4.669...
Evaluation & Planning

Complexity:
A Different Way to Understand How Programs Work and What They Accomplish*

Or as originally titled:
Drawing on Complexity Science to Design and Execute Evaluation

Professional Development Session for the Annual Meeting of the
American Evaluation Association
October 10th 2023

Jonathan A. Morell, Ph.D.
4.669 Evaluation and Planning

Phone +1 734 646-8622
Website www.jamorell.com
Email jamorell@jamorell.com
Blog https://evaluationuncertainty.com/
YouTube https://www.youtube.com/channel/UCqRIJjhqmy3ngSB1AF9ZKLg

* An inquiry into what complexity tells us about models, metrics, data interpretation, and communicating with customers and stakeholders

Overview

Part 1	Much of The Whole Workshop in Two Slides	▪ Relevant complexity concepts ▪ Core focus on joint workings of emergence, sensitive dependence, and attractors
Part 2	Evaluation's Relationship to Complexity	▪ Where do we find complex systems? ▪ Criteria for understanding complexity ▪ Intellectual history of complexity
Part 3	What can we know and how can we know it?	▪ We are already well versed in most of the tools and methodologies we will need. ▪ In a complexity framework we need to re-orient out thinking about what we can and cannot know.
Part 4	Cases to Practice on	▪ For Use Only if Participants Don't Provide Their Own.
Part 5	Does it always make sense to invoke complexity? No.	▪ Methodological issues ▪ Need to express program logic ▪ Practical considerations
Part 6	Some Other Useful Complexity Concepts	If we have time and interest
Part 7	Further reading	Nothing systematic . A collection of articles I like.

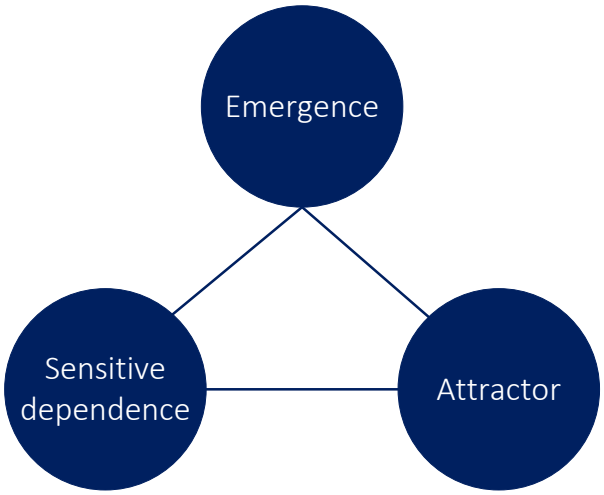
Part 1:
Much of The Workshop in Two Slides

1. Relevant complex behaviors
2. General themes

A few concepts will cover most evaluation scenarios we confront.

Scaling	Ratio of change in one factor to another
Attractors	A system's "normal" / equilibrium to which it naturally gravitates.
Stigmergy	Pattern from independent agents following rules based on what is immediately before them.
Emergence	The whole is qualitatively different than the sum of its parts.
Self-organization	Dynamics that drive attractor behavior.
Sensitivity to conditions	Small changes might change the trajectory of a system.
Evolutionary / ecological perspectives	Think of programs as organisms in an ecosystem.

For evaluation reasoning, these three are highly connected and particularly useful.



Concepts from complexity spread across themes.

	Complexity Theme		
	Pattern	Predictability	Change
Emergence			
Sensitivity to conditions			
Attractors			
Self-organization			
Stigmergy			
Scaling			
Evolutionary / ecological perspectives			

How much do you need to know?

- More is better
- At least more than minimal
- The statistics analogy holds

Part 2:
Intellectual Context – Evaluation's Relationship to Complexity

1. Domain of complexity
2. Criteria for acceptable work in a discipline
3. Historical overview

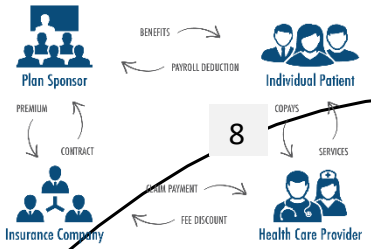
Where are complexity concepts operative?

1

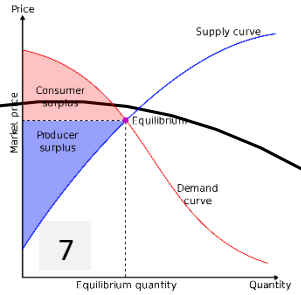
High

Complexity

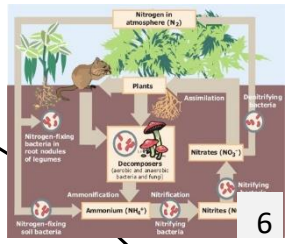
Low



8

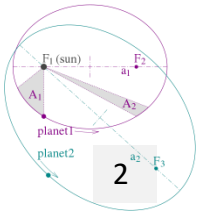


7

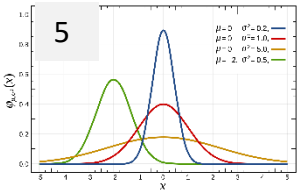


6

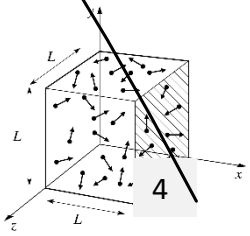
3



2



5



4

Low Random
High predictability

High Random
High predictability

1- Adapted from: David Krakauer What is Complexity?

2- https://en.wikipedia.org/wiki/Kepler%27s_laws_of_planetary_motion

3-http://barronberry.com/firm-news-and-events/celebrate-national-playground-safety-week-keep-your-kids-safe/

4- https://socratic.org/questions/what-causes-gas-pressure-in-terms-of-kinetic-theory

5- https://en.wikipedia.org/wiki/Normal_distribution

6- https://en.wikipedia.org/wiki/Ecosystem#/media/File:Nitrogen_Cycle.jpg

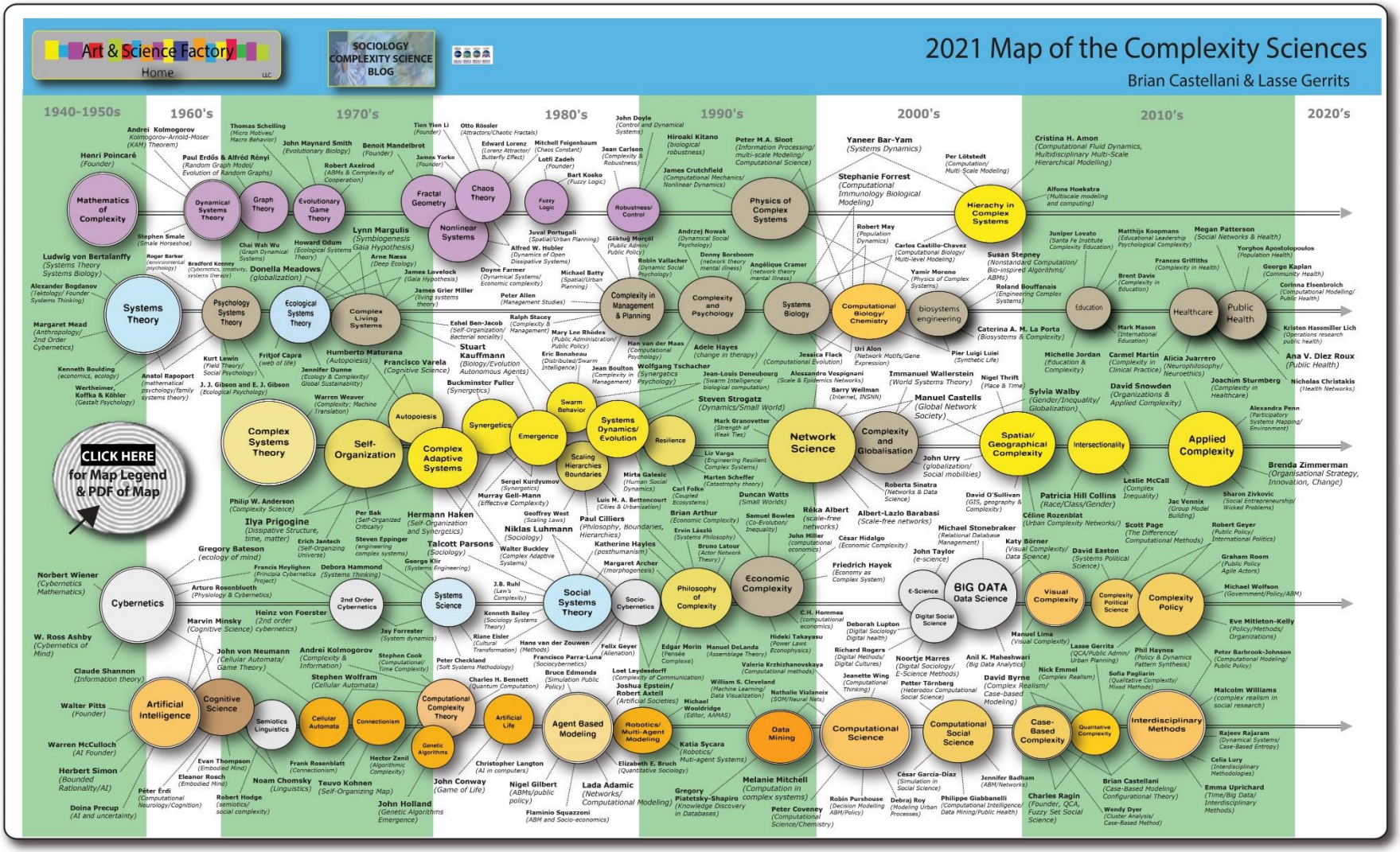
7- https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Economics

8- https://axenehp.com/accountability-hospital-health-system-pricing/

Every evaluation domain has its own criteria for accepted work. Complexity is no exception.

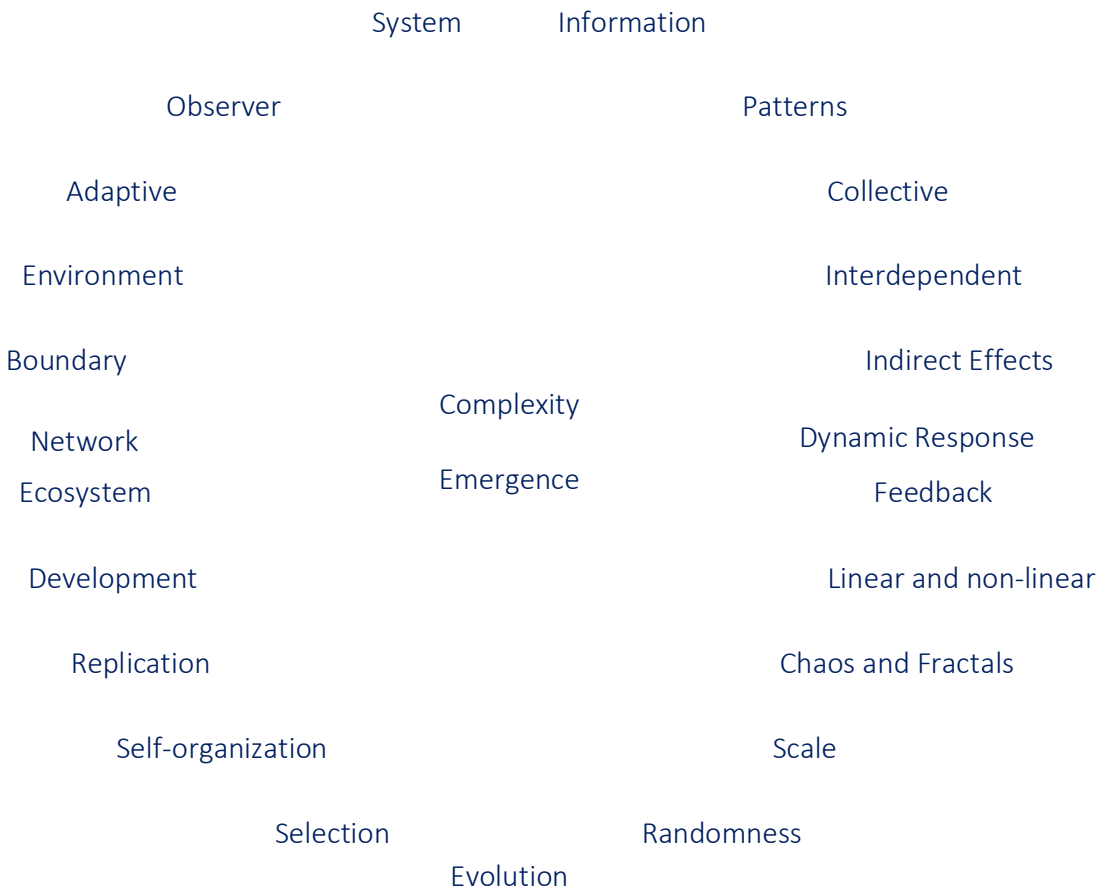
	RCT	CIPP	Realist	Process tracing	Empowerment	Developmental	Outcome harvesting	Complexity	Others
Models									
Data needs									
Methodology									
Data interpretation									
Acceptable answers									
Acceptable questions									
Convincing arguments									
Skills on research team									
Choice of research design									
Others									

