4.669…
Evaluation & Planning

Applying Complexity to Make Practical Decisions About Program Theory, Methodology, Data, and Stakeholder Engagement*

Professional Development Workshop
American Evaluation Association
November 11th – 16th, 2019
Minneapolis MN

Jonathan A. Morell, Ph.D.
4.669 Evaluation and Planning
Editor-in-Chief Emeritus, Evaluation and Program Planning

Phone 734 646-8622
Website www.jamorell.com
Email jamorell@jamorell.com
Blog https://evaluationuncertainty.com/
YouTube https://www.youtube.com/channel/UCqRIJjhqmy3ngSB1AF9ZKLg
* URL http://www.jamorell.com/documents/Complexity_workshop_AEA_2019_published.pdf

Respect data.
Trust judgement.

Part 1 Introductions

Part 2 Overview of using complexity in evaluation

Part 3 Group exercise: Collective design of an evaluation

Part 4 Redesign of evaluation with complexity taken into account

Part 5 Topics we may consider depending on interest and what we have not already covered

Part 6 Discussion of “parking lot” issues

Part 7 Resources and further reading (nothing systematic, just books and articles that I like)

Much related information sits in Evaluation Uncertainty: Surprises in Programs and Their Evaluations

Part 1: Introductions

Introduce yourselves to people in your neighborhood

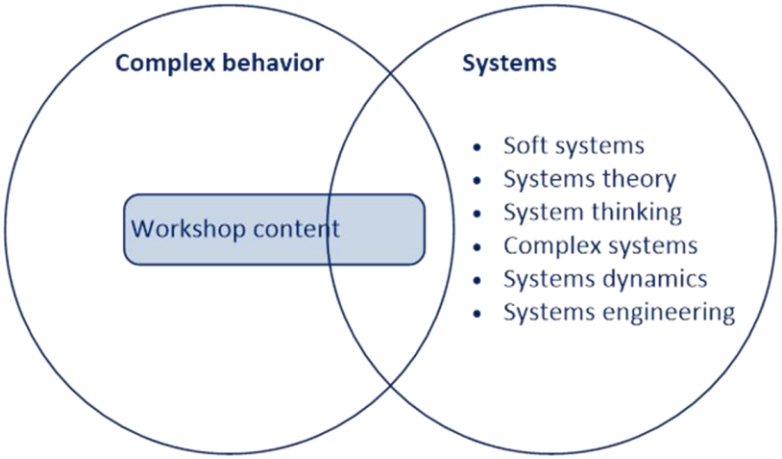
Nominate one person to make introductions to all of us

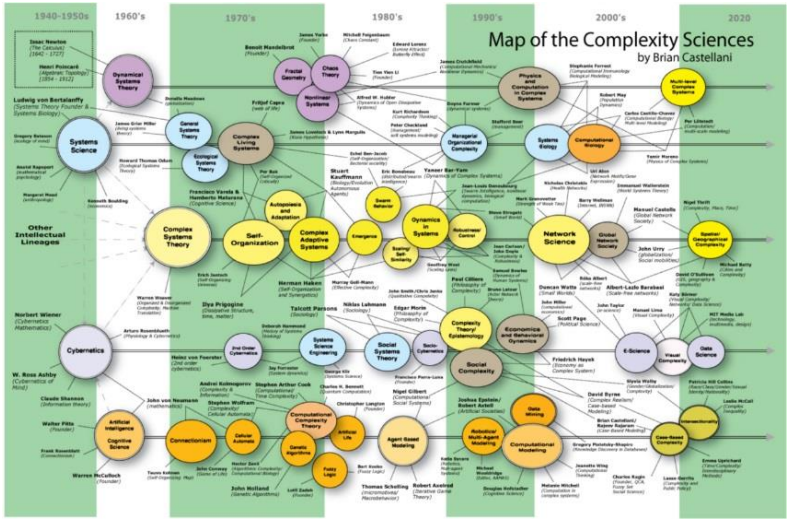
Presenter	Other group member
Name	Name
Some work experience	Some work experience
Expectations for workshop (optional)	Expectations for workshop (optional)

Part 2: Overview of Using Complexity in Evaluation

My goals for this workshop

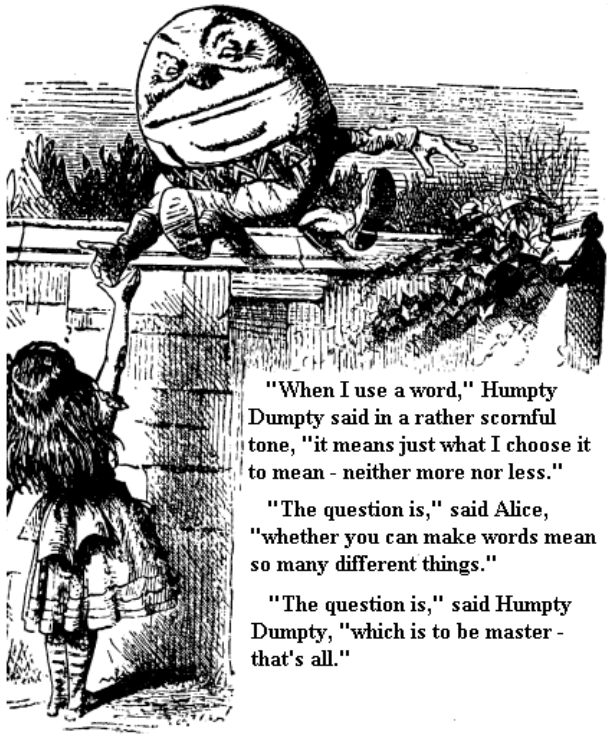
- Make you comfortable with complexity
- Apply some aspects of complexity to evaluation





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

You may learn a lot about complexity, but you won't get a definition.



"When I use a word," Humpty Dumpty said in a rather scornful tone, "it means just what I choose it to mean - neither more nor less."

"The question is," said Alice, "whether you can make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master - that's all."

Through the Looking Glass, by Lewis Carroll
http://www.authorama.com/through-the-looking-glass-6.html

A Framework for Thinking about Complexity and Evaluation

	Theme in Complexity Science			
Some complex behavior that may be useful in evaluation	Pattern	Prediction	Explanation	How change happens
Emergence				
Realistic timeframes				
Discontinuous change				
Evolutionary behavior				
Scaling and power laws				
Self-organized criticality				
Network effects and fractals				
Unpredictable outcome chains				
Consequence of small changes				
Feedback loops among outcomes				
Joint optimization of uncorrelated outcomes				
Highly optimized tolerance (robust but fragile)				

Roots in many fields, e.g. Biology, Physics, Meteorology, Mathematics, and Economics, others. Some similar constructs are discovered independently. Sometimes cross disciplinary knowledge transfer.

Often does not fit with familiar assumptions about how the world works

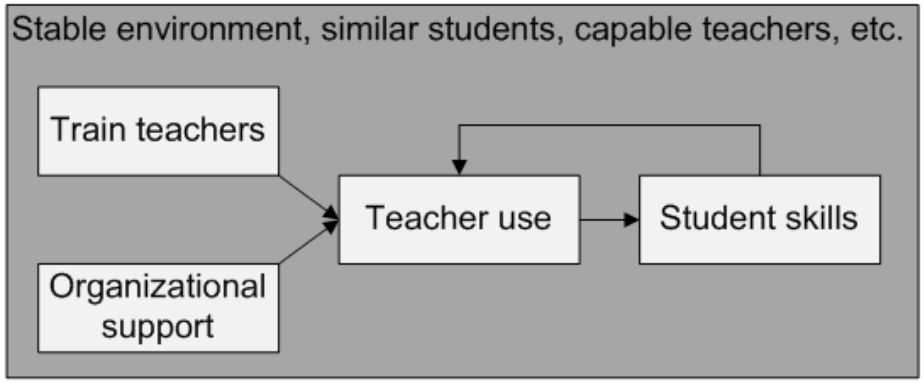
A pitch for taking complexity science seriously in evaluation

Three legitimate modes of drawing on complexity → Instrumental / Conceptual / Metaphor

Straightforward programs can exhibit complex behavior

Implications: Data / Methodology / Analysis strategies / Recommendations / Stakeholder expectations

- (1) many:1
- (1) 1:1 relationship
- (1) feedback loop
- (4) elements



Disciplines are unique

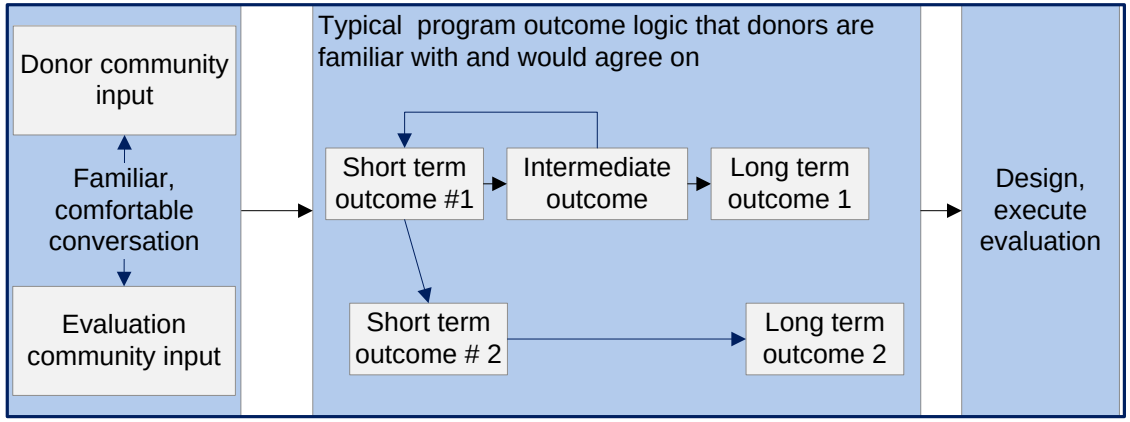
- interpreting data
- developing models
- defining data needs
- generating hypotheses
- choosing methodologies
- identifying topics to research
- specifying acceptable answers

Implications for

- Individual concepts do not matter
- I's the thought process that derives from thinking about all of them

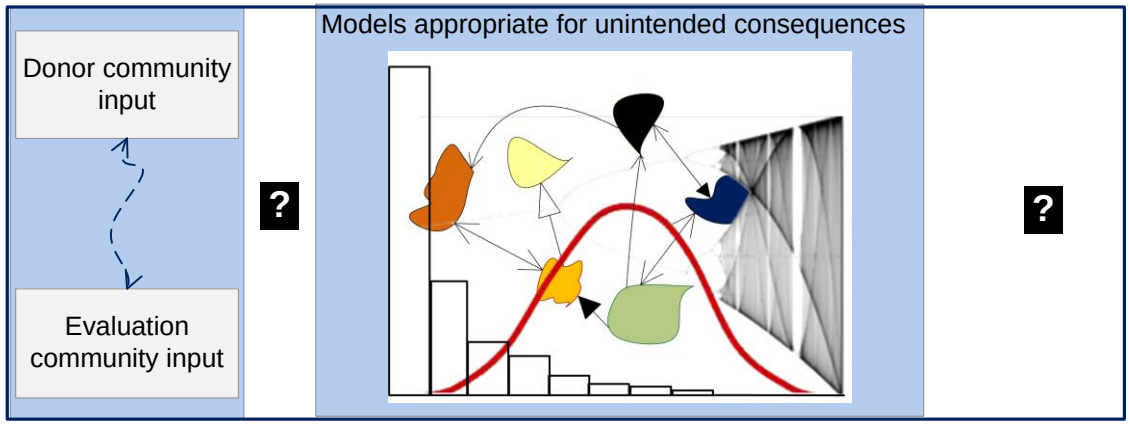
We need to get better at talking to stakeholders about complexity.

We know how to have this conversation.



But some program behaviors and outcomes can be:

- Unfamiliar
- Uncomfortable
- Deviate from common sense

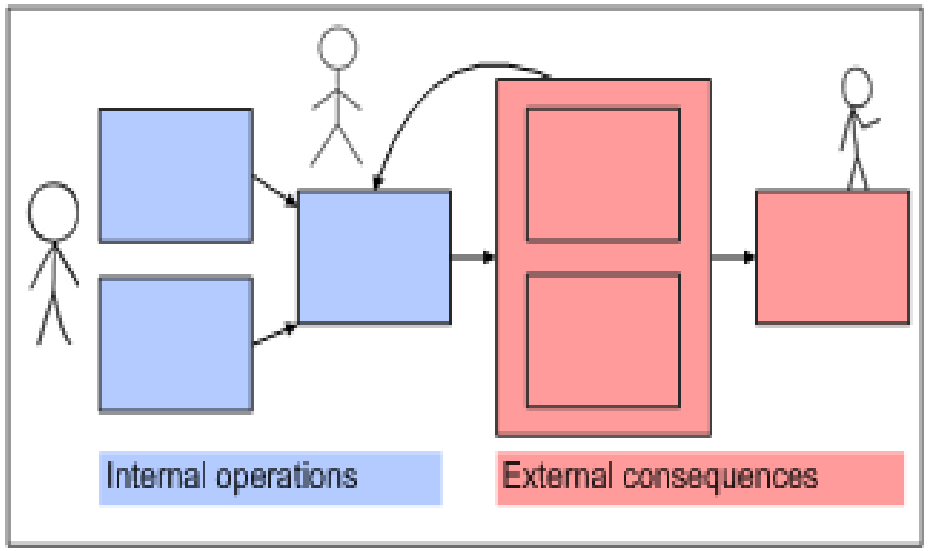


Easy to apply familiar methodology to complex behaviors.

