Overstate The Current Evaluation Capacity in Government





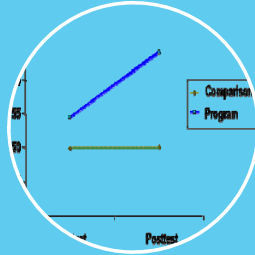
# Evaluation Capacity = Both Demand and Supply

- ▶ Who is asking for the evidence?
- ▶ How clear is the understanding between providers and requestors on what evidence is needed?
- ▶ Are there sufficient resources within agencies to respond to demand?
- ▶ What about the lack of interaction and synergies among the different potential providers of evidence - such as in the U.S. GPRA/GPRAMA reporting staff, internal evaluation staff, external evaluation contractors, SBST, data.gov teams, etc.!

# Currently Multiple Groups Undertake Monitoring and Evaluation in and for U.S. Government Agencies



Monitoring



Impact  
Evaluation



Behavioral  
Economics

They tend to operate in separate and even uncommunicative units with competing priorities!



# There are Signs of Progress in Using Evidence in Government

- ▶ Reputable Drivers at putting resources into efforts, e.g.:
  - ▶ The Pew MacArthur Results First Initiative
  - ▶ The National Academy of Sciences, Engineering, and Medicine May 2016 report: *Advancing the Power of Economic Evidence to Inform Investments in Children, Youth, and Families*
  - ▶ The Arnold and William Grant Foundations
- ▶ Recent National Science Foundation support of initiatives to help policy researchers translate their findings for government users
- ▶ Professional Associations are supporting translational efforts, e.g., APPAM, AEA
- ▶ Communities of Practice abound, especially in public health
- ▶ The Commission on Evidence-Based Policy seems to be inclusive in terms of considering what constitutes evidence

# Promising Practices from the Obama Administration

Promising Practice	Affects Supply or Demand?	Needed Support Factors
Knowledge Brokers	Both	Brokers have technical expertise, interpersonal skills, and contextual wisdom
Learning Agendas	Demand	Strong leadership backing and encouragement to be innovative
Quarterly Reviews	Supply	Credible data, stress on learning, no punitive actions
Strategic Reviews	Both	Encouragement to be innovative, stress on learning not accountability



# NAPA Transition Paper on Using Evidence-Based Approaches to Improve Program Performance

- ▶ Objective: Improve government performance by strengthening the use of data, evidence, evaluation, and innovation by government leaders, managers, front-line employees, those involved in service delivery, and other stakeholders in the allocation of resources and management of programs.

# NAPA Recommendations

1. Create a performance management framework that integrates performance management, program evaluation, futures planning, and budgeting; drive continuous improvement with benchmarking and other tools.
2. Appoint agency deputies/Chief Operating Officers and other political appointees who have the capacity and make an explicit commitment to drive performance improvement and build evidence-based decision making.
3. Empower strong Performance Improvement Officers with adequate resources to support agency Deputies/Chief Operating Officers; where the Performance Improvement Officer has other duties, ensure there is a strong Deputy Performance Improvement Officer devoting full attention to the adoption of the integrated performance management framework.
4. Continue to drive the conduct and use of appropriately tailored rigorous, independent, and relevant evaluations to improve programs; ensure a robust evaluation and data analytics capacity in every Department that, in addition to conducting rigorous program evaluations, also carries out evaluations and analyses related to and useful for performance management.
5. Enhance the accessibility, transparency, and usefulness of performance information by simplifying Performance.gov, showing clear trends in government's progress, and linking it to other, relevant sources of information about the government's performance.



# NAPA Recommendations, cont.

6. Collaborate with Congress (i.e., authorizers, appropriators, and overseers) more closely at every stage of the performance management process to facilitate more debate about the performance of programs and successful adoption of the performance management framework.
7. Encourage state and local government /other partner and stakeholder adoption of the performance management framework by supporting easy access to data and multi-stakeholder collaborations; use incentives in the grant making and contracting process for this purpose.
8. Institute annual OMB spring reviews to assess and accelerate progress on strategic goals; enhance OMB quarterly priority goal reviews to assess progress and identify opportunities to improve.
9. Ensure the Office of Management and Budget is driving the development, adoption, and implementation of cross-agency priority goals.
10. Develop a clear and concise performance management curriculum for agency and other professionals; enlist front-line employees, through employee forums or other means, to identify ways to improve performance and strengthen the adoption of the performance management framework.