Complexity: Historical View*



- https://www.art-sciencefactory.com/complexity-map_feb09.html

Complexity: Concept Map View*



Important related concepts

- Control
- Particle
- Influence
- Hierarchy
- Power Law
- Reductionist
- Newton's Laws
- Thermodynamics
- Distributed control
- Separation of scales
- Statistical mechanisms
- Nonequilibrium dynamics
- Language / formal languages

* Adapted from: Concept Map, New England Complex Systems Institute 2011 Yaneer Bar-Yam

So, is there a "Complexity Science"?

Considering the social organization of a body of knowledge people call Complexity Science and the intellectual connections across work that is done, it's fair to say "yes". [1]

But as with so much of history, there is a contrary view [2]

> The current vogue is for the third of the three C's: complexity. The buzzwords here are emergence and self-organization , as complexity theory seeks to understand how order and stability arise from the interactions of many components according to a few simple rules ... . But very often what passes today for 'complexity science' is really something much older, dressed in fashionable apparel. The main themes in complexity theory have been studied for more than a hundred years by physicists who evolved a tool kit of concepts and techniques to which complexity studies have added barely a handful of new items.

1 Phelan, Steven E. 2001 What Is Complexity Science, Really? *Emergence* v3 https://doi.org/10.1207/S15327000EM0301_08

2 Ball, P. (2003) *Critical Mass: How One Thing Leads to Another*

Part 3:
Complexity Knowledge. What can we know and how can we know it?

1. Familiar methodologies can be applied to the behavior of complex systems.
2. In a complexity framework we need to re-orient out thinking about what we can and cannot know.

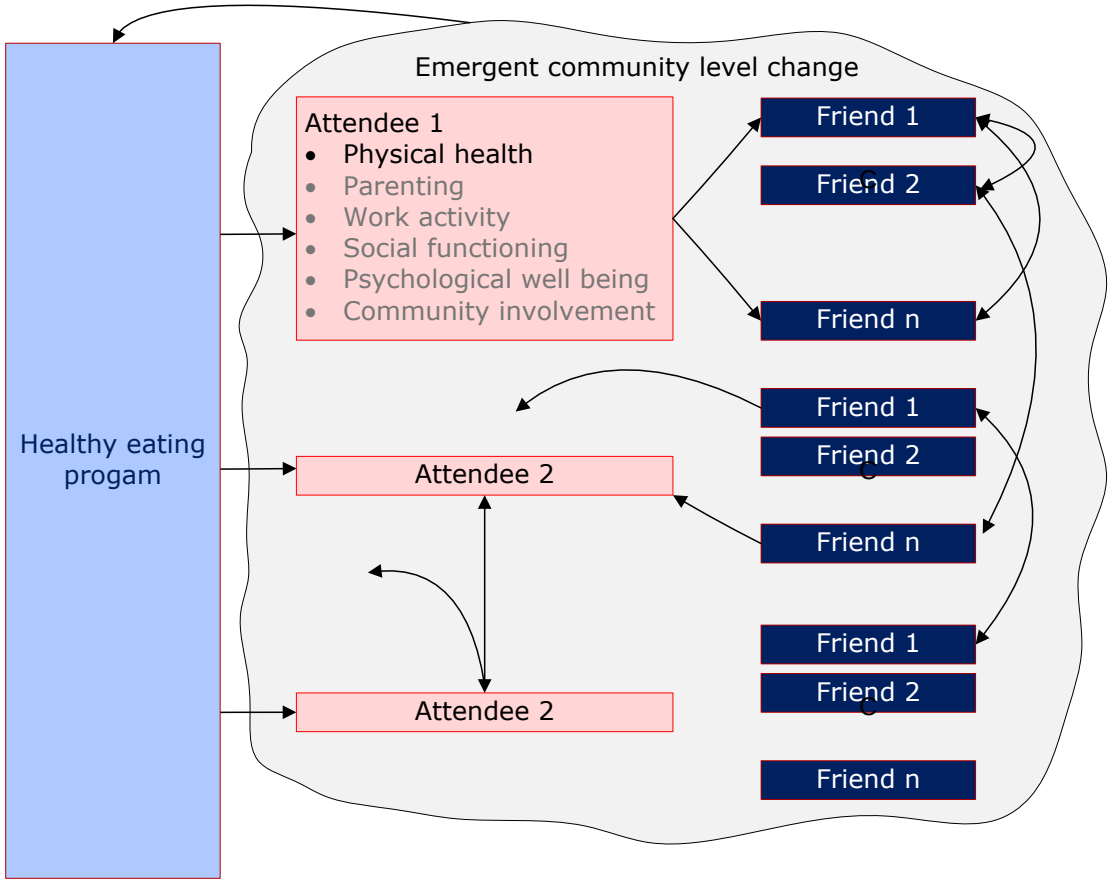
Complex behavior, plebian methodology

Data

- Health markers
- Influence on friends
- Programming change
- Community level changes
- Immediate impact on participants
- Secondary impacts on participants
- What was implemented -- # of components, order

Methodologies

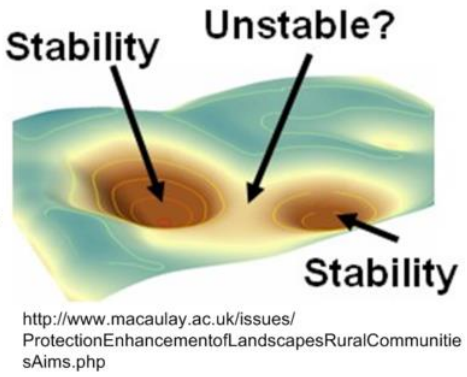
- Process tracing
- Records review
- Change over time
- Program activity monitoring
- Content analysis of social media
- Comparison with other communities
- Interviews (program staff, attendees and others)
- Observation (program staff, attendees and others)



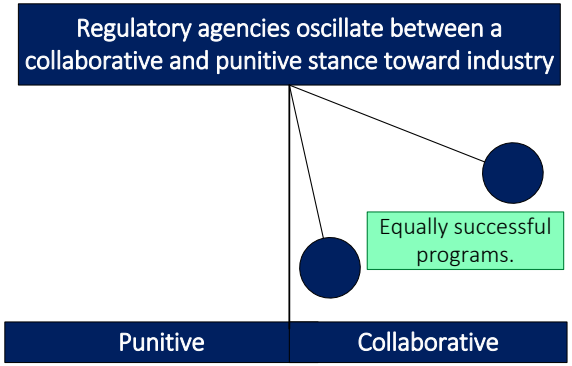
Exceptions

- Agent and system dynamic modeling
- Formal network structure
- AI applications

Attractors: A system's "normal" / equilibrium to which it naturally gravitates.



http://www.macaulay.ac.uk/issues/ProtectionEnhancementofLandscapesRuralCommunitiesAims.php



Attractor shape helps us visualize resistance to change and sustainability.

Self organization can be seen as the dynamics that shape attractors.

Can explain both resistance to change and sustainability.

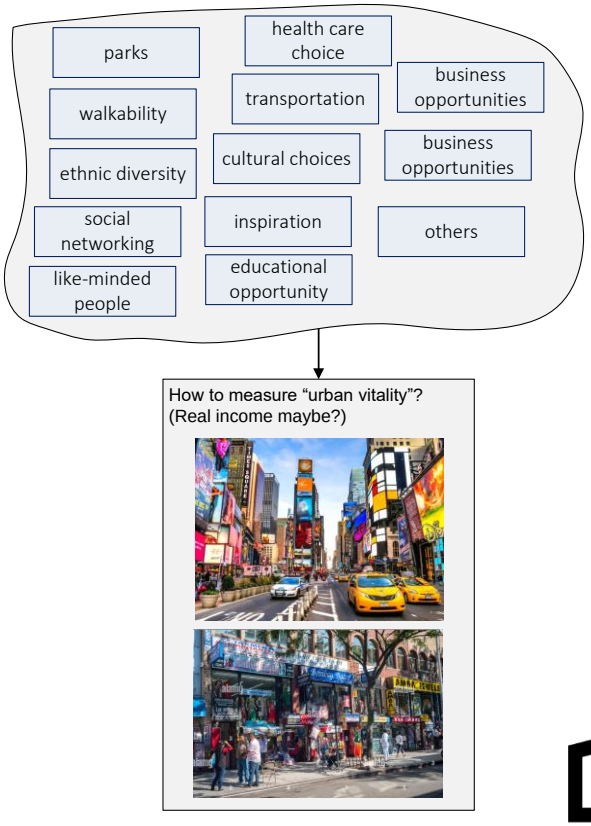
Other slides on attractors: 37, 38

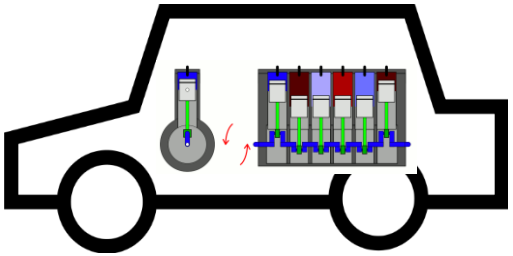
1-http://www.livescience.com/images/i/000/087/324/original/river-basins.jpg
2-https://www.mnn.com/earth-matters/animals/quiz/what-do-you-call-these-groups-of-animals
3-http://barronberry.com/firm-news-and-events/celebrate-national-playground-safety-week-keep-your-kids-safe/
4-https://i.stack.imgur.com/cj3xs.jpg
5- https://soundout.org/2015/04/15/parts-of-the-education-system/
6- https://digitalhealth.folio3.com/blog/wp-content/uploads/2023/02/benefits-of-integrated-healthcare-systems-3-1280x720.webp

Emergence: The whole is qualitatively different than the sum of its parts

- Program theory
- Measurement

- Useful to evaluate parts
- Parts maintain their identity, but
- Need a qualitatively different measure for the whole.