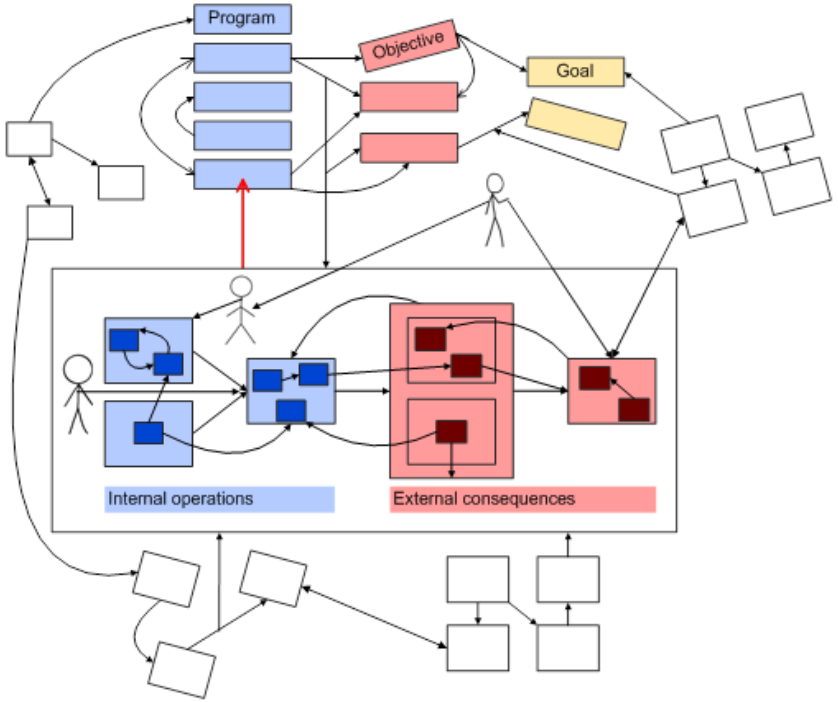
Hard for planners to take complex behavior seriously.

Ignoring complexity can be rational, adaptive behavior.

Can these people

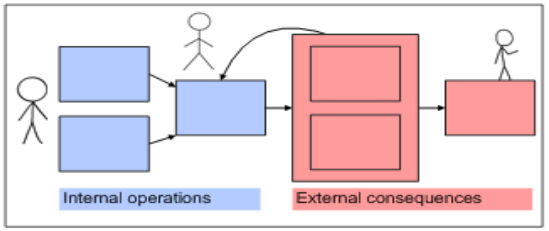


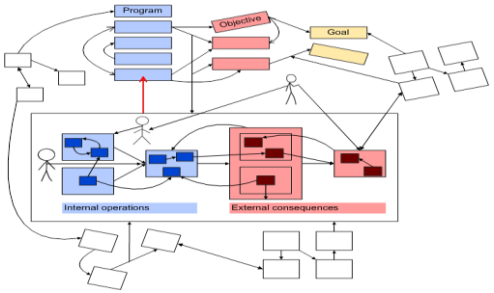
Budget, plan and implement effectively in this environment?



Tangled reality of setting priorities and program design

- 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
- Reality that programs serve purposes besides stated goals
- Different people have their favorite sources of information
- People get information from lots of sources other than formal evaluation





What are some tactics that might help

- Loose coordination
- Include unintended consequences in evaluation design
- Evaluate based on complexity even if you can't design for complexity
- Others?

What are some tactics that might help?

- Advocate loose coordination
- Support the value of short-term success
- Build on acceptance of imprecise logic models
- Advocate for methods to detect unintended consequences
- Appreciate different patterns that complex behavior may generate
- Others?

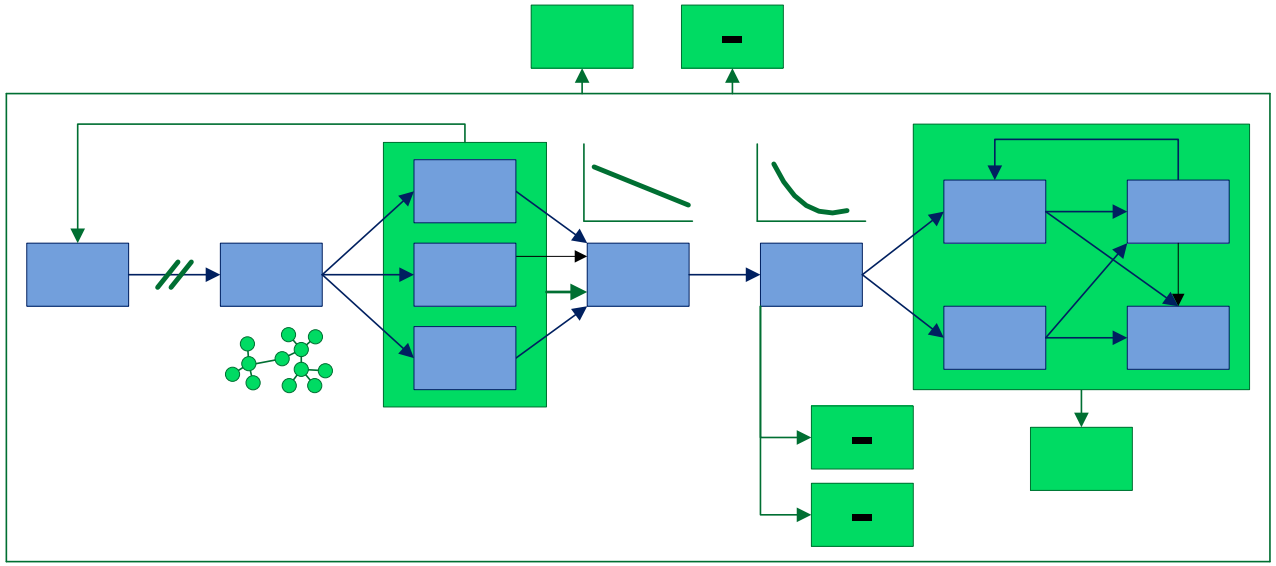
Best if evaluators and planners appreciate evaluation, but plenty of value if it's only the evaluators

If program designers build models that do not incorporate complex behavior they will

- Miss important relationships
- Not be able to advocate effectively
- Not be effective in making changes to improve their programs
- Misunderstand how programs operate and what they may accomplish

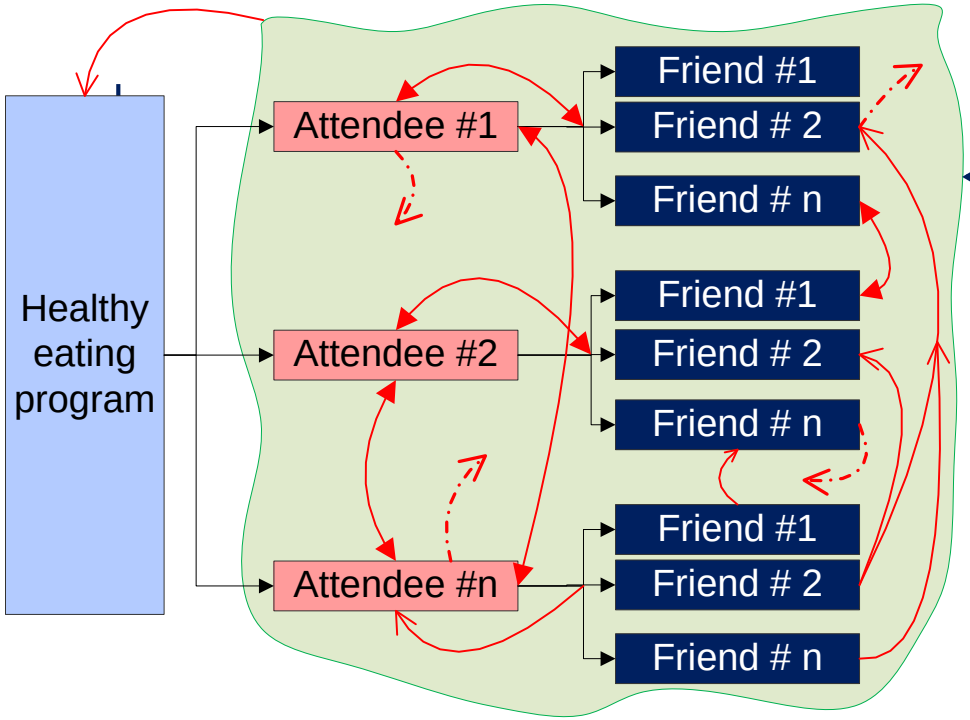
These problems cannot be fixed in an evaluation, but

- It is still possible to evaluate the complex behavior that is in their models.



Complex behavior, plebian methodology

Feedback
- Track changes in services over time
- Interview staff on perceptions of need



Unpredictable change emanating from community activity
- Monitoring
- Observation
- Open ended interviewing
- Content analysis of community social media

Program theory
- Desired outcomes can be specified.
- Path to desired outcomes cannot be specified. (Sensitive dependence)

Exceptions
- Agent and system dynamic modeling
- Formal network structure

Parts 3 and 4
Collective Design of an Evaluation / Redesign based on complexity

Three community initiatives

- Healthy eating
- New parent support
- Teen alcohol use

All three implemented at about the same time in the same community

- Each needs to be evaluated independently, but
- Some interest in overall community level change
- Evaluations begin at or near to start of program implementation
- Each program funded for three years with another two for evaluation
- Funding from different agencies but collective interest in cross-program effects and community level change.
- Separate evaluation teams, but some professional and personal overlaps.

Minimal constraints on the evaluation

- Budgetary
- Methodological

Healthy Eating by Working with Civic Organizations

Teaching/coaching about healthy eating:

- Proven "best practice" curricula
- Training on material and effective delivery
- Funding support for providing training

Choices to meet community needs

- Assistance for each organization set up their own healthy eating programs
- Each organization provides necessary resources.

New Parent Support – Birth to First Grade

Classes for parents

- Based on best practices from the research literature
- Civic organizations and local health centers
- Meet once per week for three months
- Attendance at more than one series allowed

Coaching

- In home visits by a social worker
- Telephone support
- Referrals to other services as needed.

Minimizing Teen Consumption of Alcohol

Peer to Peer

- Identify teachers, and students to run program
- Provide materials, examples of successful programs
- Provide ongoing technical assistance

Enforcement

- Publicize danger of providing alcohol to teens
- Training to help retailers recognize false IDs
- High-publicity increased enforcement actions
- Publicizing names of offending stores

Community level change: Each evaluation has some (but not a lot) of resources to study community level change not directly attributable to any specific program.