# The NAPA Recommendations Reflect the Political Context for Evidence-Based Policy Making in the U.S.

- ▶ Program managers and other decision makers are caught between two masters- The President and Congress, and these entities have different priorities and values
- ▶ The implementation of virtually all federal programs and policies is undertaken through states, local governments, nonprofits and even private agents
- ▶ Federalism affects the flow of money to implement federal policies and programs -for example, formula grants given to states are hard to change into evidence-based grants
- ▶ The President relies on his or her Office of Management and Budget to “drive management reforms” and it is hard to not have the rest of government view these directives as compliance exercises



# How and When do Decision-makers learn from Evidence?

## Information Processing

- Automatic Operations ("Fast Thinking")
- Controlled Operations (Slow)
  - Worldview & Epistemology
  - Expertise
- Judgmental Heuristics
- Emotional State
- Presentism
  - Pure rate of time preference

## Transmission Process

- Brokering/delivering the information
- Priming (timing matters!)
- Timeliness of Access
- Presentation of Data
  - Logic visualization
  - Data visualization

## Nature of Information

- Source of Evidence
  - Presumed credibility
  - Reputed credibility
- Trustworthiness of Evidence
  - Weight of evidence
  - Strength of evidence
  - Reliability of data
- Match Between Evidence and Receiver's Epistemological Preferences
- Signaling about Priorities in Research Designs Rigor from Respected Sources

## Organization and Social Context

- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>-Organizational Culture<ul style="list-style-type: none"><li>• Leadership modeling of use of evidence</li><li>• Priority given data in decision making</li><li>• Focus on learning</li><li>• Support for risk taking</li><li>• Treatment of "errors"</li></ul></li></ul> | <ul style="list-style-type: none"><li>-Social Supports<ul style="list-style-type: none"><li>• Similarity in worldviews within group</li><li>• "Like-minded" peers</li><li>• Priority given to diversity of views</li></ul></li></ul> |
|--|--|

# How Can Evaluators Contribute to Helping Decision-makers learn from Evidence?

## Information Processing

- Automatic Operations ("Fast Thinking")
- Controlled Operations (Slow)
  - Worldview & Epistemology
  - Expertise
- Judgmental Heuristics
- Emotional State
- Presentism
  - Pure rate of time preference

## Organization and Social Context

- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>-Organizational Culture<ul style="list-style-type: none"><li>• Leadership modeling of use of evidence</li><li>• Priority given data in decision making</li><li>• Focus on learning</li><li>• Support for risk taking</li><li>• Treatment of "errors"</li></ul></li></ul> | <ul style="list-style-type: none"><li>-Social Supports<ul style="list-style-type: none"><li>• Similarity in worldviews within group</li><li>• "Like-minded" peers</li><li>• Priority given to diversity of views</li></ul></li></ul> |
|--|--|

## Transmission Process

- Brokering/delivering the information
- Priming (timing matters!)
- Timeliness of Access
- Presentation of Data
  - Logic visualization
  - Data visualization

## Nature of Information

- Source of Evidence
  - Presumed credibility
  - Reputed credibility
- Trustworthiness of Evidence
  - Weight of evidence
  - Strength of evidence
  - Reliability of data
- Match Between Evidence and Receiver's Epistemological Preferences
- Signaling about Priorities in Research Designs Rigor from Respected Sources



# Transmission Process

- ▶ Just as there are many producers, there are many potential users of the evidence provided, e.g., different policy designer and implementers in complex service delivery and regulatory networks units
- ▶ **Knowledge brokering** is critical
- ▶ Understanding and strengthening the linkage between the producers of evaluative data and the many potential users of that information requires time and resources
- ▶ For Example: the network of 57 evaluation brokering units in Poland overseeing 900 evaluations of EU cohesion policy investments