Ways to think about emergence

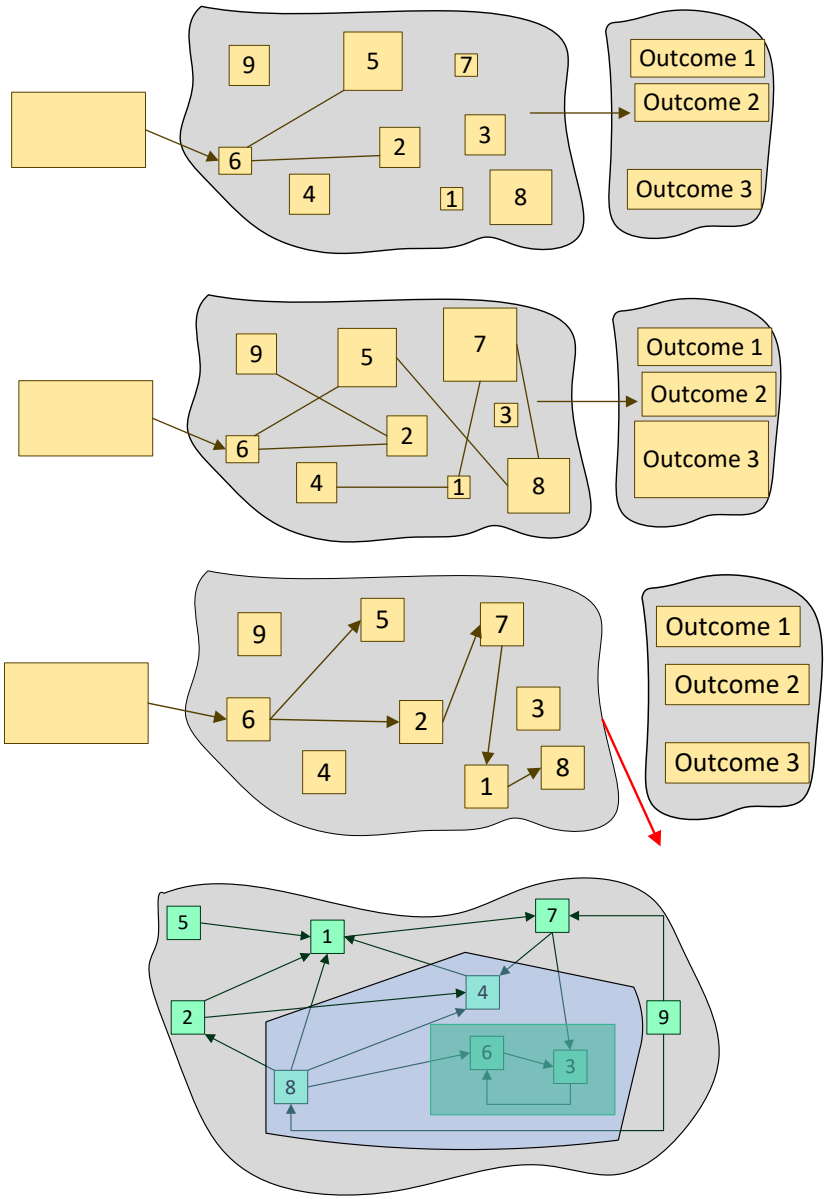
- Whole is greater than the sum of its parts
- "Emerged" phenomenon cannot be predicted.
- "Emerged" phenomenon cannot be decomposed

Other slides on emergence: 34, 35

Image sources:
Times Square
Jackson Heights Queens

Image sources:
cylinder, car, dog

Sensitivity to conditions: Small changes *might* change the trajectory of a system



With sensitive dependence, what can we know?

History
- What path led to success or failure
- Whether multiple paths worked
- Similarities in successful paths

Connections
- Change in network relationships over time

Magnitude of change
- Program elements
- Outcomes

Causal relationships
- Among close elements of a model

A complexity-based understanding of realist evaluation

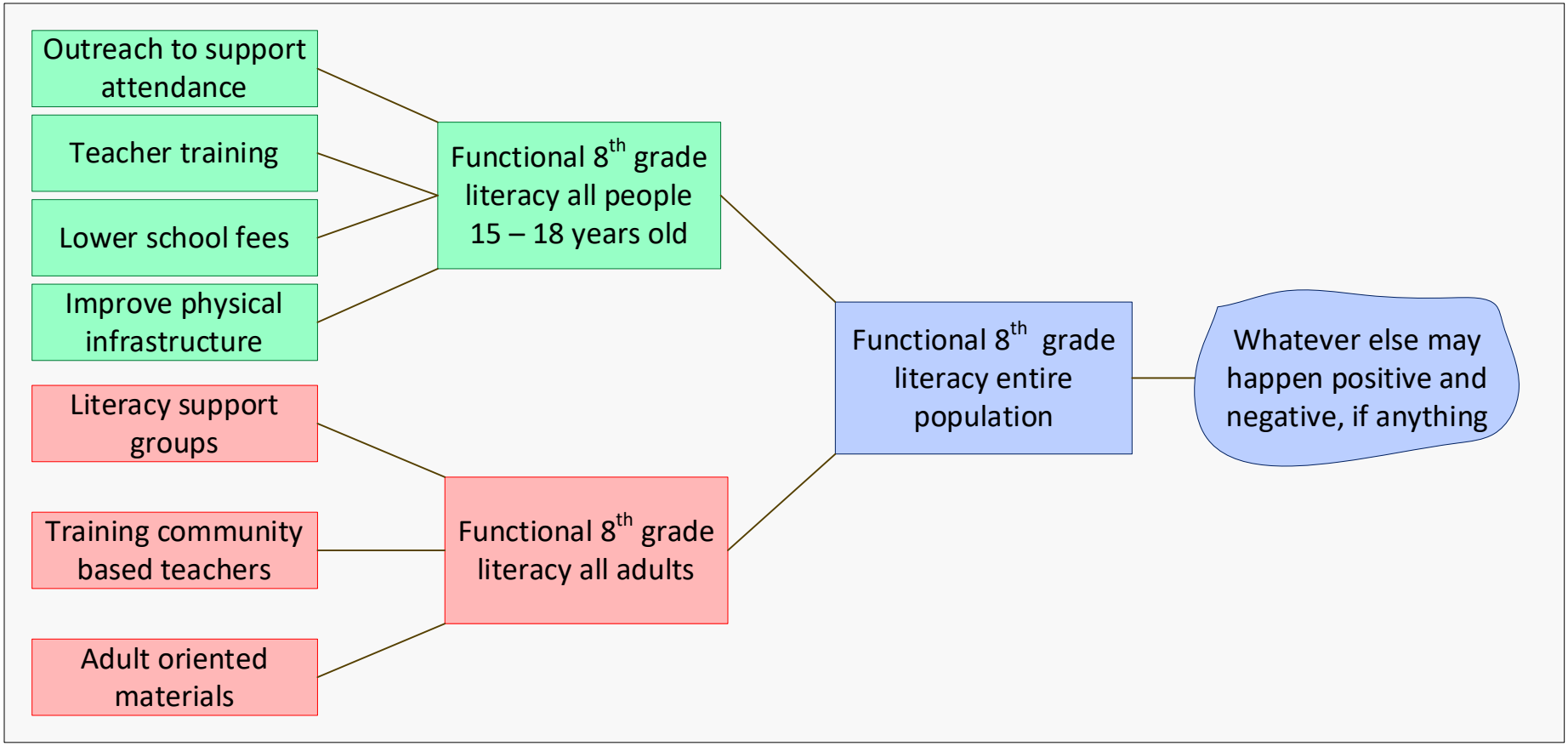
Part 4:
Cases to Practice on
For Use *Only* if Participants Don't Provide Their Own.

1. Universal Literacy in a Geo-political Region
2. Portfolio of Community Support for Parents and Children

Case #1 Universal Literacy in a Geo-political Region

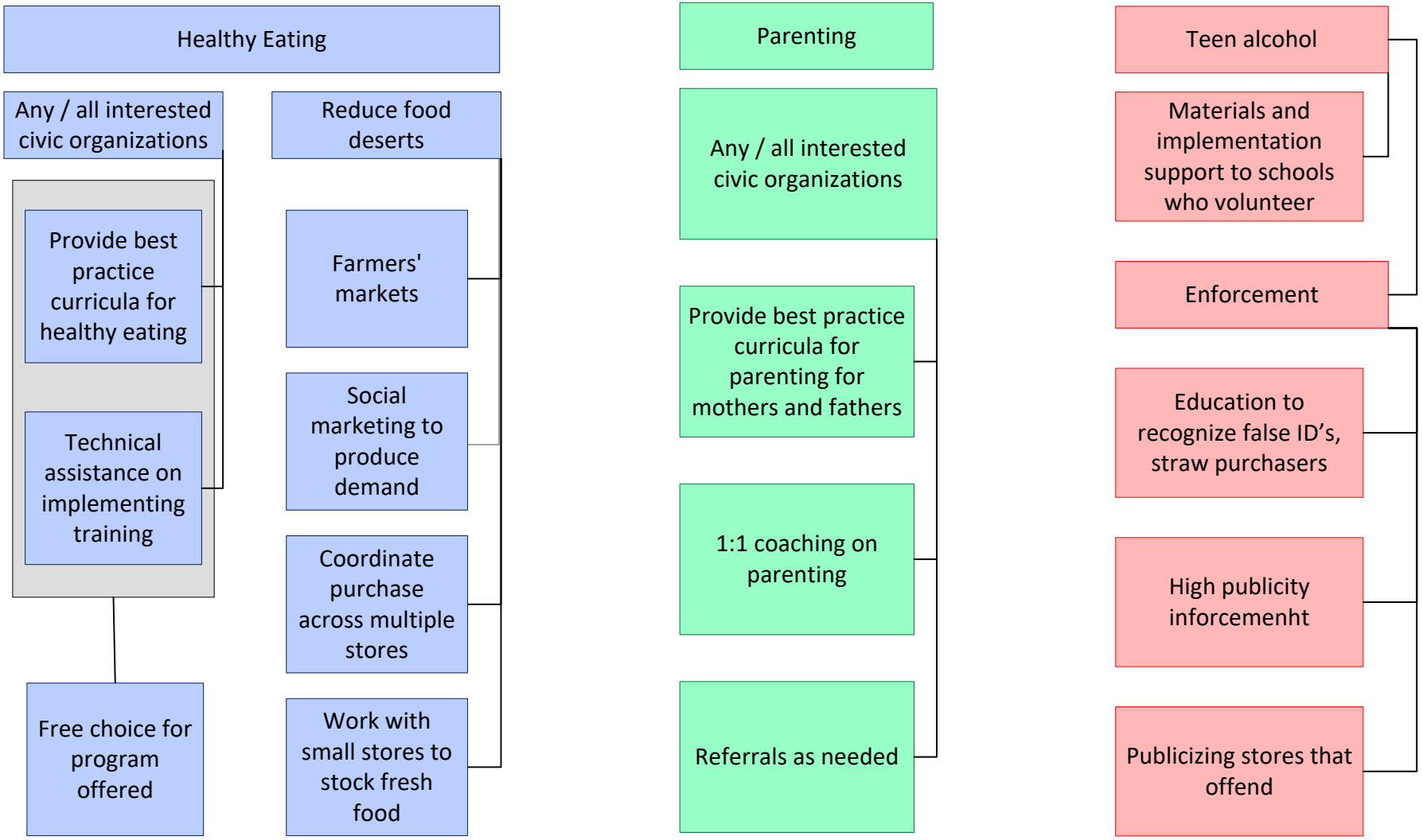
Theory:

- Literacy itself is desirable
- Widespread literacy will result in much change, most (if not all) of it desirable



Case #2 Portfolio of Community Support for Parents and Children

- Any single program can have unanticipated positive and negative consequences.
- Programs can combine to have widespread consequences



Part 5:
Does it always make sense to invoke complexity?

1. Would it present methodological difficulties?
2. Is it needed to adequately express program logic?
3. Are there practical considerations such as time, money, expertise required, etc.

Does invoking complexity always make sense?