Part 5: Topics We May Consider Depending on Interest and What we have not Already Covered

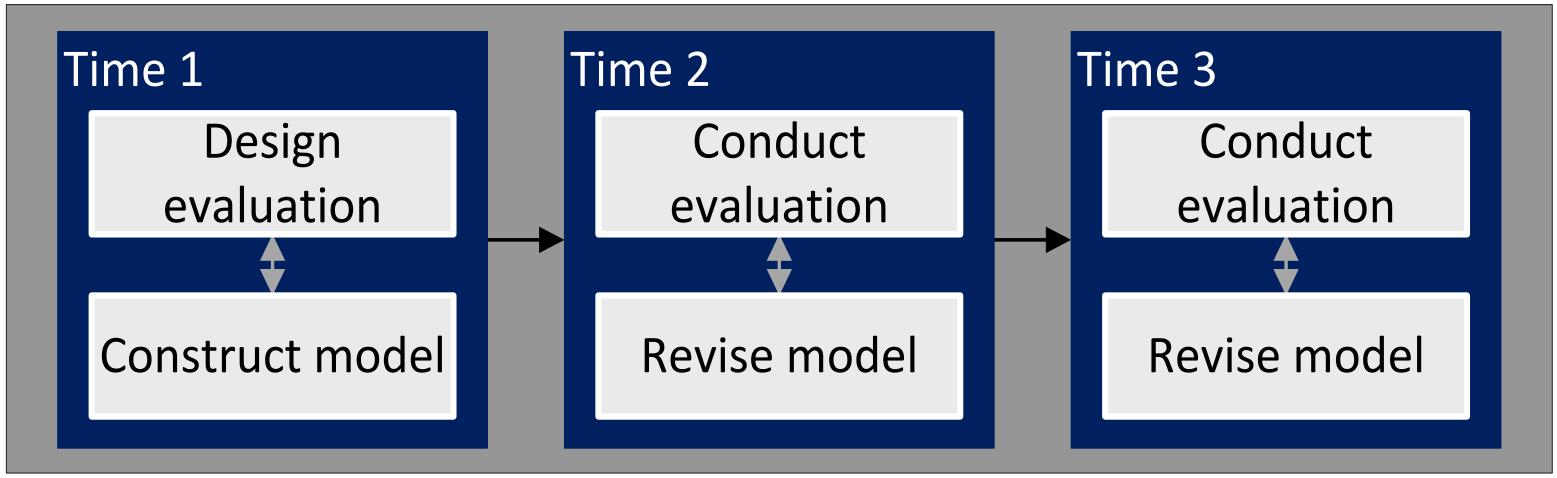
Topic	Slide #s
Types of change	15
Model change over time (maybe toward complexity but one way or another they do change)	16 - 17
Model specificity (the less specific, the more honest)	18 - 20
Model range of “trustability” (limited maybe, broad no)	21 - 22
Sensitive dependence and outcome chains (The best-laid plans of mice and men often go awry.)	23 - 26
Nested models (don’t trust aggregating up)	27
Effective programs may not have identifiable intermediate outcomes (easy for us, hard for our customers)	28
Evolutionary biology and ecology – uses in evaluation (some, if we focus on groups of programs)	29 - 32
Power law distributions (analysis easy, acceptance by stakeholders, maybe not so much)	33 - 35
Feedback loops and timing (something we should take seriously and don’t)	36
Network as constructs for explaining impact (often profound, often not)	37 - 39
Emergence (when can we and can’t we use a system’s parts to understand the system)	40
Attractors (as instrumental for program theory and analysis, and as metaphor for explanation	41 - 44
The agent-based view of how the world works (individual units yes, means and variances, no)	45 - 46
Is it worth evaluating with respect to complexity? (Sometimes, sometimes not)	47

Types of change

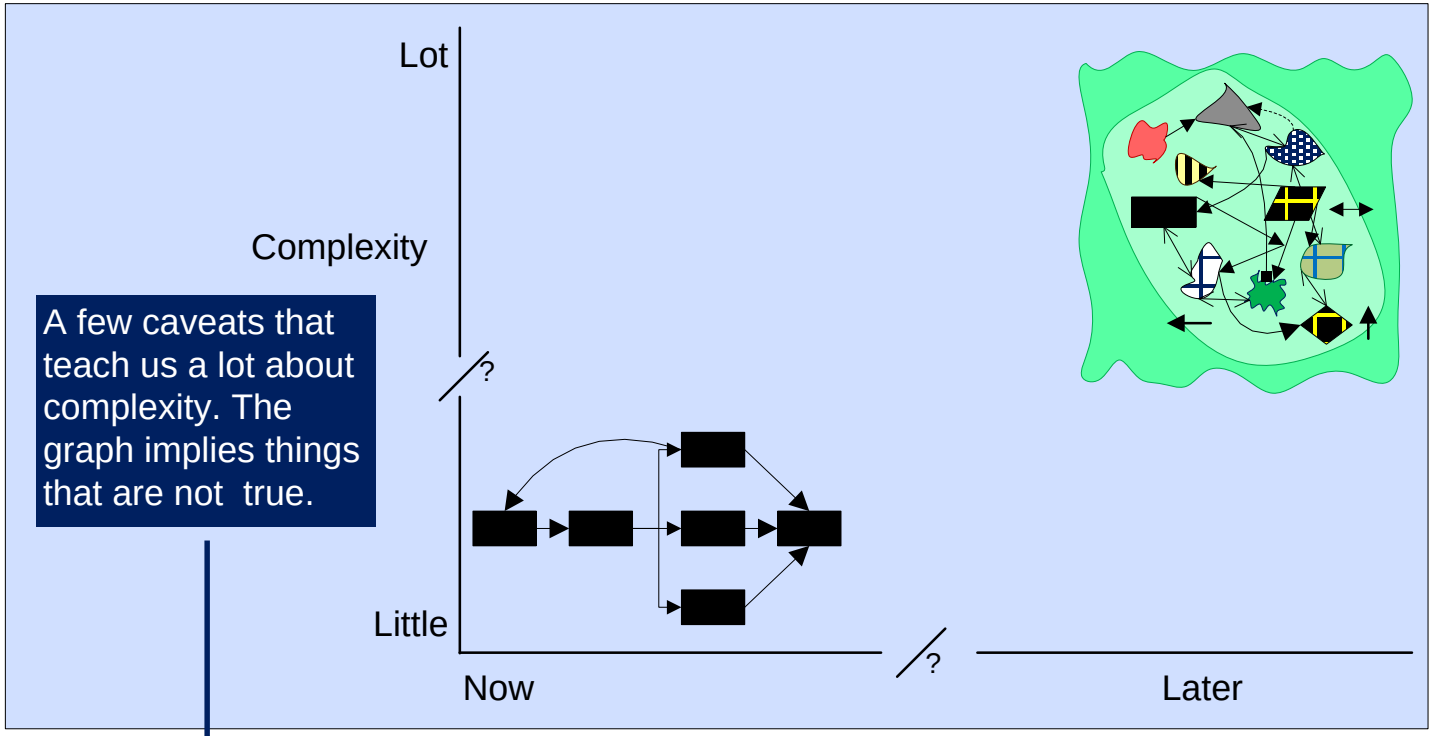
Phase shift	Sudden change from one condition to another
Emergence	Change such that each component loses its unique identity
Expected causal	Identified model content in terms of elements and relationships
Unknown causal	Because models simplify, causal dynamics are going are unknown
From the outside	Events in a program's environment can make a difference
Sensitive dependence	Small (often random) changes affects an entire trajectory over time.
Preferential attachment	The rich get richer, e.g. snowflakes, Internet nodes

Systematic iteration between data and model can

- maximize time to adjust methodology and
- keep program theory relevant.



The commonly accepted view of logic models and program theory may be less and less correct as time goes on.

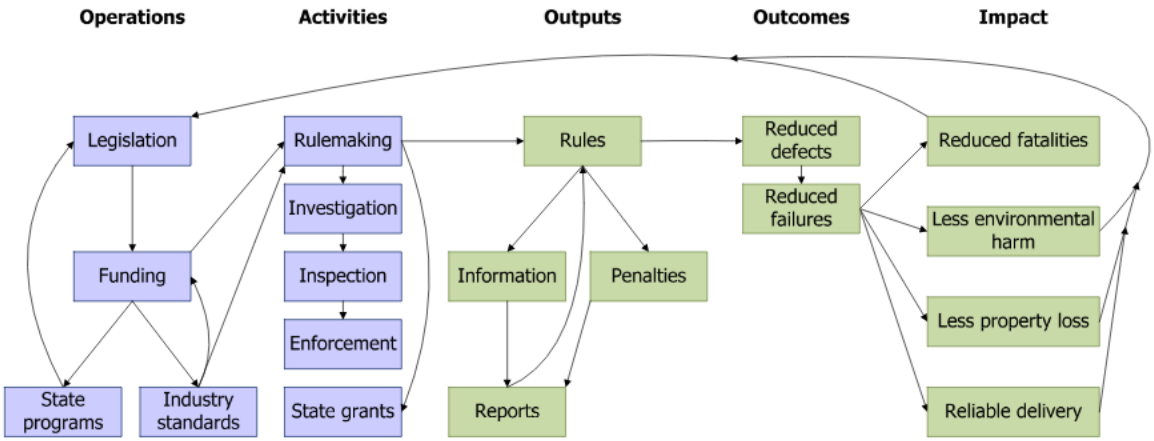


A few caveats that teach us a lot about complexity. The graph implies things that are not true.

- There are degrees of complexity
- There is a smooth transition to complexity.
- A transition to complex behavior is inevitable.
- For any place on the graph there is only one applicable model.
- Complex behavior can be ignored early in the project life cycle.

There are limits on what we can predict or explain

- Multiple causal paths within an attractor
- Limits on our detailed knowledge

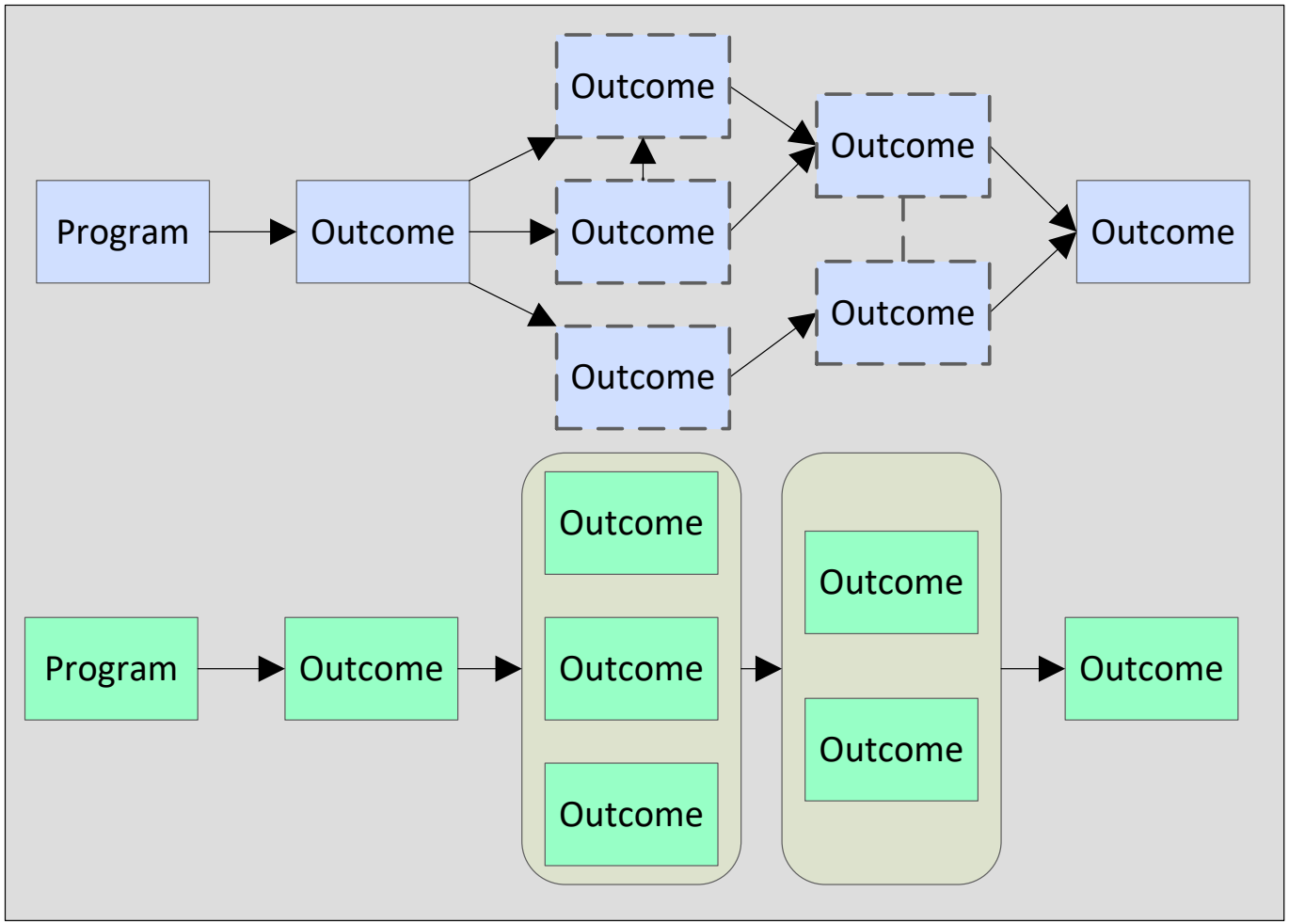


If things happen here.

Inputs	Throughput	Outputs	Outcomes	Long Term
Legislation	Rulemaking	Rules	Reduced defects	Reduced fatalities
Funding	Inspection	Reports	Reduced failures	Less environmental harm
Industry	Enforcement	Penalties		Less property loss
Industry standards	Investigation	Information		Reliable delivery
State programs	State grants			

Things will happen here.