# How do we Meet Challenges for Evidence to Inform Decision-Making?

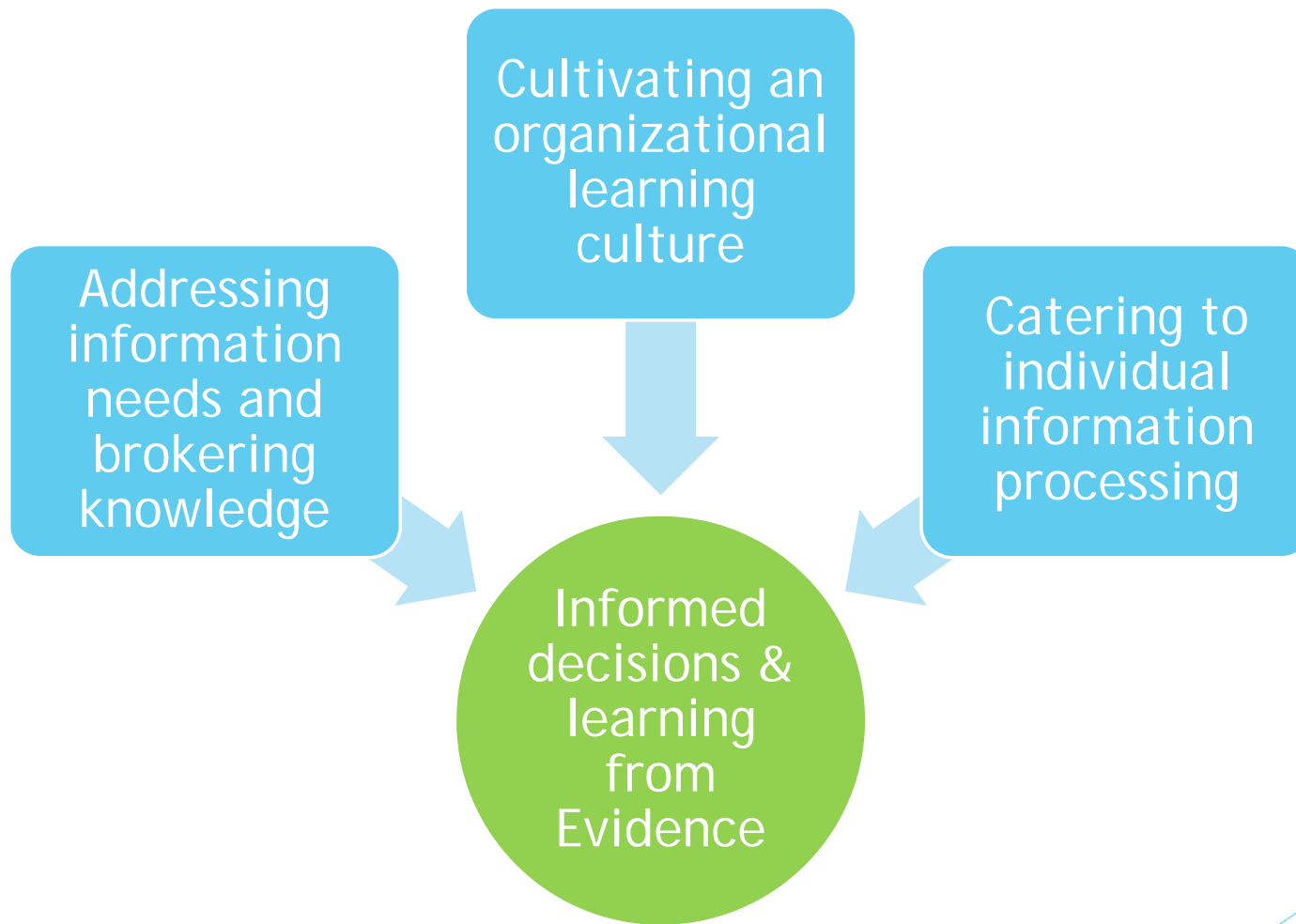


**MEET THE CHALLENGE!**





## A simple framework....

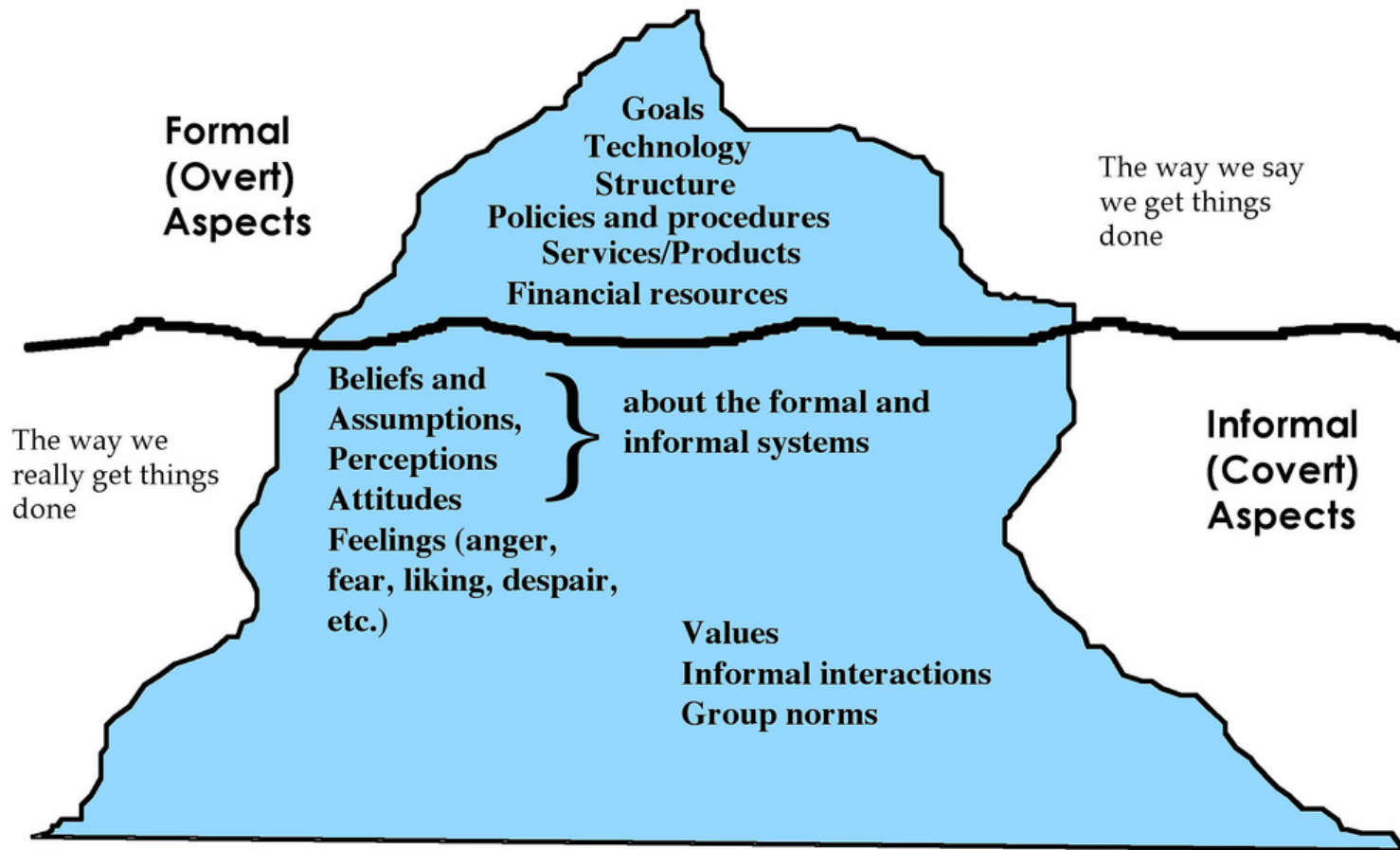


# Remember Evaluation Capacity = Both Demand and Supply

- ▶ Consider who is asking for the data/evidence and **who** might use the information provided and **how** and **when** they may use it
- ▶ Probe the extent to which there is a clear understanding between providers and requestors for what sorts of evidence is needed, e.g., brokering
- ▶ Assess whether or not sufficient resources are available to meet demand
- ▶ Address the lack of interaction and facilitate synergies among the different potential providers of evidence - such as monitoring and reporting staff, internal evaluation staff, external evaluation contractors, etc.