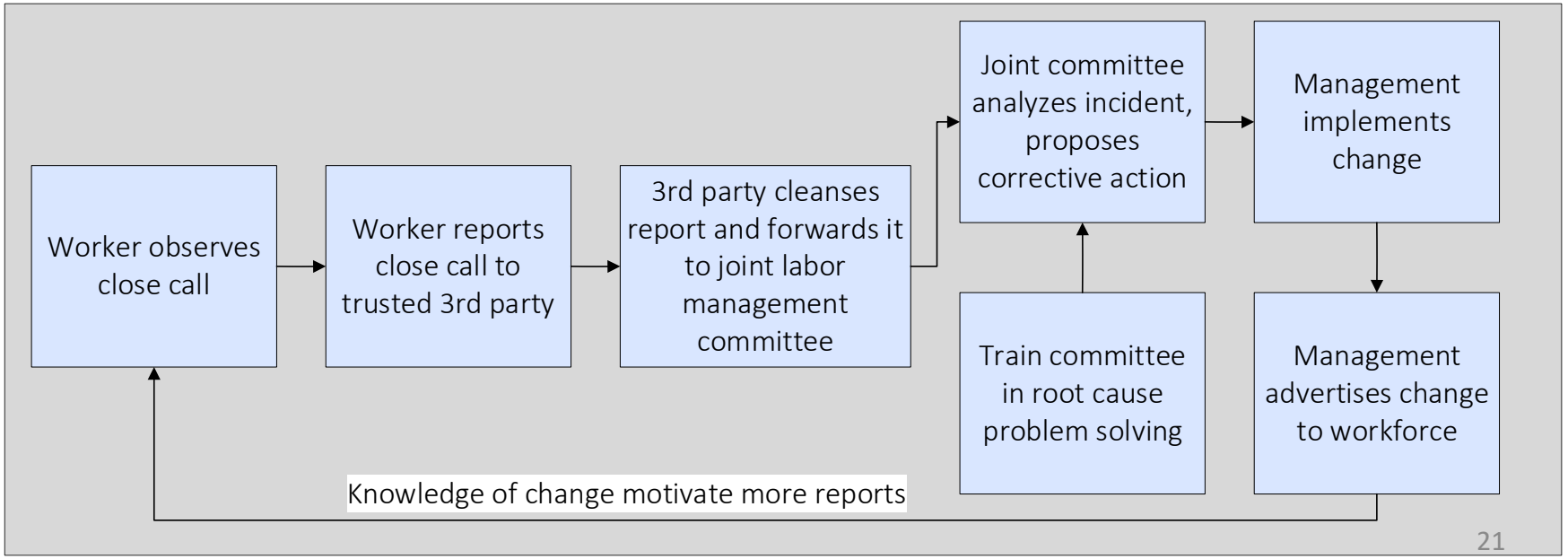
No. If it does not affect program logic.

Context

- Accidents are common in railroads.
- For legal and cultural reasons railroading is a blame-based culture.
- Labor / management relations are fraught.
- Frank discussions of close calls do not take place

Program theory

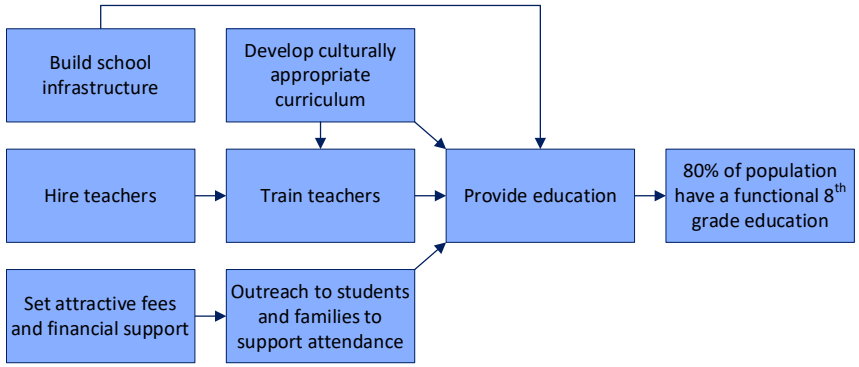
- Knowing the root cause of “close calls” increases safety.
- Filtering reports through a trusted third party to scrub identifying detail will lead to knowledge of what happened and why.
- A diverse labor /management group can devise workable corrective actions



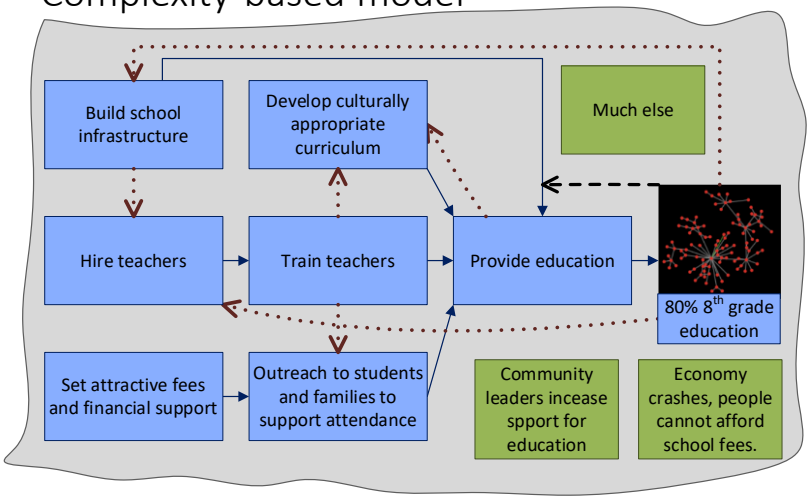
Does invoking complexity always make sense?

No, for technical and practical reasons.

Traditional if → then model



Complexity-based model



	Most Likely Answer			
	Traditional		Complex	
	Yes	No	Yes	No
Environment considered?				
Cost easily accommodated?				
Growth patterns recognized?				
Data requirements manageable?				
Easily understandable to stakeholders?				
Important elements and connections missing?				
High % of findings provide actionable information?				

Does invoking complexity always make sense? No.

- I would have done that anyway.
- It would be too resource intensive.
- It won't add value to the evaluation.
- My customer won't like it or benefit from it.
- It won't lead to different recommendations.
- I don't have access to the necessary expertise.
- It won't lead to different understanding of how the program works.

Part 6:
Some Other Useful Complexity Concepts
if we have time and interest

1. Scaling
2. Stigmergy
3. Predictability
4. Growth dynamics
5. Conditionals in models
6. More on attractors and emergence
7. When can we do without complexity?
8. Ideas from Ecology and Evolutionary Biology
9. Sneaking in complexity to a traditional evaluation design.
10. Progression from models with no complexity to models with complexity
11. I would have done it anyway. What is the value added of invoking complexity?

Evolutionary / Ecological Behavior

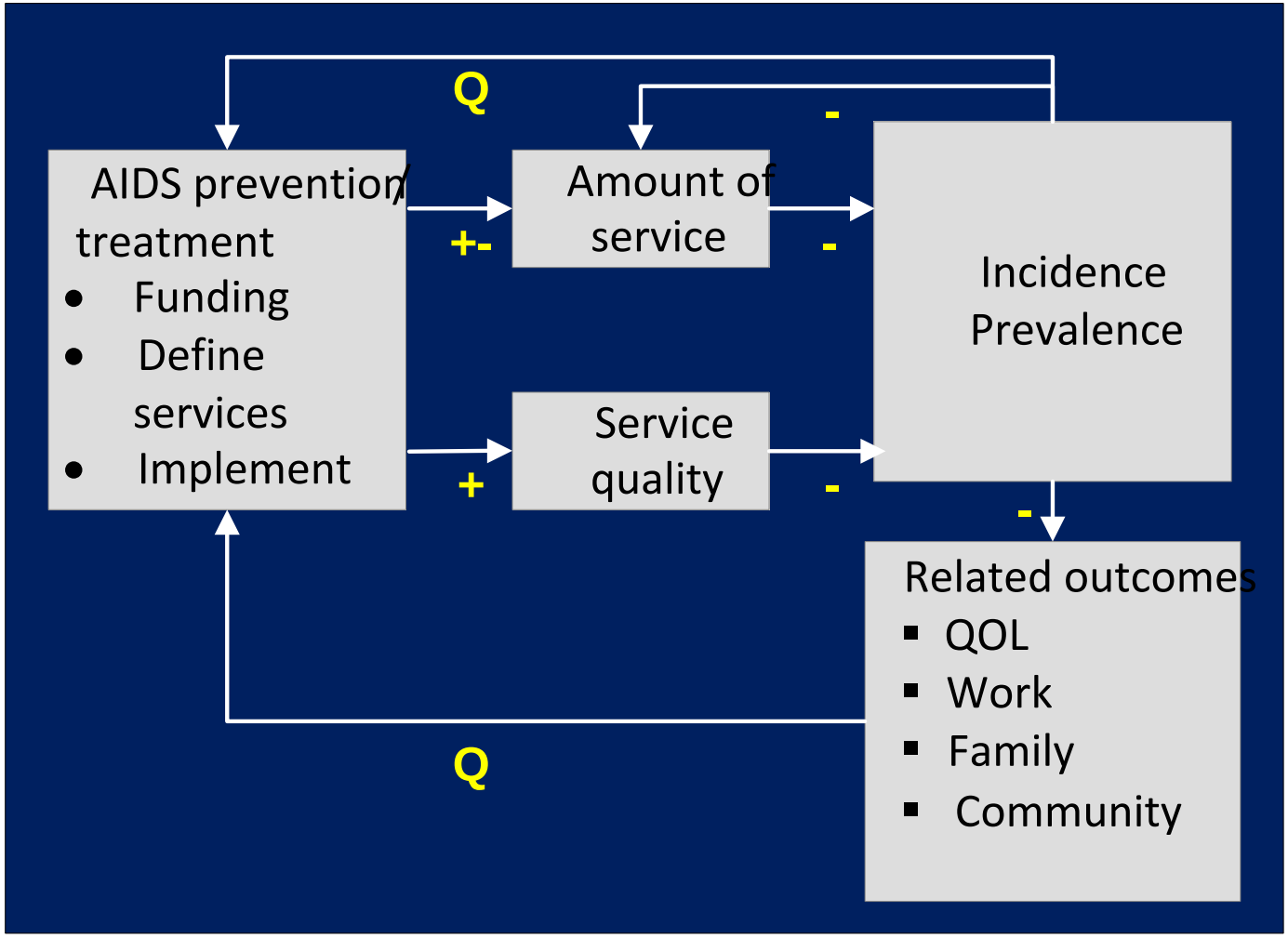
Thinking in ecological terms conjures many related questions that may not derive from other ways of approaching an evaluation

Adaptation	▪ How do programs in an “ecosystem” react when other programs are introduced
Population size	▪ How many of a kind of program are there?
Rates of change	▪ At what rate does the number of programs change? ▪ How quickly do programs change what they do? ▪ How do programs change how they are structured?
Timing of changes	▪ How much time elapses between a policy change and program change?
Fitness landscape	▪ How much of a change is required to make a dramatic difference for the viability of a program?
Diversity of programs	▪ How many kinds of programs are there?