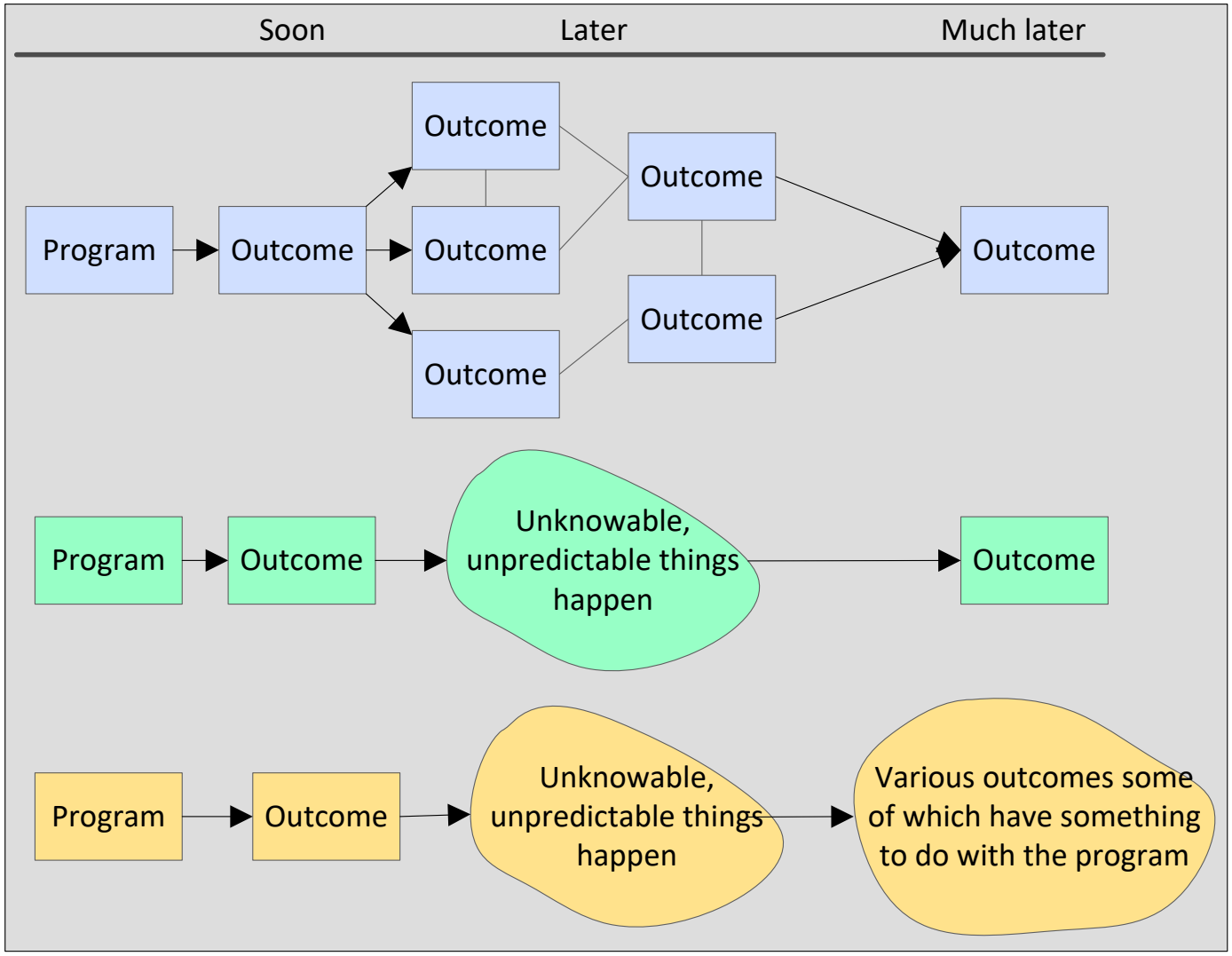
How much do we really know about a program?
Maybe some modesty is called for.



Who has a harder time with this, funders or evaluators?

Three models for the same program.

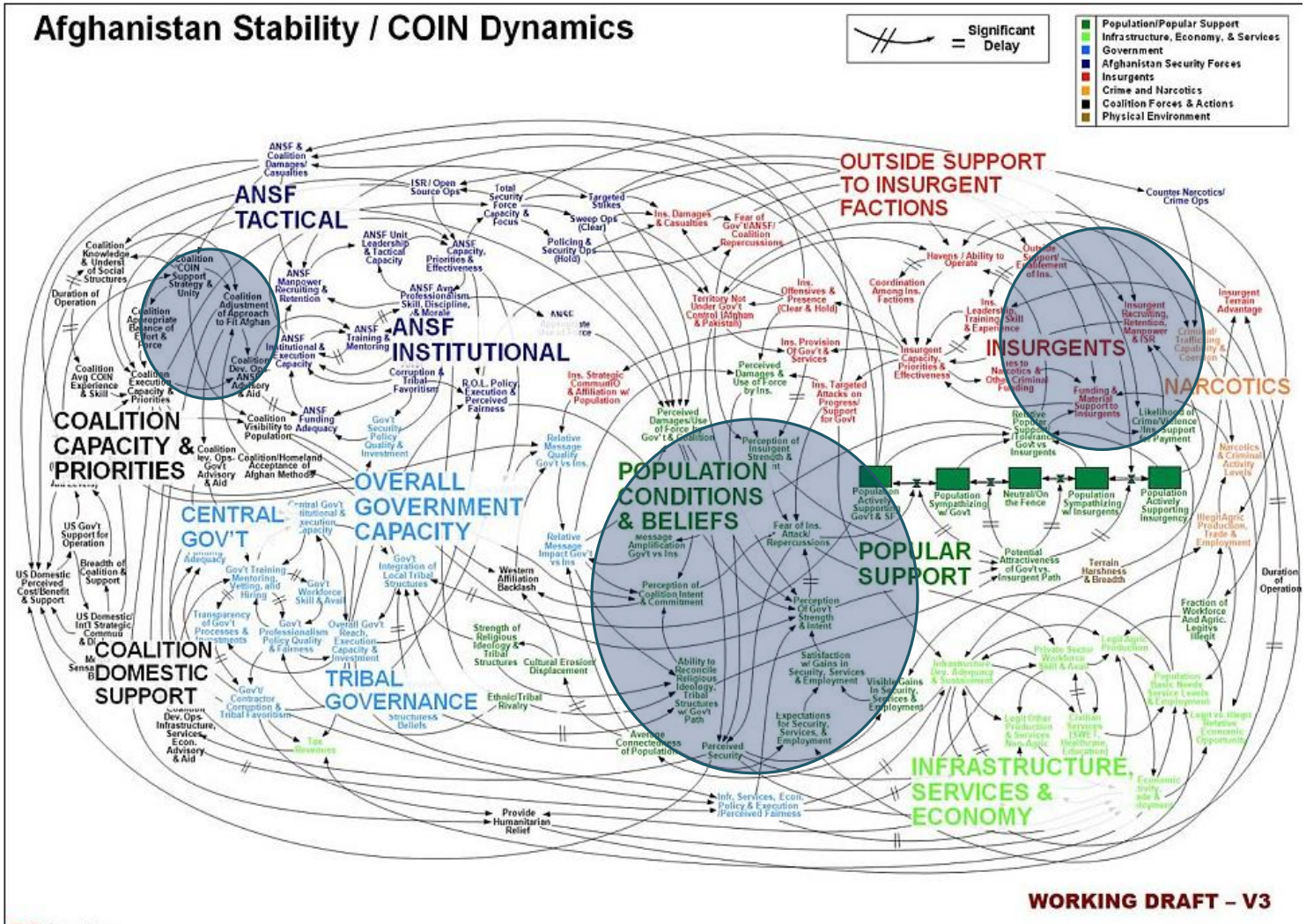
- Knowing nothing about the particulars, which model would you bet your $5.00 on?
- Could you design an evaluation to encompass more than one model?
- Could you convince your customer to buy an evaluation that encompassed more than one model?



Models can be locally correct but not globally correct.

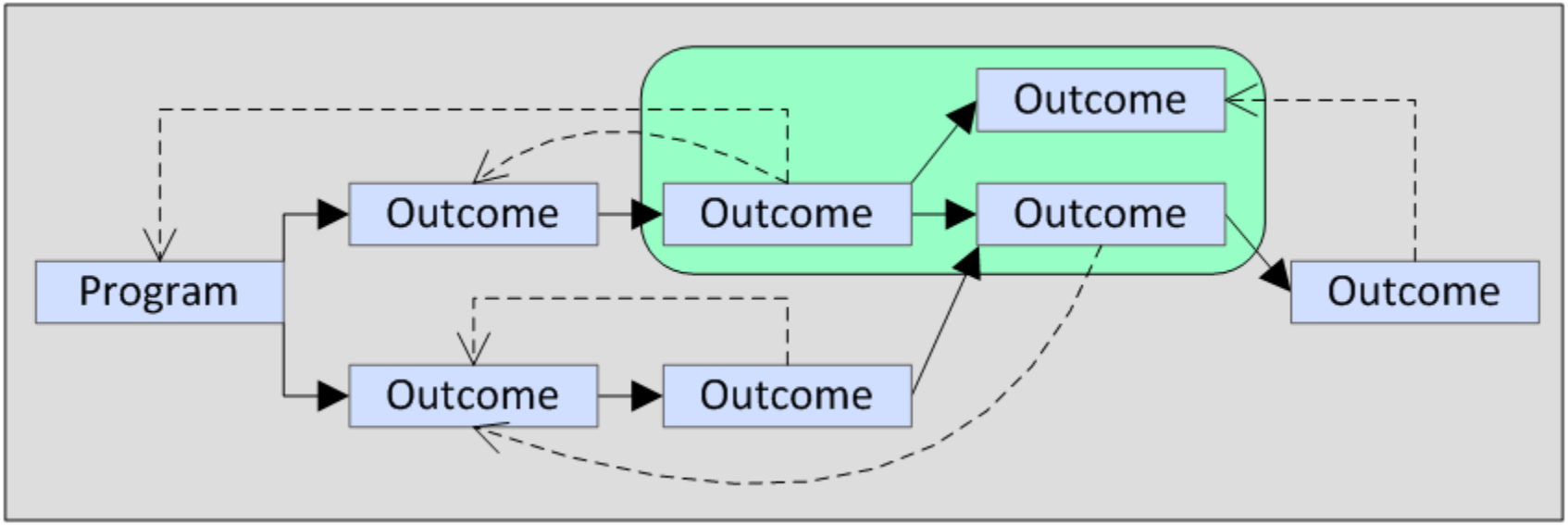
- Emergent behavior
- Adaptive network
- Sensitive dependence

- Non-linear feedback loop effects
- Phase shifts
- Embedded levels of detail



This version comes from the Daily Mail, April 28, 2010.

Because of feedback loops and sensitive dependence, a program theory may be everywhere locally correct, but never globally correct.



No matter how many specific outcomes stakeholders can specify, they cannot understand long term impact by combining the impacts they are sure of.

Does it matter to stakeholders?
Does it matter to evaluators?
If it does matter, what can be done about it?

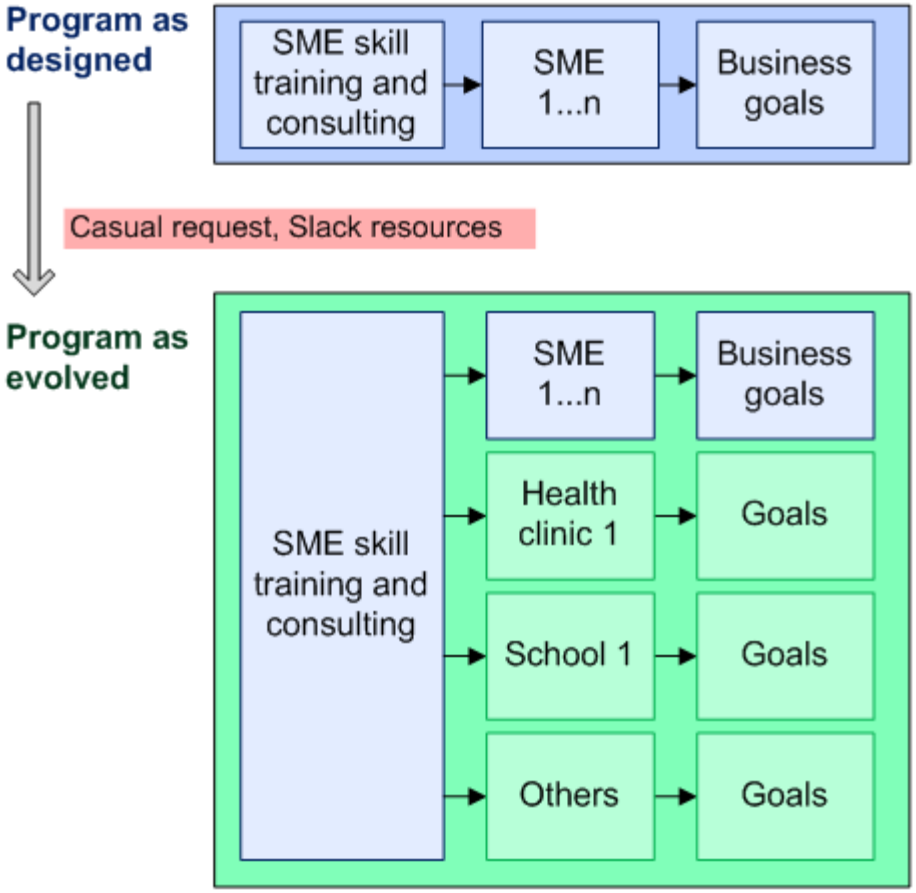
Because of sensitive dependence, it may be impossible to specify an outcome chain.