# Organizational Culture is Difficult to Change



*Created by Stanley N. Herman. TRW Systems Group, 1970*

# What are Evaluation-Receptive Organizational Cultures?

- ▶ Engage in self-reflection & self-examination
  - ▶ Deliberately seek evidence on what it's doing
  - ▶ Use results information to challenge or support what it's doing
  - ▶ Promote candor, challenge and genuine dialogue
- ▶ Engage in evidence-based learning
  - ▶ Make time to learn
  - ▶ Learn from mistakes and failures
  - ▶ Encourage knowledge sharing
- ▶ Encourage experimentation and change
  - ▶ Support deliberate risk-taking
  - ▶ Seek out new ways of doing business

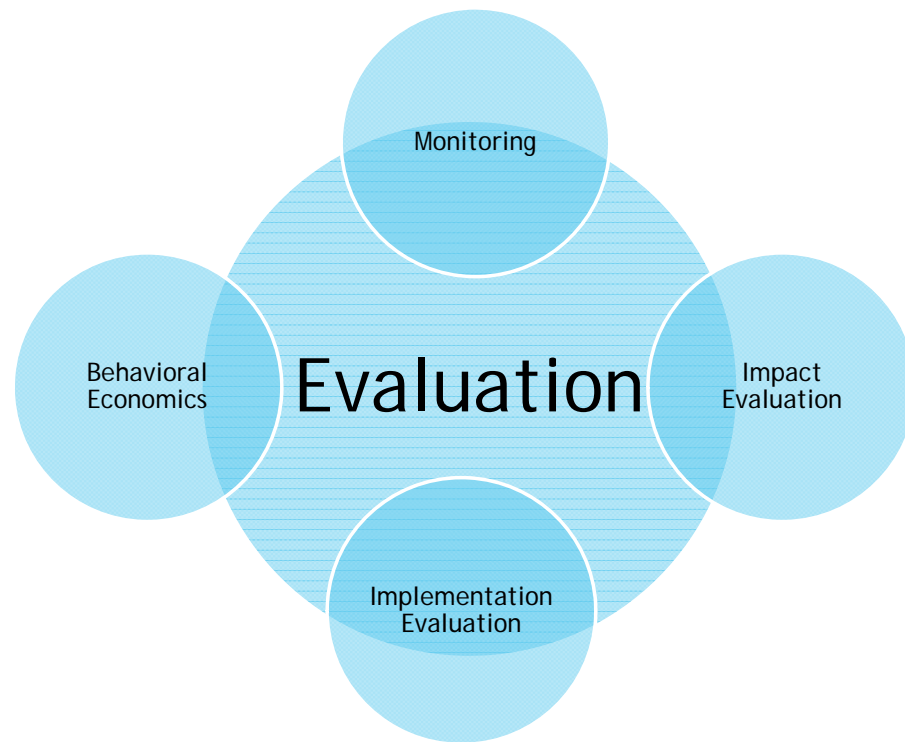
(See John Mayne, 2010)



# Be Strategic and Intentional about Cultivating Evaluation-Receptive Cultures

- ▶ Assess and address the factors perpetuating a compliance mentality among potential users, especially clients
- ▶ Reward learning from monitoring and evaluation, e.g., Learning Audits in the Netherlands
- ▶ Cultivate capacity to support both the demand and supply of information, e.g., the Canadian approach
- ▶ Match evaluation approaches to questions appropriately and transparently
- ▶ Reward mixed methods approaches that integrate data collected via differing methods

# Move To Strategic and Synergistic Use of Evaluation!



For Example, the U.S. Department of Labor.