Evolutionary / Ecological Behavior

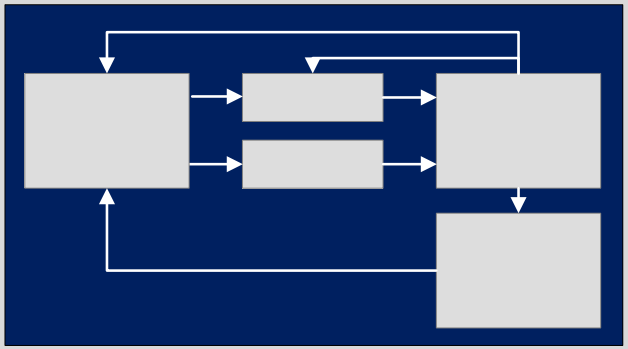
Shifting From

- Program Model to
- Model of the Program as a Charge in an Ecosystem



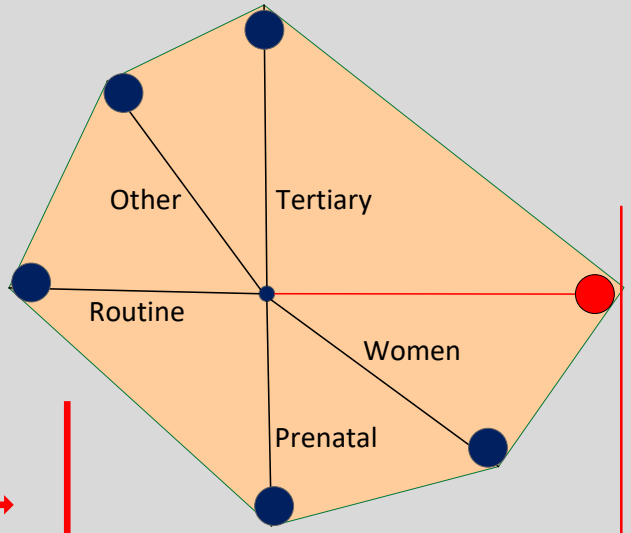
Evolutionary / Ecological Behavior

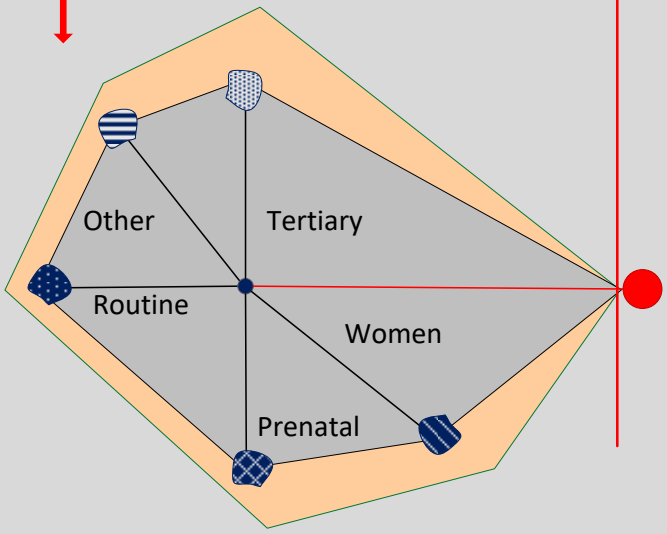
Outcome Maximization Program Theory



	Public Health	AIDS
After: Public health system		X
▪ Career choices		X
▪ Policy synergies		X
▪ Political capital		X
▪ Intellectual effort		X
▪ Skills people develop		X
▪ Informal relationships		X
▪ Supporting structures		X

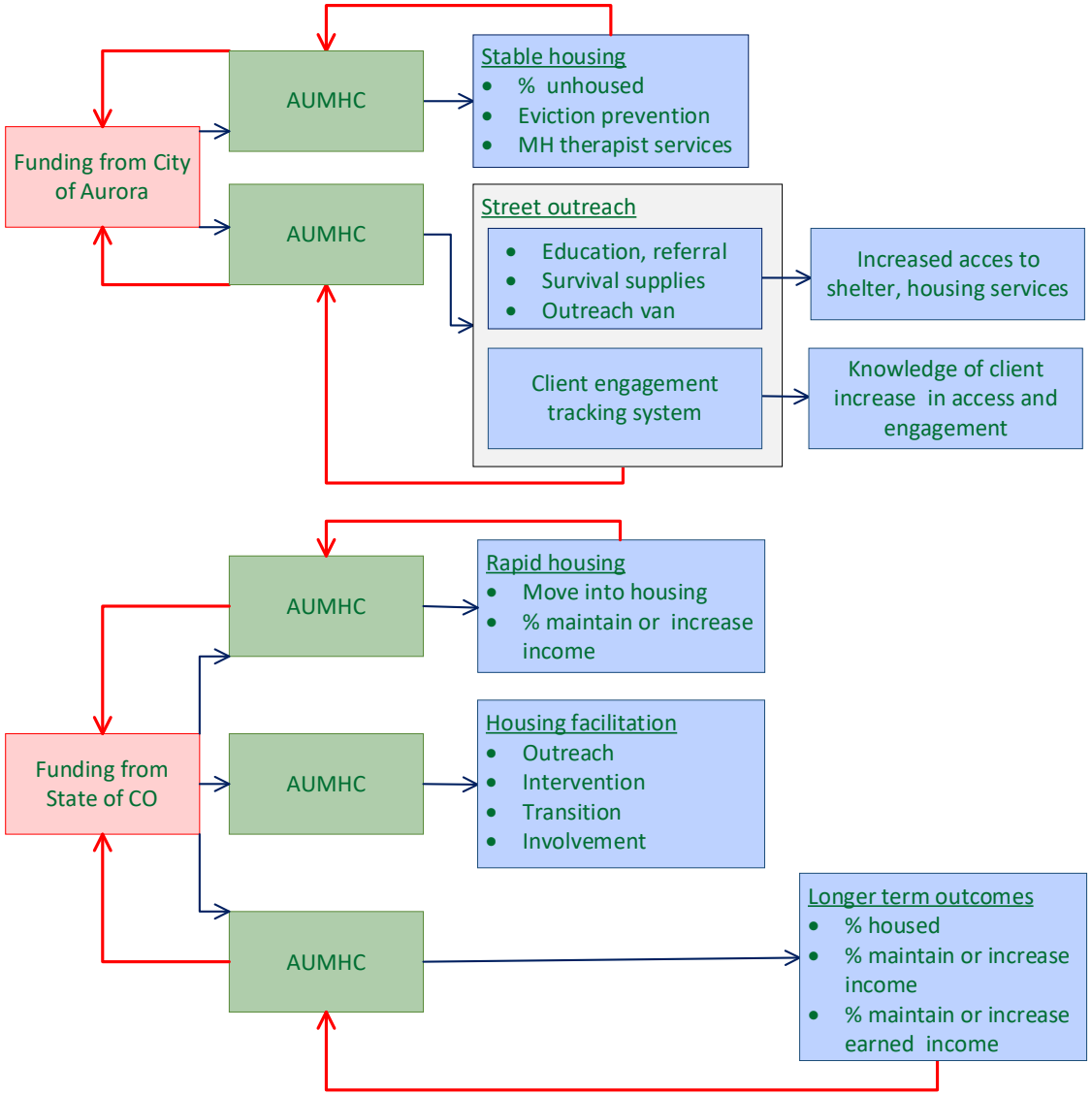
Evolutionary / Adaptive Program Theory





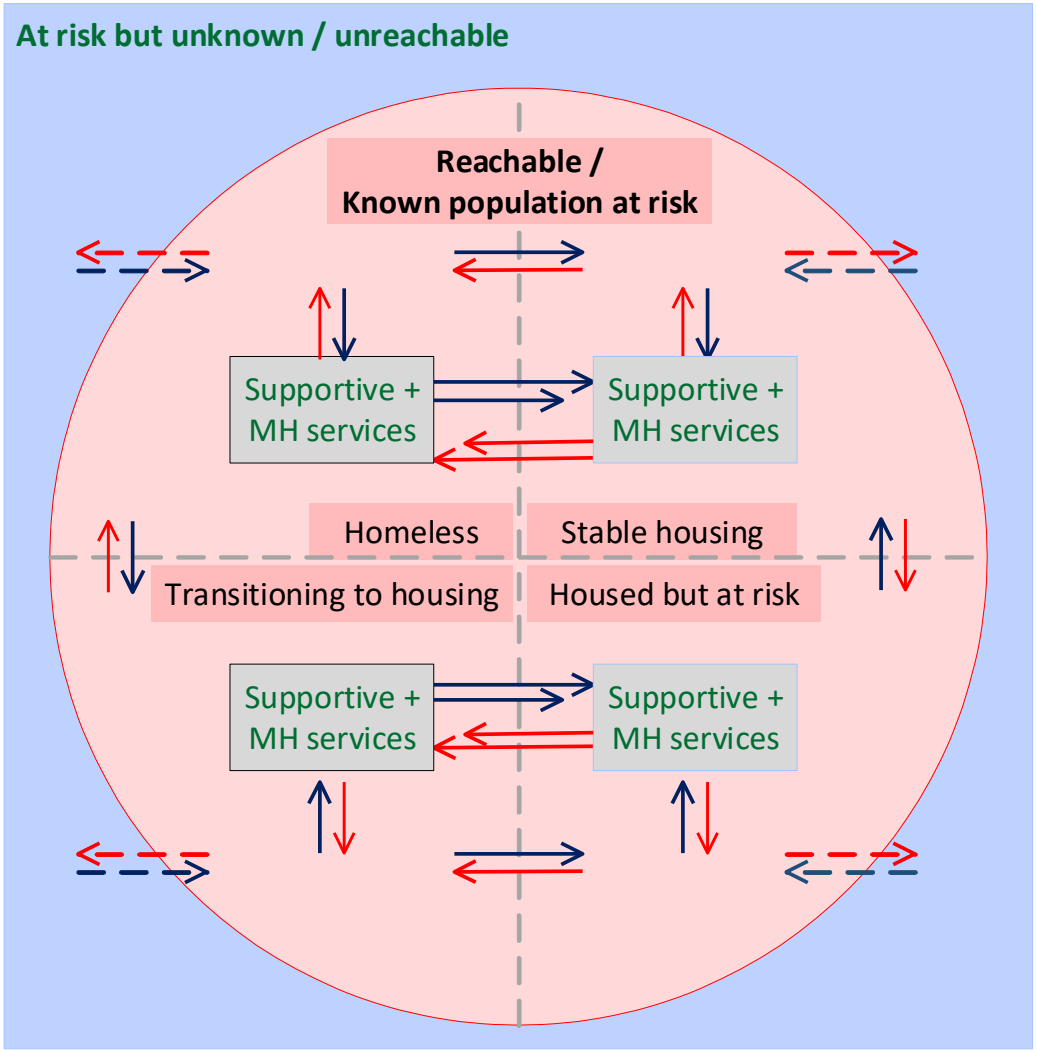
Evolutionary / Ecological Behavior – Population Flow

Traditional support program for the unhoused*

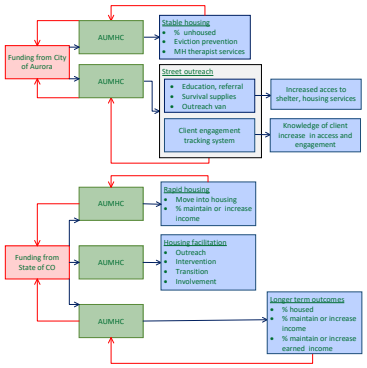


* Pathways to Home: Providing Wraparound Case Management and Clinical Services for Homeless Individuals in Aurora, CO, For information contact Antonio Olmos AntonioOlmos@aumhc.org

Evolutionary / Ecological Behavior – Population Flow



Example of analysis suggested by model			
From	To	Rate of Boundary Crossing Scenario 1	Rate of Boundary Crossing Scenario 2
unreachable	reachable	up	up
receives services	leaves services	down	up

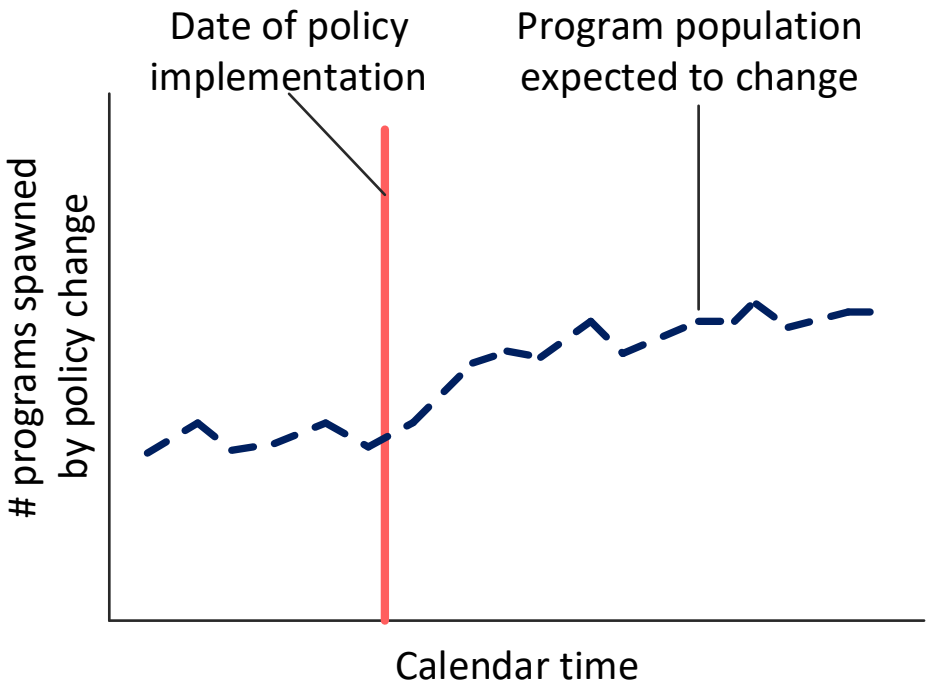


Evolutionary / Ecological Behavior

A common model for evaluating the impact of a policy change.