Planners and funders

- Expertise
- Timelines
- Advocacy
- Coordination
- Program theory
- Funding sources
- Societal benefits

Evaluation

- Lead time to implement changes to the evaluation
- Event sequence may be unique but knowing it can help with future planning



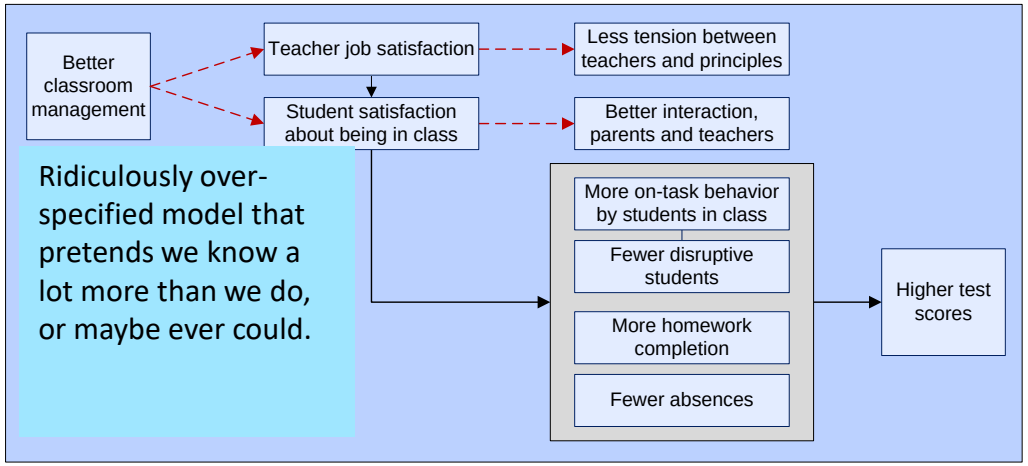
Because of sensitive dependence, it may be impossible to specify an outcome chain.

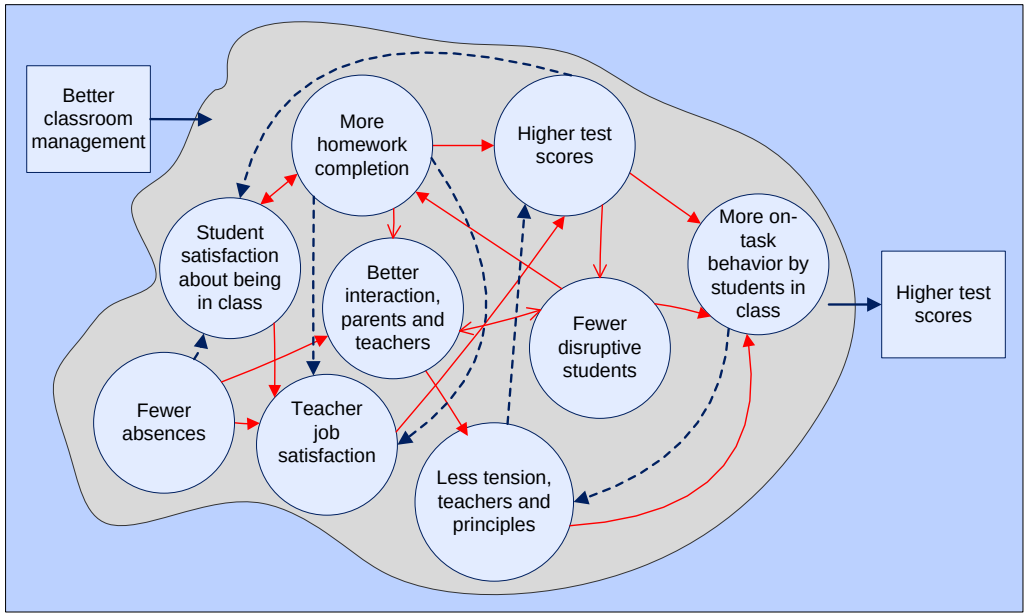
Example 2: uncertain network effects

We and our stakeholder are comfortable with models like this.

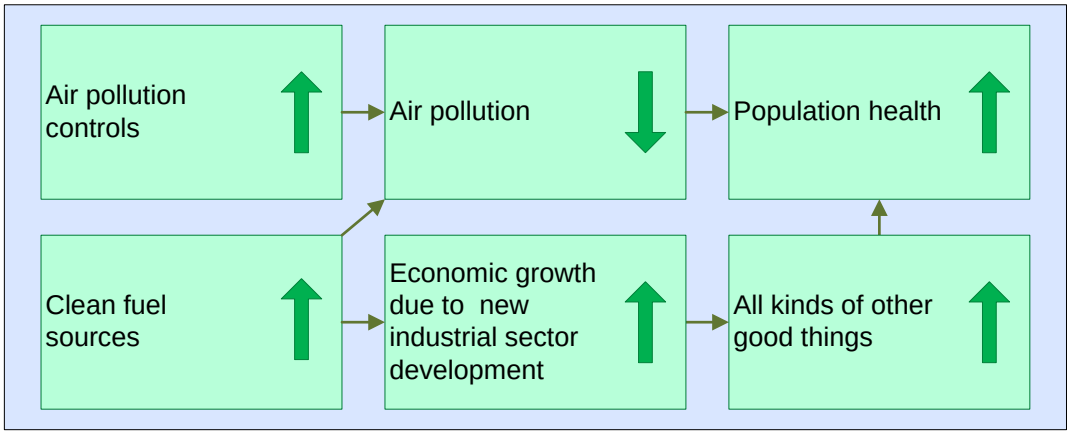
When is it appropriate to use each of these?

Will funders and other stakeholders be OK with this?
What might convince them to accept it?





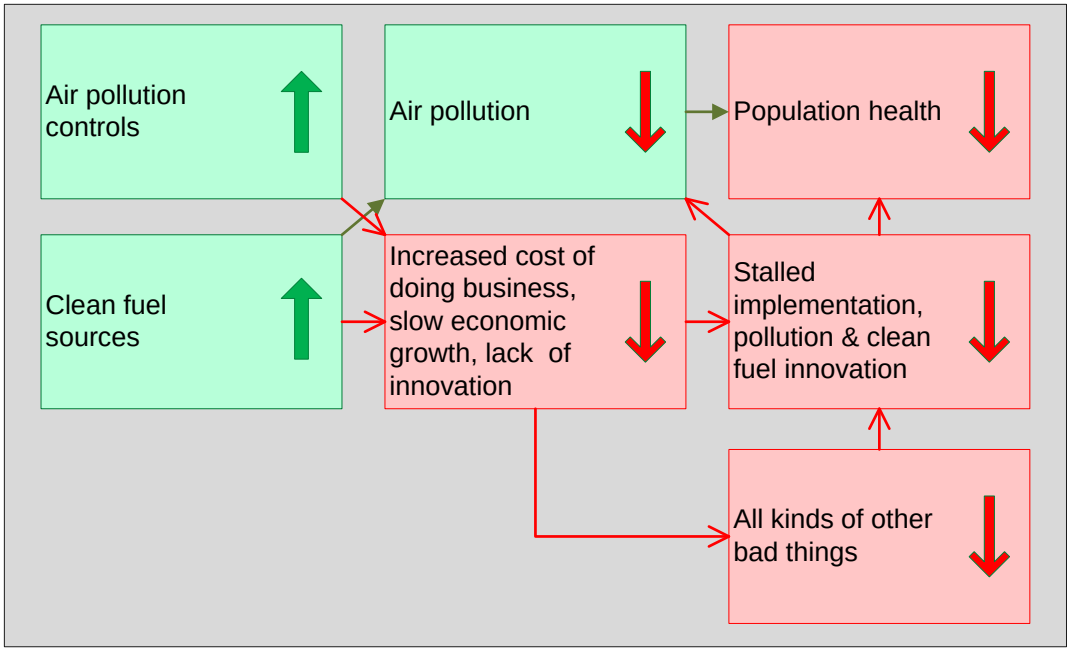
It may not be possible to predict which competing program theory will be correct



How many factors

- Large and small
- identifiable and unidentifiable

Would have to line up to activate one or another of these models?



Is there any reason for an evaluation to test only one?

And in case anyone thinks they can beat the odds and predict an outcome chain…

TO: President George W. Bush

April 12, 2001 10:22 AM

CC: Vice President Richard B. Cheney
Honorable Condoleezza Rice

FROM: Donald Rumsfeld

SUBJECT: Predicting the Future

You will recall that we had Andy Marshall come over, and we briefly discussed the defense strategy review and what the future might hold.

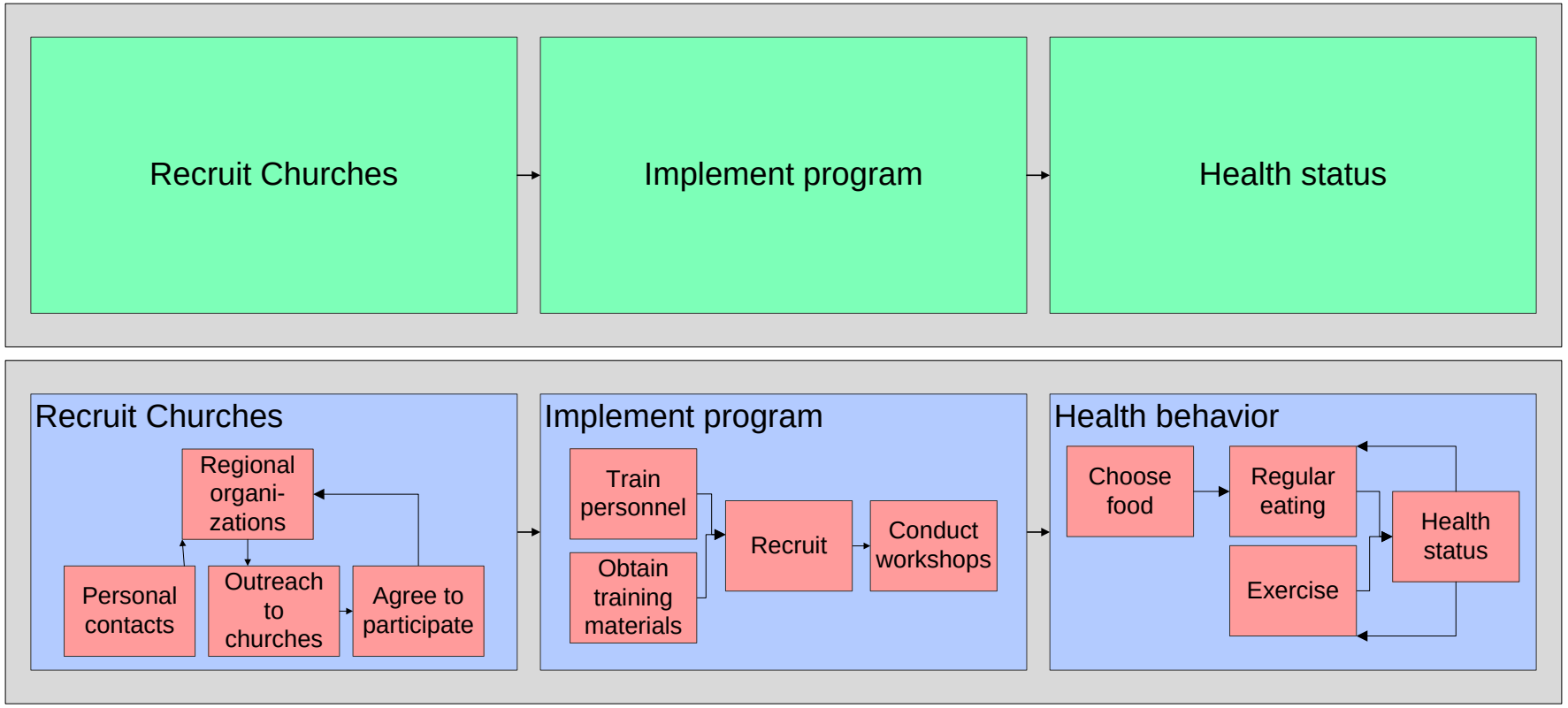
I ran across this piece on the difficulty of predicting the future, written by one of the folks here at the Pentagon, Lin Wells. I thought you might find it interesting.