# Help Users Frame Pertinent Questions and then Match the Questions with the Appropriate Evaluation Approach

Questions Relevant  
to Users



Evaluation  
Design

# Match Evaluation Approach to Questions

Objective	Illustrative Questions	Possible Design
<b>#1: Describe program activities</b>	<ul style="list-style-type: none"> <li>• How extensive and costly are the program activities?</li> <li>• How do implementation efforts vary across sites, beneficiaries, regions?</li> <li>• Has the program been implemented sufficiently to be evaluated?</li> </ul>	<ul style="list-style-type: none"> <li>• Monitoring</li> <li>• Exploratory Evaluations</li> <li>• Evaluability Assessments</li> <li>• Multiple Case Studies</li> </ul>
<b>#2: Probe targeting &amp; implementation</b>	<ul style="list-style-type: none"> <li>• How closely are the protocols implemented with fidelity to the original design?</li> <li>• What key contextual factors are likely to affect achievement of intended outcomes?</li> <li>• How do contextual constraints affect the implementation of a intervention?</li> <li>• How does a new intervention interact with other potential solutions to recognized problems?</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple Case Studies</li> <li>• Implementation or Process evaluations</li> <li>• Performance Audits</li> <li>• Compliance Audits</li> <li>• Problem-Driven Iterative Adaptation</li> </ul>
<b>#3: Measure the impact of policies &amp; programs</b>	<ul style="list-style-type: none"> <li>• What are the average effects across different implementations of the intervention?</li> <li>• Has implementation of the program or policy produced results consistent with its design (espoused purpose)?</li> <li>• Is the implementation strategy more (or less) effective in relation to its costs?</li> </ul>	<ul style="list-style-type: none"> <li>• Experimental Designs/RCTs</li> <li>• Non-experimental Designs: Difference-in-difference, Propensity score matching, etc.</li> <li>• Cost-effectiveness &amp; Benefit Cost Analysis</li> <li>• Systematic Reviews &amp; Meta-Analyses</li> </ul>
<b>#4 : Explain how/ why programs &amp; policies produce (un)intended effects</b>	<ul style="list-style-type: none"> <li>• How/why did the program have the intended effects?</li> <li>• To what extent has implementation of the program had important unanticipated negative spillover effects?</li> <li>• How likely is it that the program will have similar effects in other communities or in the future?</li> </ul>	<ul style="list-style-type: none"> <li>• Impact Pathways and Process tracing</li> <li>• System dynamics</li> <li>• Configurational analysis,</li> </ul>



# Thank You!

I can be reached at [newcomer@gwu.edu](mailto:newcomer@gwu.edu)

**Trachtenberg School  
of Public Policy  
& Public Administration**  

---

**THE GEORGE WASHINGTON UNIVERSITY**

# Relevant References

- ▶ Dahler-Larsen, Peter. 2012. *The Evaluation Society*. Stanford University Press.
- ▶ Donaldson, S., C. Christie, and M. Mark (editors) 2015. *Credible and Actionable Evidence*, 2<sup>nd</sup> Edition. Sage.
- ▶ Head, B. 2015. "Toward More "Evidence-Informed" Policy Making?" *Public Administration Review*. Vol.76, Issue 3, pp. 472-484.
- ▶ Kahneman, D. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux Publishers.
- ▶ Mayne, J. 2010. "Building an evaluative culture: The key to effective evaluation and results management." *Canadian Journal of Program Evaluation*, 24(2), 1-30.
- ▶ Newcomer, K. and C. Brass. 2016. "Forging a Strategic and Comprehensive Approach to Evaluation within Public and Nonprofit Organizations: Integrating Measurement and Analytics within Evaluation." *American Journal of Evaluation*, Vol. 37 (1), 80-99.
- ▶ Olejniczak, K., E. Raimondo, and T. Kupiec. 2016. "Evaluation units as knowledge brokers: Testing and calibrating an innovative framework." *Evaluation*, Volume 22 (2)., 168-189.
- ▶ Sunstein. C. and R. Hastie. 2015. *Wiser: Getting Beyond Groupthink to Make Groups Smarter*. Harvard Business Review Press.
- ▶ World Bank Group. *Mind, Society and Behavior*. 2015.