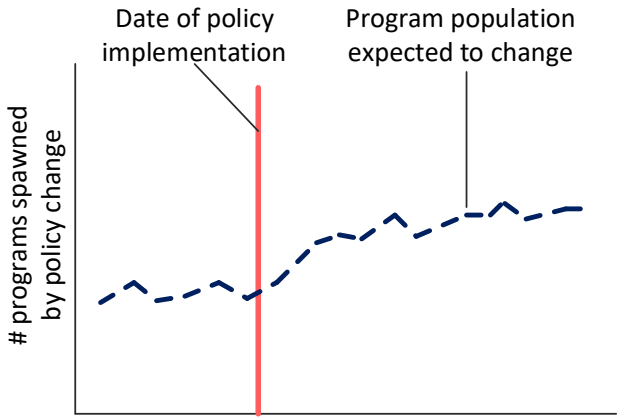
Shifting From a Policy → Program Model to a Policy → Ecosystem Model

- Program of primary interest is identified.
- Program growth is tracked over time.

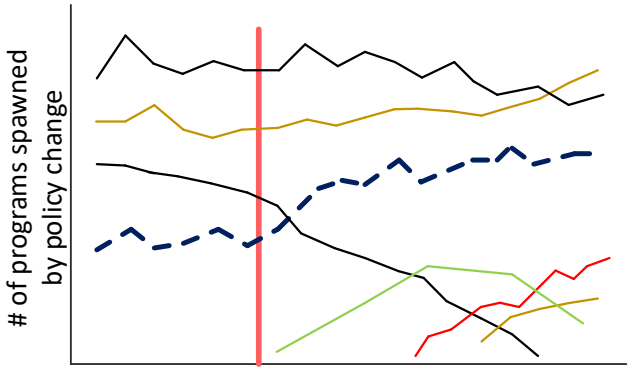


Evolutionary / Ecological Behavior – Policy evaluation from an ecosystem perspective

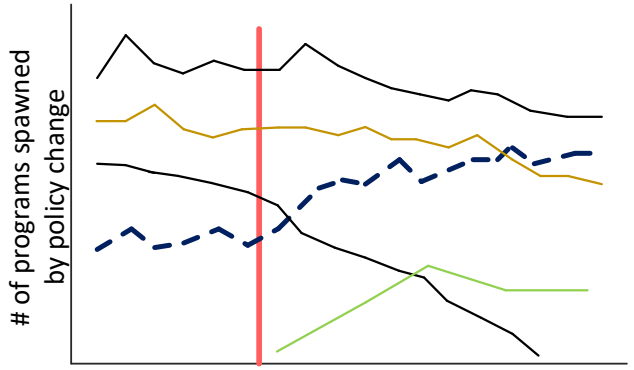
- Primary measurement remains
- Measurements of other programs are added
- A different view of the policy appears



- Three new types of programs appear
- 2/3 of the new programs thrive
- 1 program appears soon after policy change
- 2 programs lag policy chabge
- Only 1 other program declines



- Three programs decline
- Only 1 other new program appears



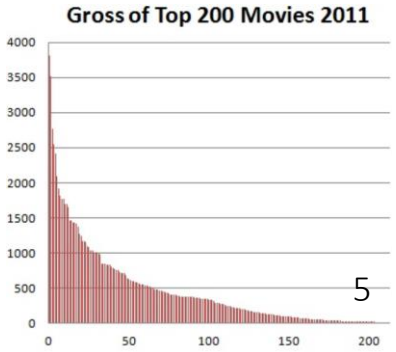
If we can recognize complex behaviors, we can predict with high accuracy.

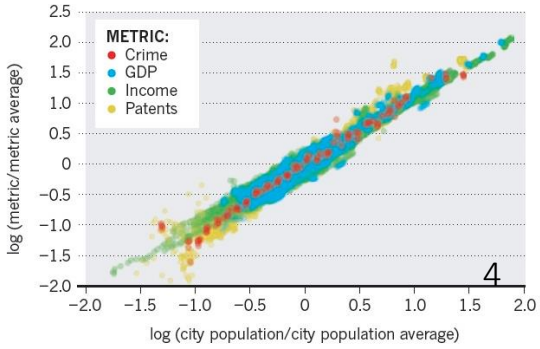
Preferential attachment yields fractal patterns

Snowflake

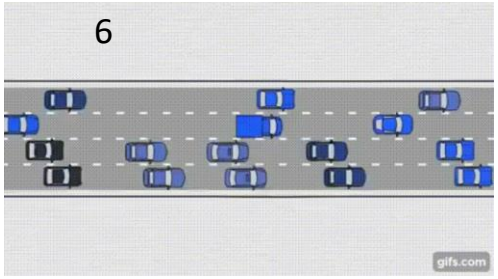
Internet

Scaling factors are common over a wide range of variables





Emergence can often be expected



2 snowflake https://www.wbur.org/news/2022/02/25/snowflake-man-photographs-wilson-bentley

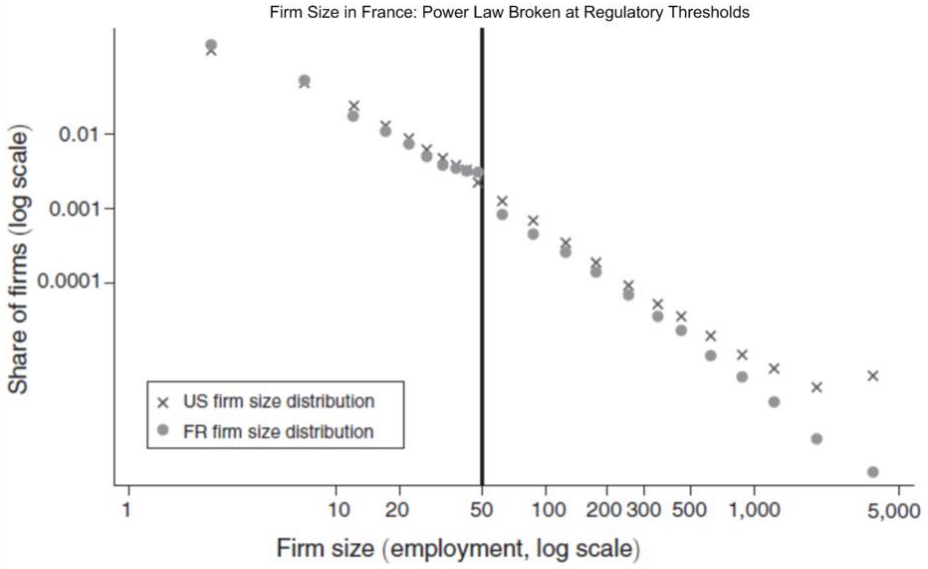
3 internet https://www.kaggle.com/general/177015

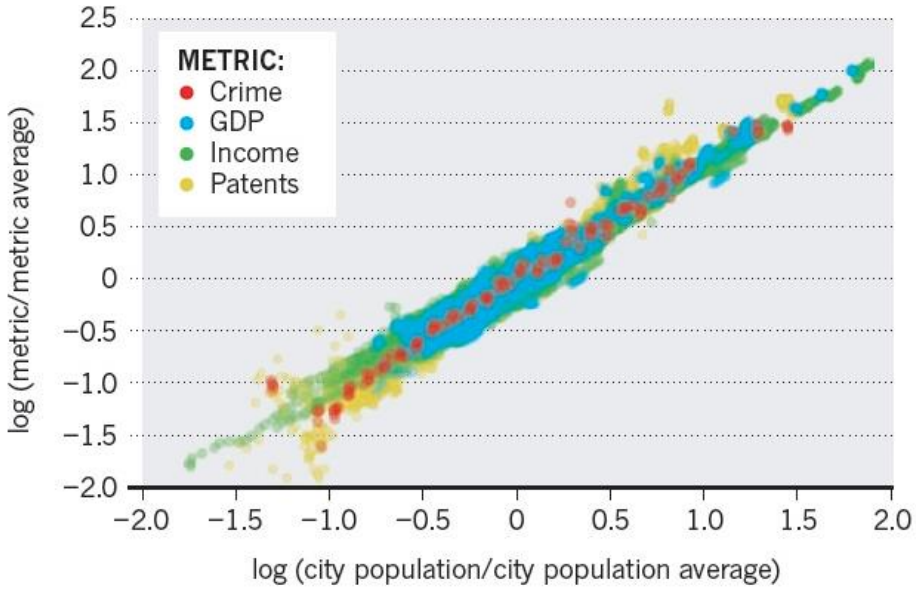
4 http://www.interculturalurbanism.com/?p=2879

5 https://praxtime.com/2012/12/16/digital-economics-the-hollow-middle/

6 https://motorlease.com/article/traffic-jams-explained/

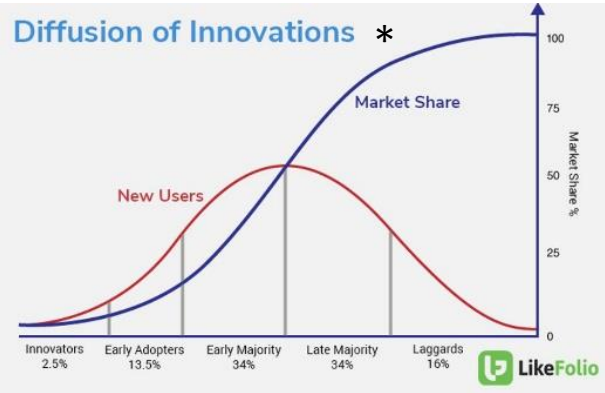
Scaling patterns are common to much of what we evaluate.