Attach.
Lin Wells: "Thoughts for the 2001 Quadrennial Defense Review"

Thoughts for the 2001 Quadrennial Defense Review

- If you had been a security policy-maker in the world's greatest power in 1900, you would have been a Brit, looking warily at your age-old enemy, France.
- By 1910, you would be allied with France and your enemy would be Germany.
- By 1920, World War I would have been fought and won, and you'd be engaged in a naval arms race with your erstwhile allies, the U.S. and Japan.
- By 1930, naval arms limitation treaties were in effect, the Great Depression was underway, and the defense planning standard said "no war for ten years."
- Nine years later World War II had begun.
- By 1950, Britain no longer was the world's greatest power, the Atomic Age had dawned, and a "police action" was underway in Korea.
- Ten years later the political focus was on the "missile gap," the strategic paradigm was shifting from massive retaliation to flexible response, and few people had heard of Vietnam.
- By 1970, the peak of our involvement in Vietnam had come and gone, we were beginning détente with the Soviets, and we were anointing the Shah as our protégé in the Gulf region.
- By 1980, the Soviets were in Afghanistan, Iran was in the throes of revolution, there was talk of our "hollow forces" and a "window of vulnerability," and the U.S. was the greatest creditor nation the world had ever seen.
- By 1990, the Soviet Union was within a year of dissolution, American forces in the Desert were on the verge of showing they were anything but hollow, the U.S. had become the greatest debtor nation the world had ever known, and almost no one had heard of the internet.
- Ten years later, Warsaw was the capital of a NATO nation, asymmetric threats transcended geography, and the parallel revolutions of information, biotechnology, robotics, nanotechnology, and high density energy sources foreshadowed changes almost beyond forecasting.
- All of which is to say that I'm not sure what 2010 will look like, but I'm sure that it will be very little like we expect, so we should plan accordingly.

It can be problematic to assume that for a program to succeed, its models must be correct at all levels of detail.



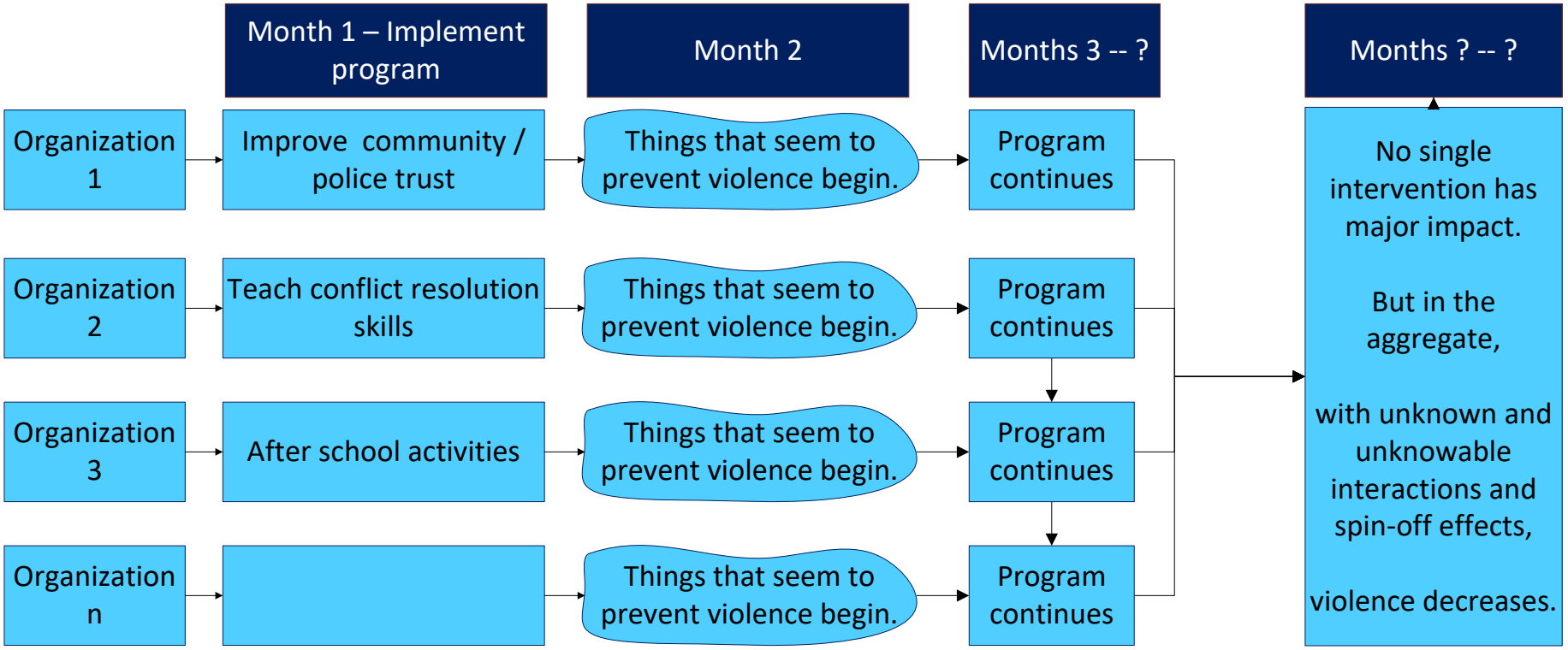
Effective programs may have no foreseeable processes or outcomes.

Good programs may have to be funded on faith.

Problem: Violence in Gotham City

Program theory: Because local community organizations

- know local conditions, and
- have credibility with community leaders
- they can conceive, implement, and run successful interventions.



$ and opportunity costs committed to uncertain programs. Can this be justified? When?

Can Knowledge of Evolutionary Biology and Ecology Inform Evaluation?

Quick overview
Longish (~ 3,500 words)

Disciplines have unique ways of

- Interpreting data
- Developing models
- Defining data needs
- Generating hypotheses
- Choosing methodologies
- Identifying research topics
- Specifying acceptable answers

Some useful concepts from evolutionary biology and ecology

- Population
- Coevolution
- Birth/death rates
- Selection pressure
- Species and species variation
- Emphasis on rates of change
- Understanding of diversity, adaptability, context, and related concepts

It's not a matter of choosing any one of these concepts.

What matters is the thought process that derives from thinking about all of them.

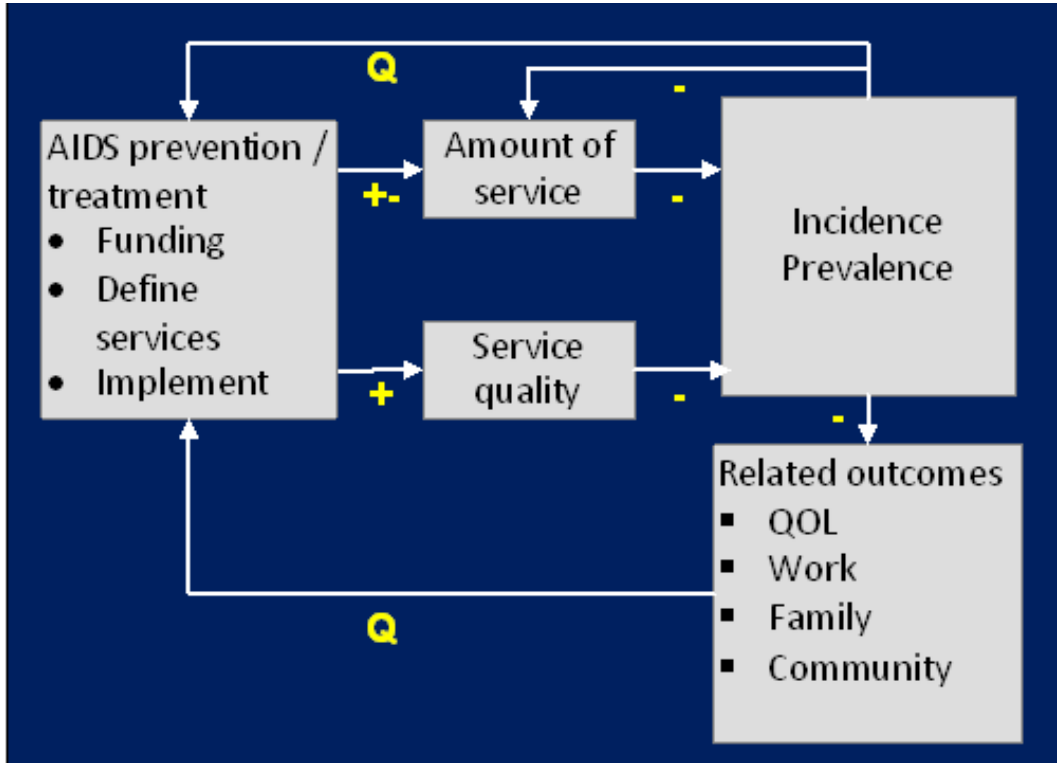
Thinking of theories of change in terms of evolutionary biology and ecology

A nice, traditional, comfortable model.

All outcomes are highly correlated

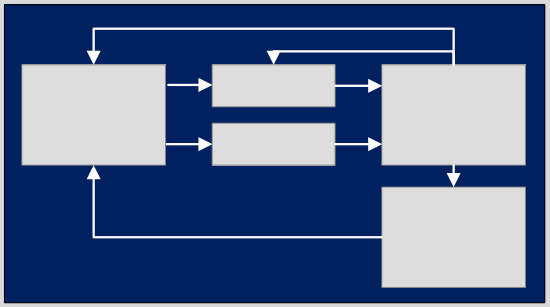
This is a fine program theory. I'd love a chance to do this work.

But let's recast the program theory in adaptive, evolutionary terms.



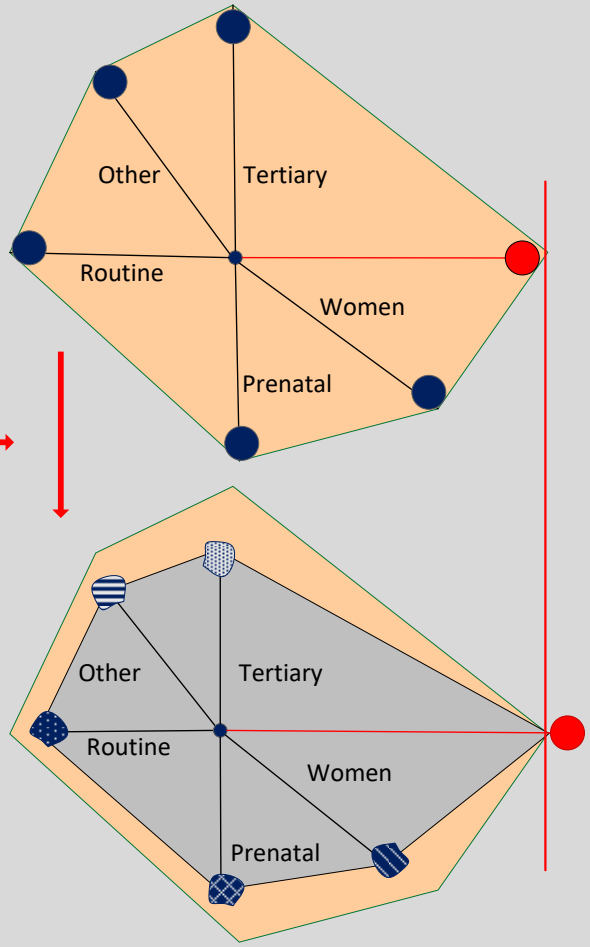
Neither program theory is inherently good or bad.

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

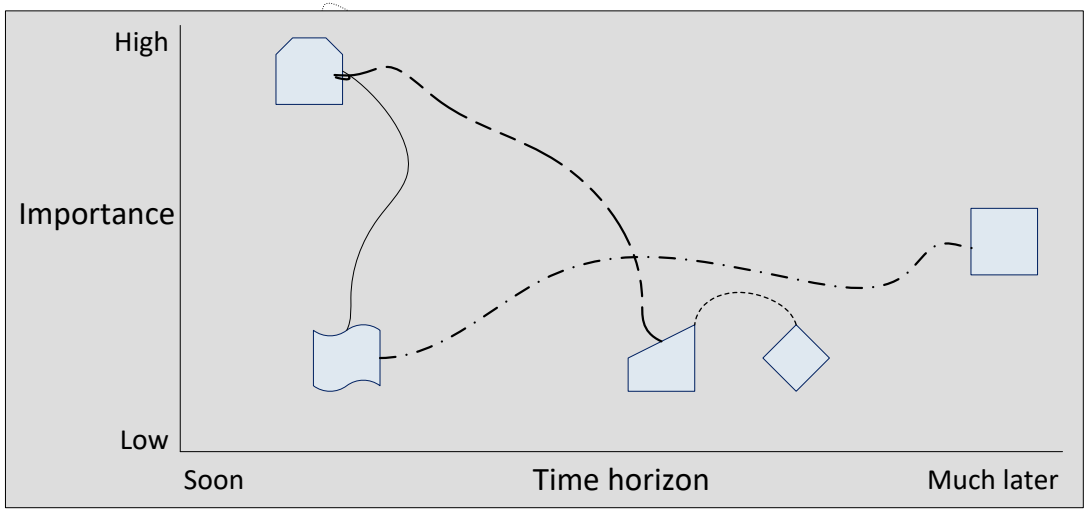


How can we decide which to use for any given evaluation?

Another view of why pursuing single-program goals won't work.

Goal structure Time 1

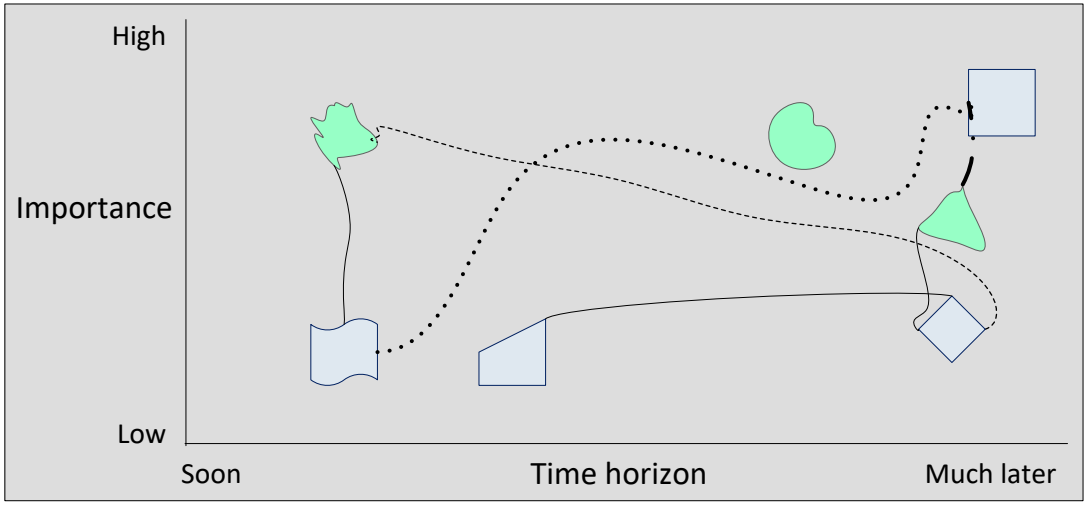
Total resources available = 100



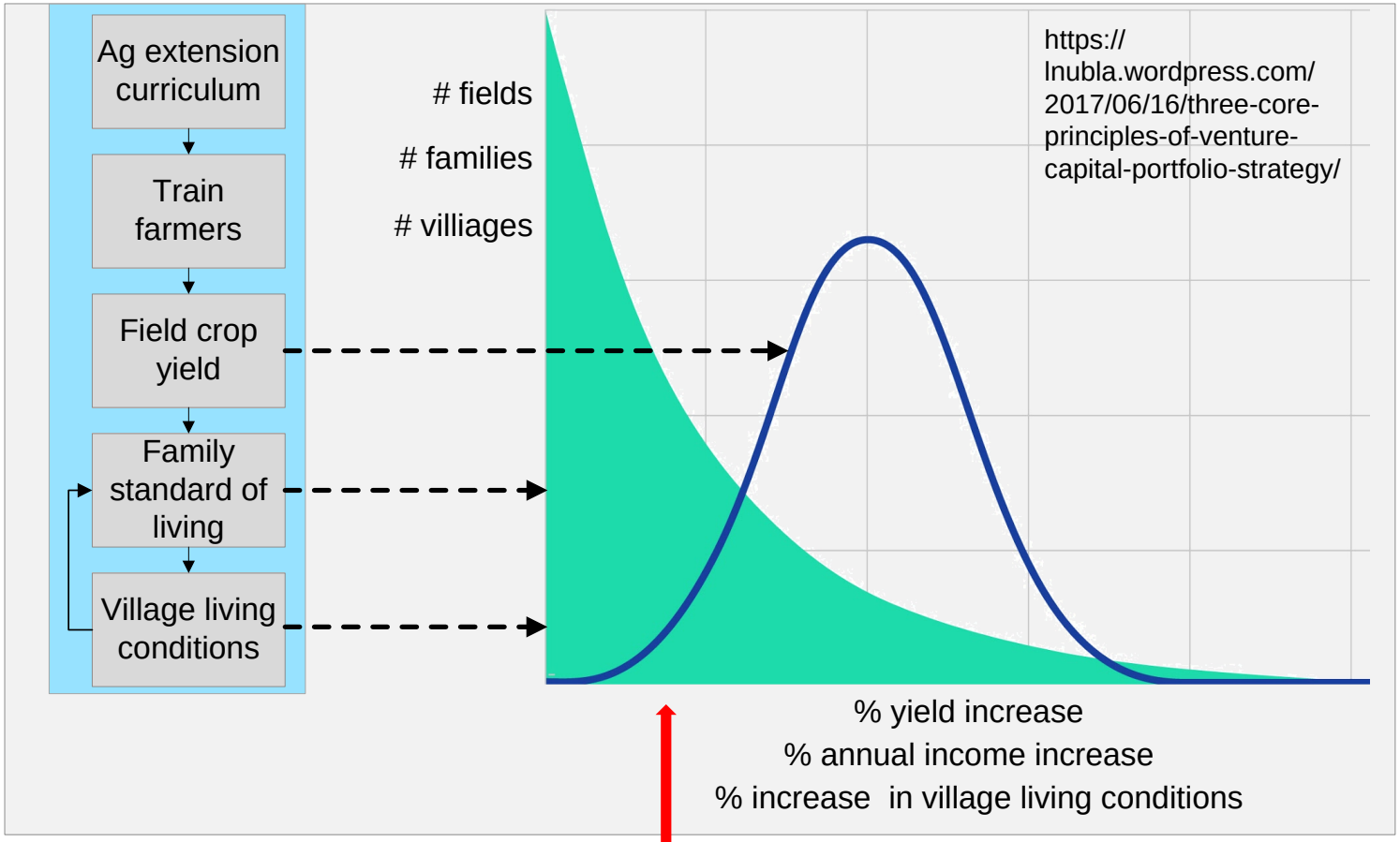
- Goals connected (or not) in odd and often unknowable ways
- Priorities change
- Some old goal disappear
- Some new goals appear

Goal structure Time 2

Total resources available still = 100

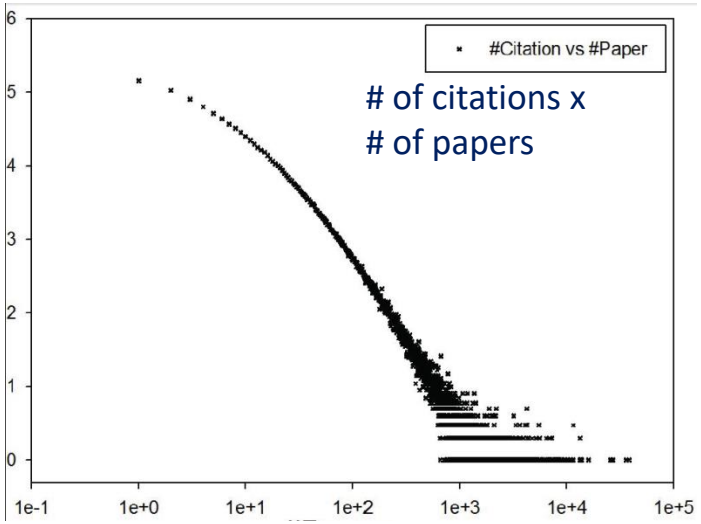


- Success may mean that the rich get richer.
- In a very successful program benefits may be not be symmetrically distributed.
- Evaluation methodology, straightforward. The politics and values? Not so much.



Will stakeholders overtly acknowledge this outcome distribution?
What are the consequences for evaluation if they do or don't?
What are the consequences for public policy if they do or don't?

Power law have consequences for planning, program theory, and expected outcomes



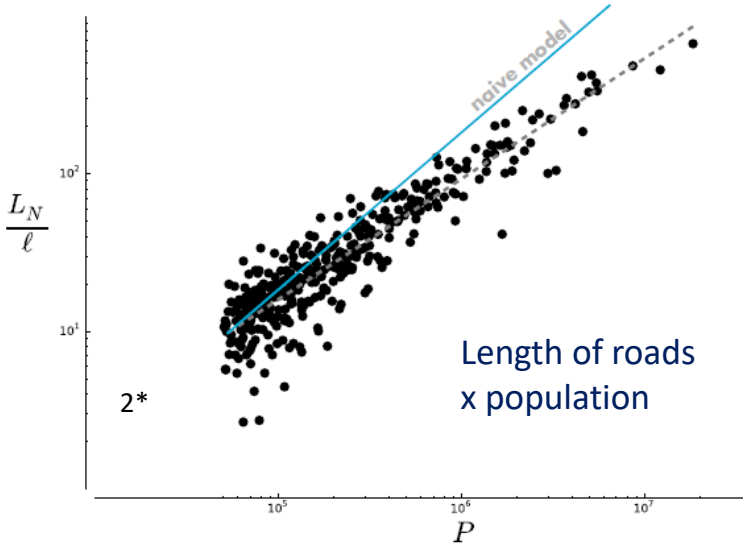
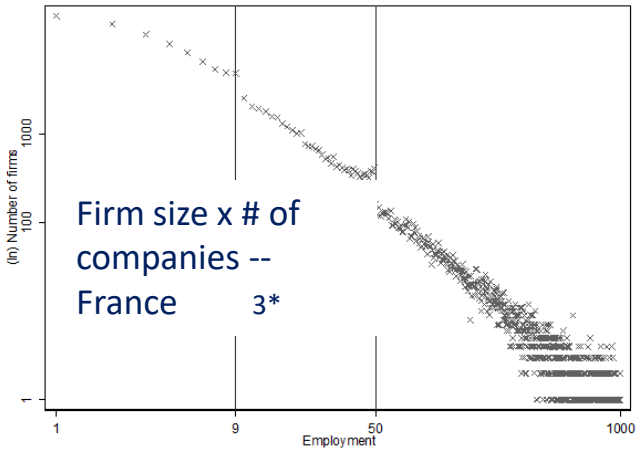
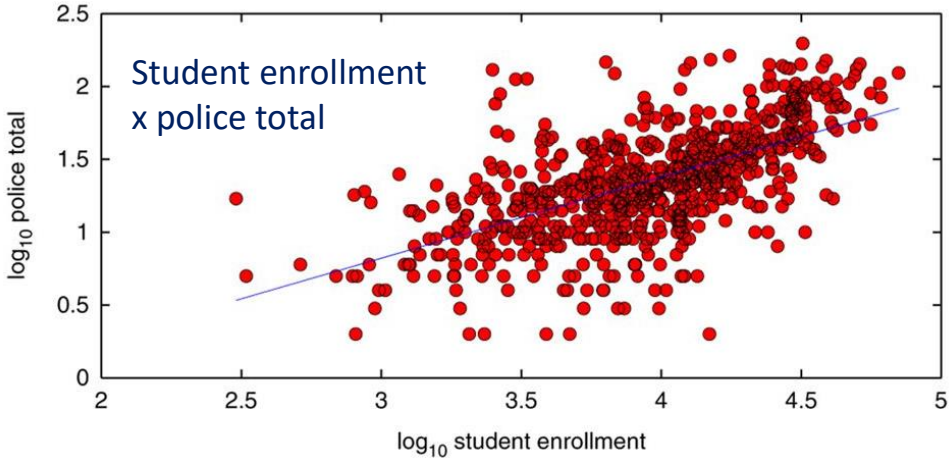


Figure 1. Firm-size distribution in France: Power law broken at regulatory thresholds





1*- Yan, Tang, Liu, Shan, & Li, 2011)Yan, R., Tang, J., Liu, X., Shan, D., & Li, X. (2011). Citation count prediction: Learning to estimate future citations for literature.

2*- Louf, R. (2015). Wandering in cities: a statistical physics approach to urban theory.

3*- https://voxeu.org/article/small-isn-t-always-beautiful-cost-french-regulation

4*- Banerjee, S., Van Hentenryck, P., & Cebrian, M. (2015). Competitive dynamics between criminals and law enforcement explains the super-linear scaling of crime in cities. Palgrave Communications, 1, 15022. doi:10.1057/palcomms.2015.22