Some other examples of emergence

Same adoption curve, different impact metrics



Connectivity in Internet Deseret

- New business activity
- Participation in civic groups

Qualitative transition. Need metrics to assess community impact

Flat screen TV

- # owned
- % penetration

No new metrics needed to assess impact

* Forbes: How To Spot Companies Accelerating Through The Adoption Curve

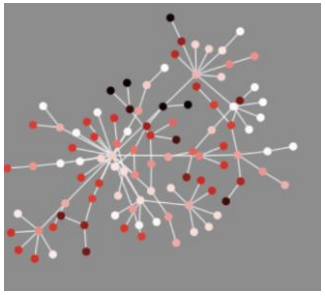
** https://en.wikipedia.org/wiki/Emergence

*** Wykis - Own work, Public Domain,https://commons.wikimedia.org/w/index.php?curid=1618169

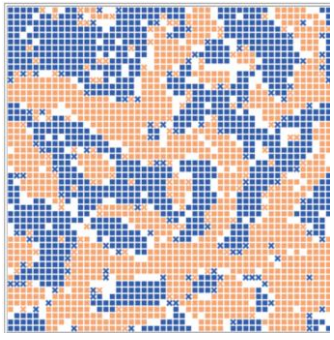
NetLogo models library

https://ccl.northwestern.edu/netlogo/

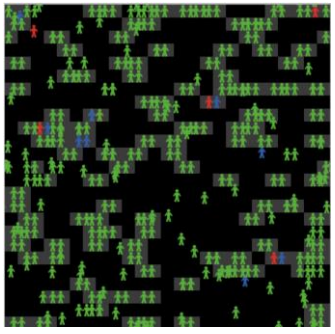
Language change



Segregation

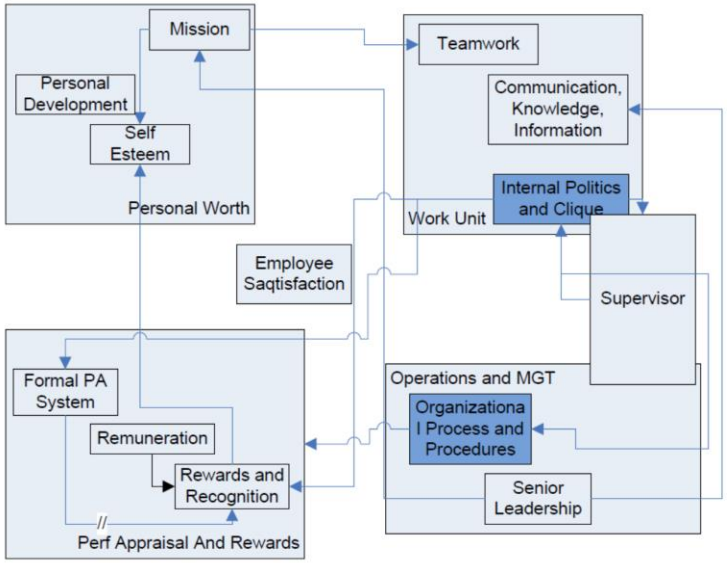


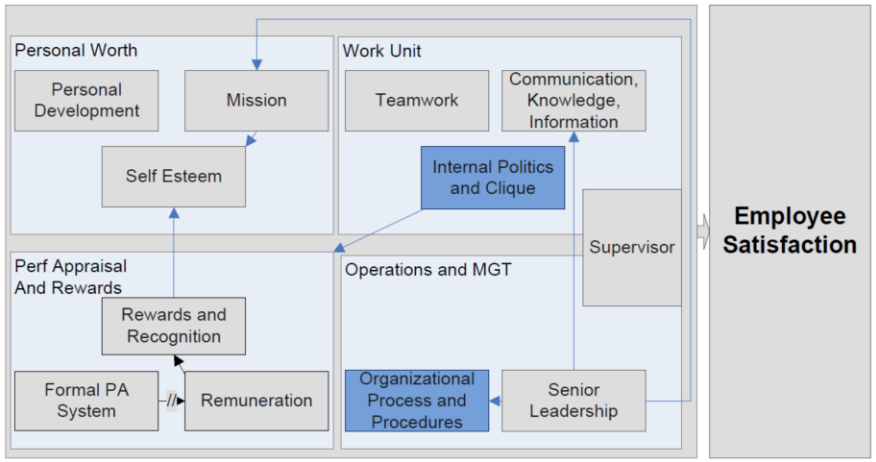
HIV



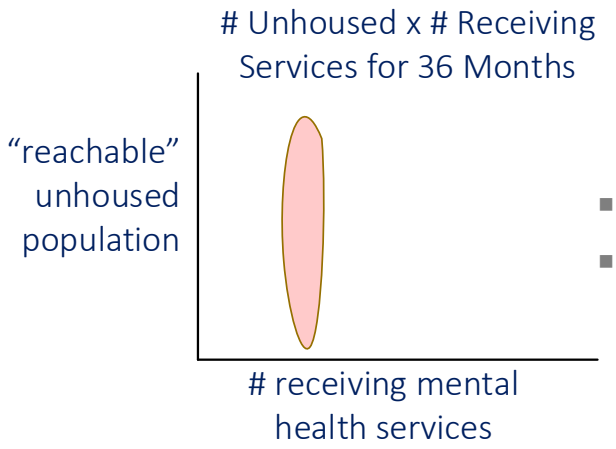
Where is complexity operating?

- Brainstorm model
- Clean model
- Perform sniff test – now that we understand what we said, do we really mean it?
- Decide if/where complexity behavior might be worth considering.

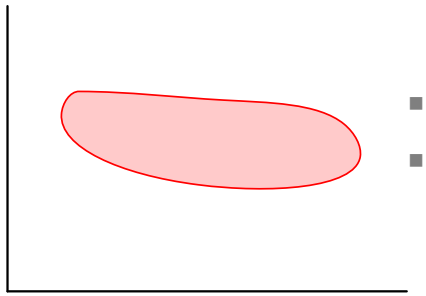




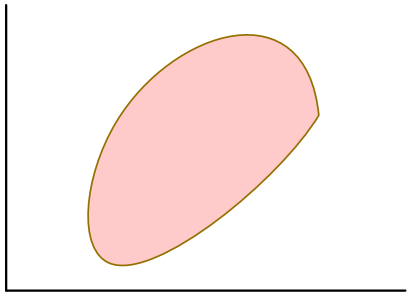
Social Attractor



- # of unhoused changes
- narrow range of people served



- amount of service varies
- despite a constant need



- Need and service scale pretty well.

What are the implications for each scenario?

- Service planning
- Resource planning
- Understanding need?

Why is there monthly variation?

- weather
- # unhoused
- office space
- # counselors
- monthly budget
- # student interns
- police behavior
- # outreach workers
- willingness to seek help
- much more

Why invoke attractors? I would have done that anyway.

Compare across programs / program types	▪ What other types of programs, pursuing other outcomes, have the same attractor shape?
Is the change stable?	▪ How sustainable will change be? ▪ How much resistance to change will there be?
What attractor space describes the program's outcome?	▪ What is the range of values the outcomes can take? ▪ What patterns characterize those outcomes?
Why not think in terms of attractors?	▪ I would have done that anyway. No value added. ▪ The topics it raises don't matter.
Self organization is a related concept.	▪ A process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system. ▪ Moreover, the rules specifying interactions among the system's components are executed using local information, without reference to the global pattern.*

* Glossary at the Santa Fe Institute's Complexity Explorer

Stigmergy

a mechanism of indirect coordination, through the environment, between agents or actions. The principle is that the trace left in the environment by an individual action stimulates the performance of a succeeding action by the same or different agent.*

Why should evaluators care about stigmergy?

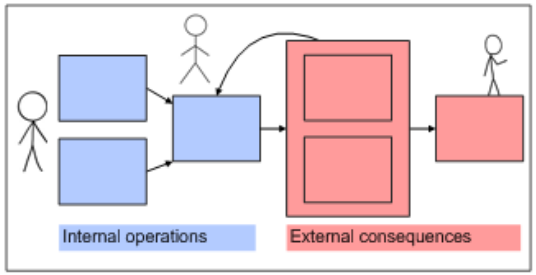
- Explanation of change patterns and program effects
- Humility with respect to predicting or designing change patterns
- Framework for inquiry when evaluating planned and unplanned change
- Input for designing change efforts

Some examples of stigmergic processes

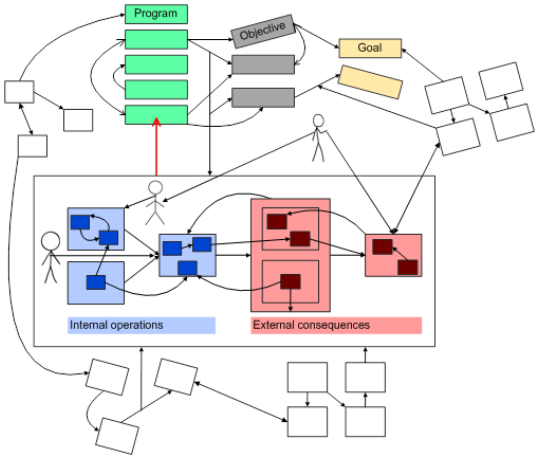
- Style in art
- Urban vitality
- Online communities
- Intellectual trends in science

Help planners escape the trap of "stovepipe" conflict with "stigmergent coordination".

Model we almost always use



Model nature uses



- Different time horizons
- Stovepipes are efficient
- Goals that may conflict
- Different organizational cultures
- Few personal working relationships
- Unknown, unknowable interactions
- Cost of coordination people, $, time
- Different contingencies to prepare for
- Different schedules for decision making
- Different stakeholders with different priorities
- Programs serve purposes besides stated goals
- People have their favorite sources of information

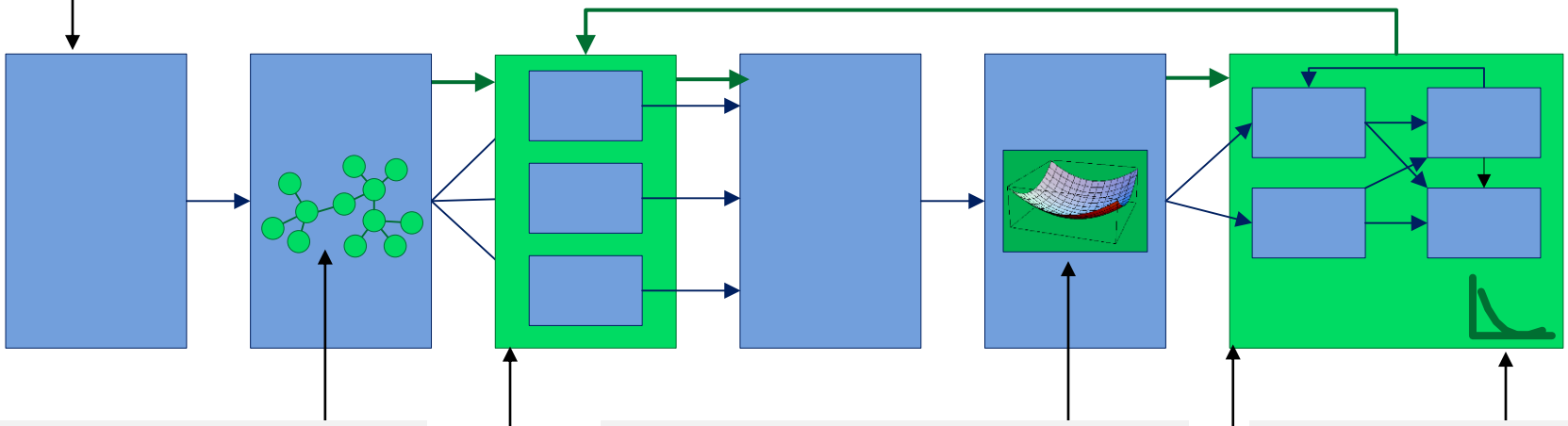
Stigmergic coordination*

- A few (of many) common metrics
- No compulsion, maintain independent decision making

* Reed, John H., and Gretchen Jordan. 2007. Using systems theory and logic models to define integrated outcomes and performance measures in multi-program settings. *Research Evaluation* 16 (3): 169–181.

Sometimes one invoke complexity even if funders insist on a traditional evaluation *

Traditional model desired by customer.



Observe growth patterns

Data interpretation wrt attractors

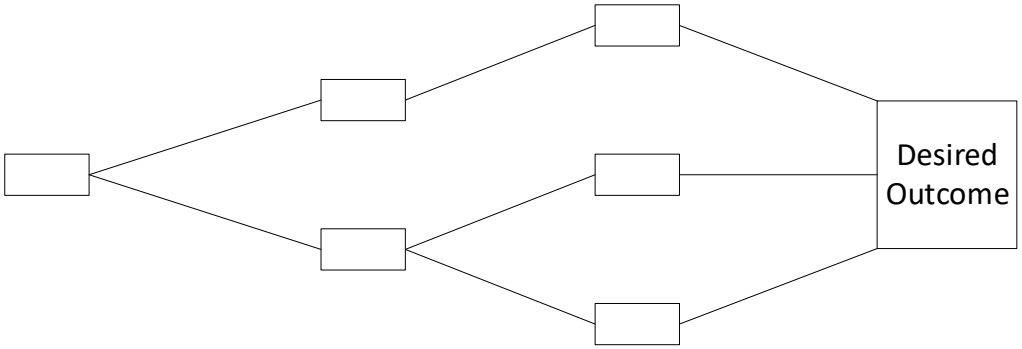
Effect size distribution

Measurement on an emergent level

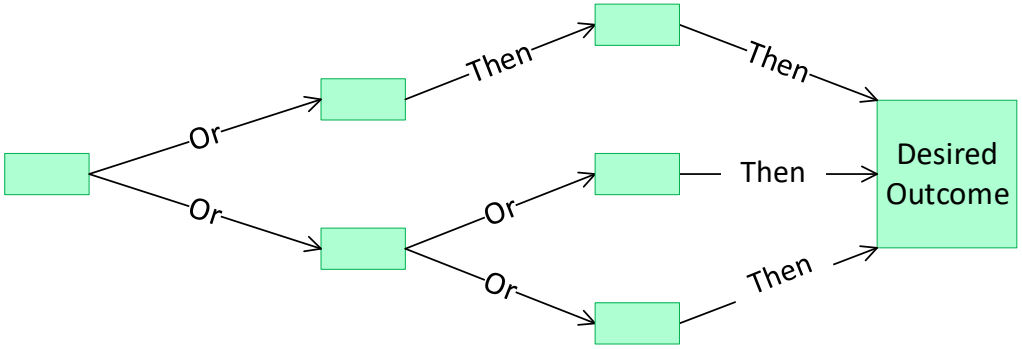
* It depends on whether additional data collection or a parallel design is needed.

Asking stakeholders to specify conditionals facilitates expectation setting

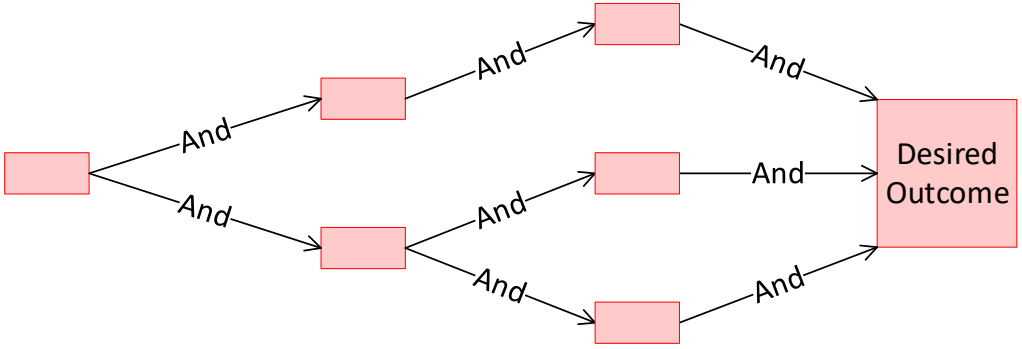
Common model form has minimal predictive or explanatory ability.



This program has a reasonable chance of working



This program is doomed to failure

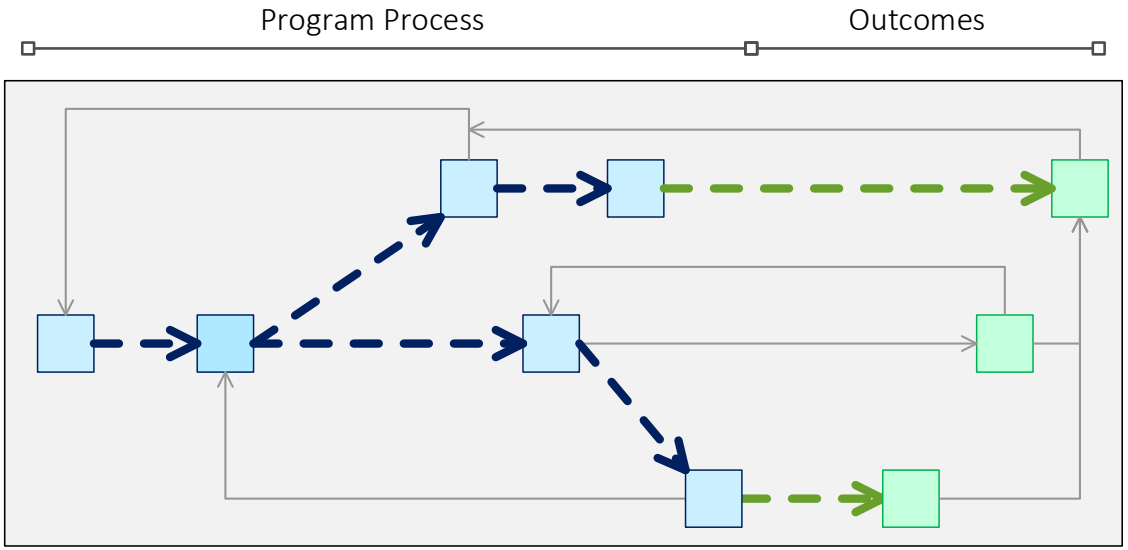


Asking stakeholders to state confidence in linkages facilitates expectation setting

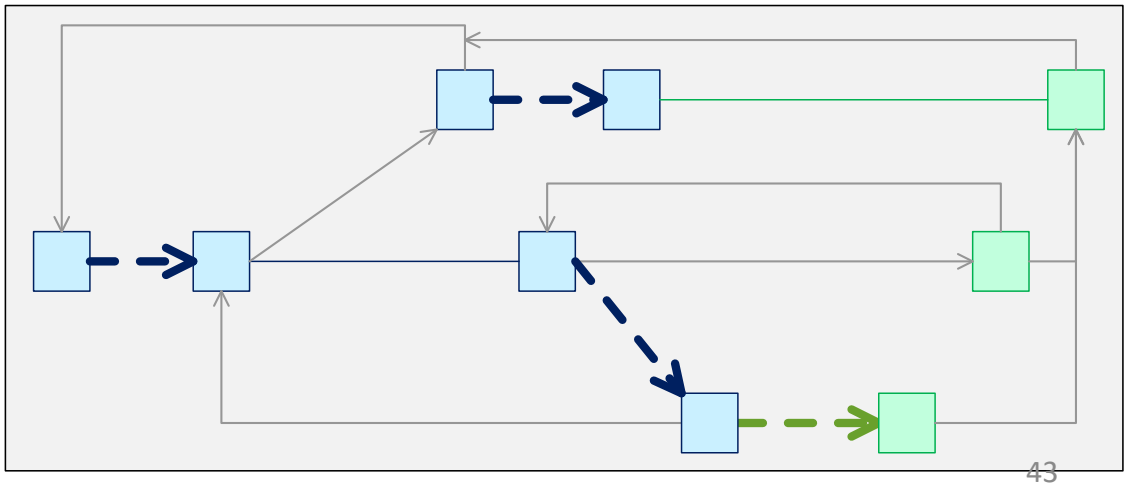
High confidence that linkage will succeed

Not high confidence that linkage will succeed

Scenario 1: Stakeholders
- believe that program can be implemented successfully



Scenario 2: Stakeholders are
- uncertain about successful implementation
- confident that only 1/3 of the program – outcome links will work



Part 7:
Some Reading

There is nothing systematic about this list. It’s just a collection of of articles I like.

Bamberger, M., Vaessen, L., & Raimondo, E. (Eds.). (2015). *Dealing With Complexity in Development Evaluation: A Practical Approach*. Thousad Oaks CA.: Sage.

Barbrook-Johnson, P., Proctor, A., Giorgi, S., & Phillipson, J. (2020). How do policy evaluators understand complexity? *Evaluation, 26*(3), 315-332. doi:10.1177/1356389020930053

Bicket, M., Christie, I., Gilbert, N., Hills, D., Penn, A., & Helen Wilkinson. (2020). *Supplementary Guide: Handling Complexity in Policy Evaluation* (H. Treasury Ed.): CECAN – the Centre for Evaluating Complexity Across the Nexus.

Comfort, L. K. (1994). Self-Organization in Complex Systems. *Journal of Public Administration Research and Theory: J-PART, 4*(3), 393-410. Retrieved from http://www.jstor.org/stable/1181895

Cox, J. (2019). How does the commissioning process inhibit the uptake of complexity-appropriate evaluation? https://www.cecan.ac.uk/news/how-does-the-commissioning-process-inhibit-the-uptake-of-complexity-appropriate-evaluation/

Gates, E. F., Walton, M., Pablo, V., & McNall, M. (Eds.). (2021). *Special Issue: Systems and Complexity-Informed Evaluation: Insights from Practice* (Vol. 170): Wiley.

Hertz, T., Brattander, E., & Rose, L. (2021). Complexity-Aware Monitoring and Evaluation. *Journal of Multidisciplinary Evaluation, 17*(41), 35-50. Retrieved from https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/679

Ling, T. (2012). Evaluating Complex and Unfolding Events in Real Time. *Evaluation*. doi:DOI: 10.1177/1356389011429629

Morell, J. A. (2021). A Complexity-based Metatheory of Action for Transformation to a Green Energy Future. In R. D. Van den Berg, C. Maro, & M.-H. Adrien (Eds.), *Transformational Evaluation for the Global Crises of our Times* . (pp. 345 - 370). Exeter, UK: IDEAS.

Morell, J. A. (2021). Using Evolutionary Biology and Ecology to Think About Evaluation. https://www.youtube.com/channel/UCF0s_RM9Vmbu-j8T5pemjRw?view_as=subscriber

Morell, J. A. (2023). A Complexity-based Plan for Evaluating Transformation. *Journal of Multidisciplinary Evaluation*. https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/867

Morell, J. A., Hilscher, R., Magura, S., & Ford, J. (2010). Integrating Evaluation and Agent-Based Modeling: Rationale and an Example for Adopting Evidence-Based Practices. *Journal of Multidisciplinary Evaluation, 6*(14), 35 -- 37.

Morell, J. A., Parunak, H. V. D., Hwalek, M., & Knestis, K. (Producer). (2016a). Part 1 Combining Agent Based Modeling and Traditional Evaluation Part 1, Part 2, Part 3.

Picciotto, R. (2021). Towared a Complexity Framework for Transformative Evaluation. *Journal.of Multidisciplinary Evaluation, 16*(35), 54 - 76.

Preiser, R., Biggs, R., De Vos, A., & Folke, C. (2018). Social-ecological systems as complex adaptive systems

organizing principles for advancing research methods and approaches. *Ecology and Society, 23*(4). Retrieved from https://www.jstor.org/stable/26796889

Renger, R. (2022). *System Evaluation Theory: A Blueprint for Practitioners Evaluating Complex Interventions Operating and Functioning as Systems*. Charlotte NC: Information Age Publishing.

Zazeueta, A. (2017). *Principles for the Development of Integrated Transformational Projects in Climate Change and Chemicals & Waste*. Retrieved from Washington DC: http://www.stapgef.org/sites/default/files/publications/Final_Principles%20for%20%20Development%20of%20Integrated%20CCM%20and%20ChW%20Projects_October%202017.pdf

Zazueta, A. E., Le, T. T., & Bahramalian, N. (2021). Development Trajectories and Complex Systems–Informed Theories of Change. *American Journal of Evaluation, 42*(1), 110-129. doi:10.1177/1098214020947782