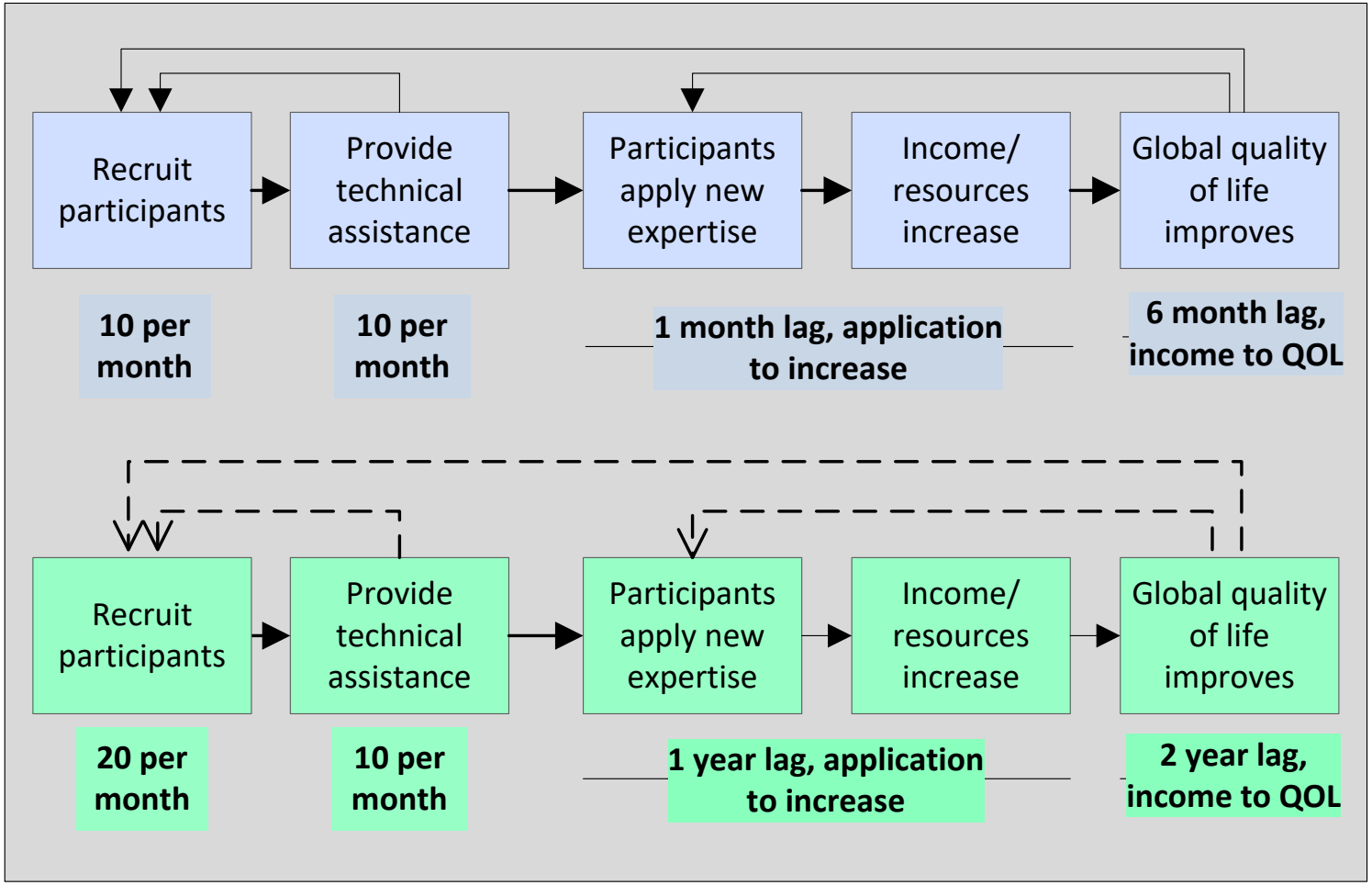
Power laws are found in a wide variety of settings

Natural science			
Cities	Traffic jams	Coastlines	Brush-fire damage
Water levels in the Nile	Hurricanes & floods	Earthquakes	Asteroid hits
Sun spots	Galactic structure	Sandpile avalanches	Brownian motion
Bach's music	Epidemics	Genetic circuitry	Metabolism of cells
Networks in the brain	Tumor growth	Biodiversity	Circulation in plants, animals
Langton's Game of Life	Fractals	Punctuated equilibrium	Mass extinction, explosions
Brain functioning	Predicting premature births	Laser technology evolution	Fractures of materials
Social science			
Language	Social networks	Internet	Blockbuster drugs
Sexual conquests	'Fordist' power	Wealth	Citations
Co-authorships	Actor networks	Job vacancies	Salaries
Firm size	Supply-chains	Growth rates of firms	Growth & internal structure
Casualties in war	Countries' GDP growth rates	Stock price movements	Delinquency rates
Movie profits	Consumer product sales	Size of villages	Cotton prices
Economic fluctuations	Biotech alliance networks	Entrepreneurship/innovation	

*Sources are given in Andriani and McKelvey (2005).

https://www.researchgate.net/figure/Listing-of-some-power-law-discoveries_tbl1_30051663

Because feedback loops can produce nonlinear behavior, the details of their operation matter.



It is entirely possible that the different latencies of these feedback loops will result in very different performance of the same logic.

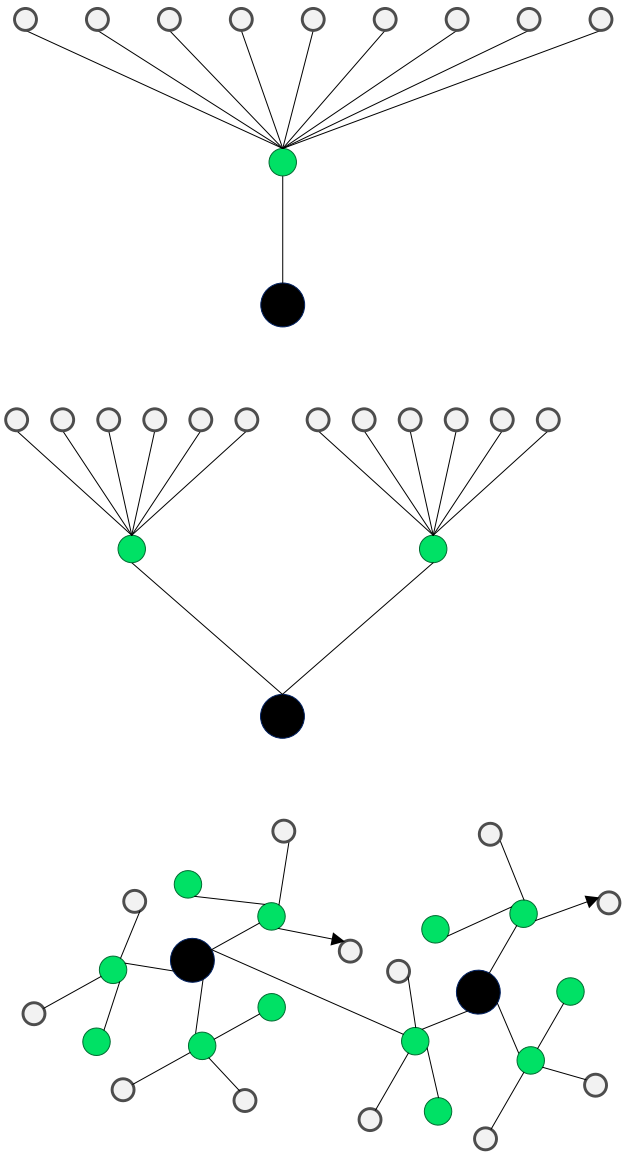
Evaluation that wanted to add robustness and efficiency as metrics might consider network structures.

Programs

- Primary, secondary, tertiary health care
- Disaster supplies stockpiling
- Technical assistance centers
- Etc.

Metrics

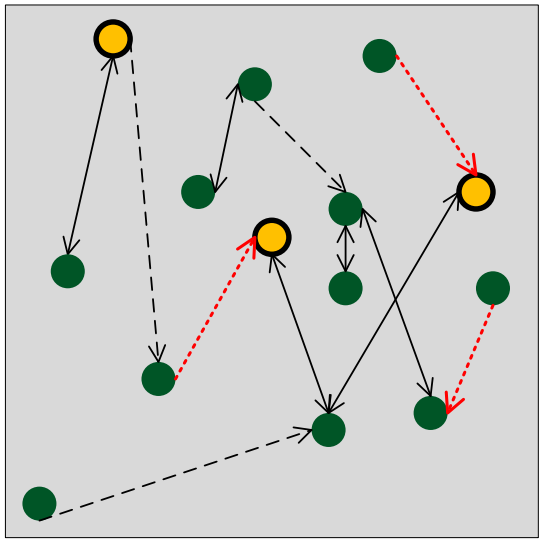
- Cost
- Response time
- Robustness
- Population level indicators
- Quality of service to individuals
- Etc.



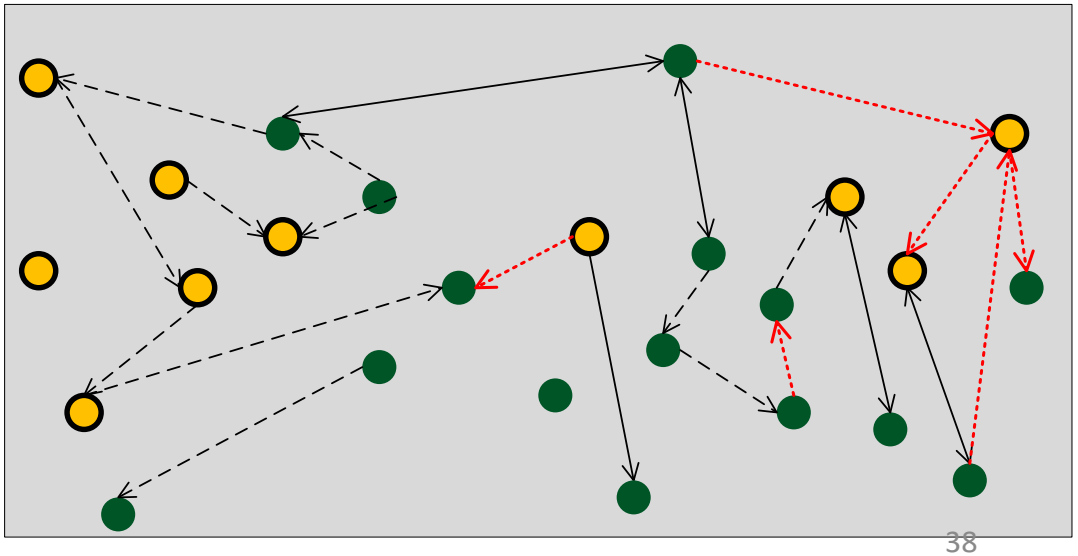
Some evaluations may care about network structure and behavior

- Size
- Density
- Directionality
- Types of nodes
- Connectedness
- Change over time
- Types of relationships

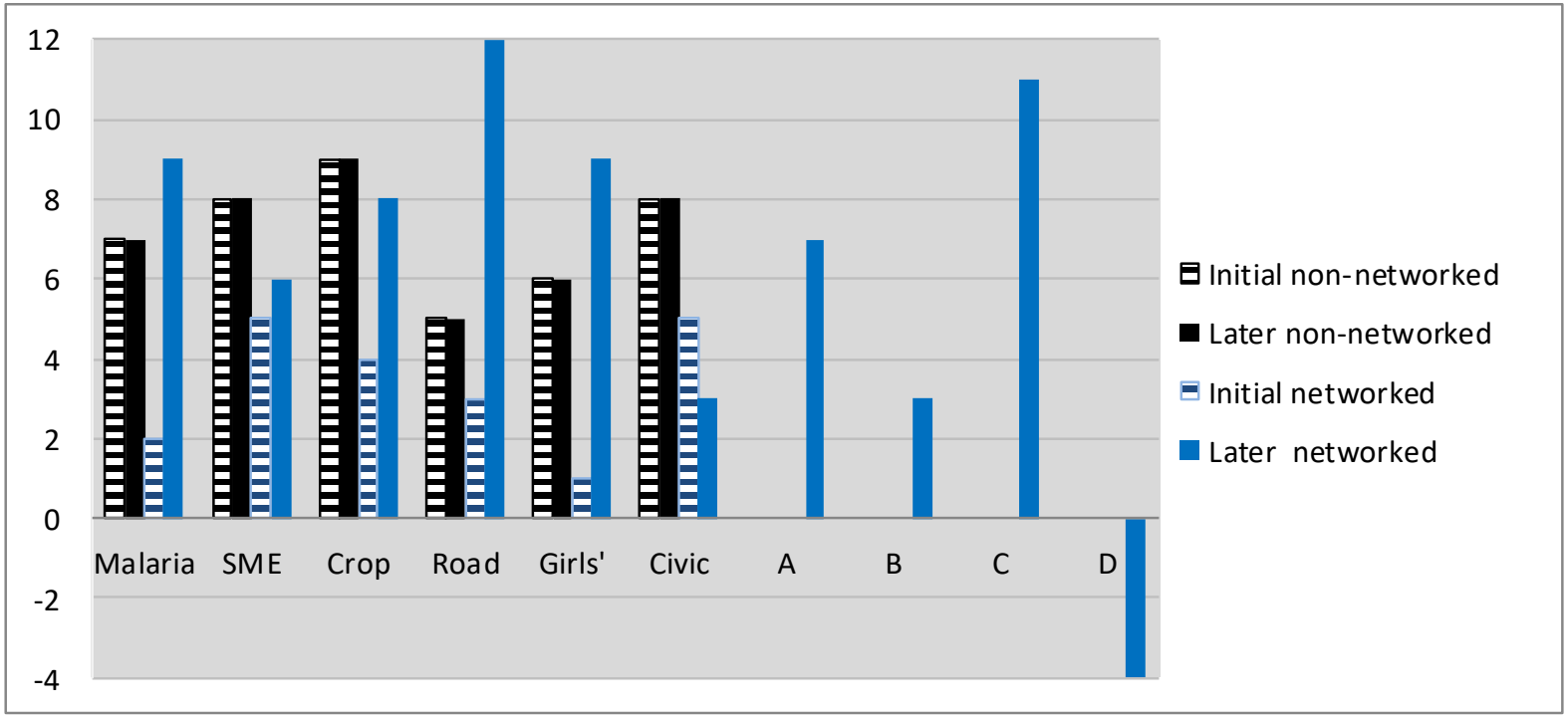
Now



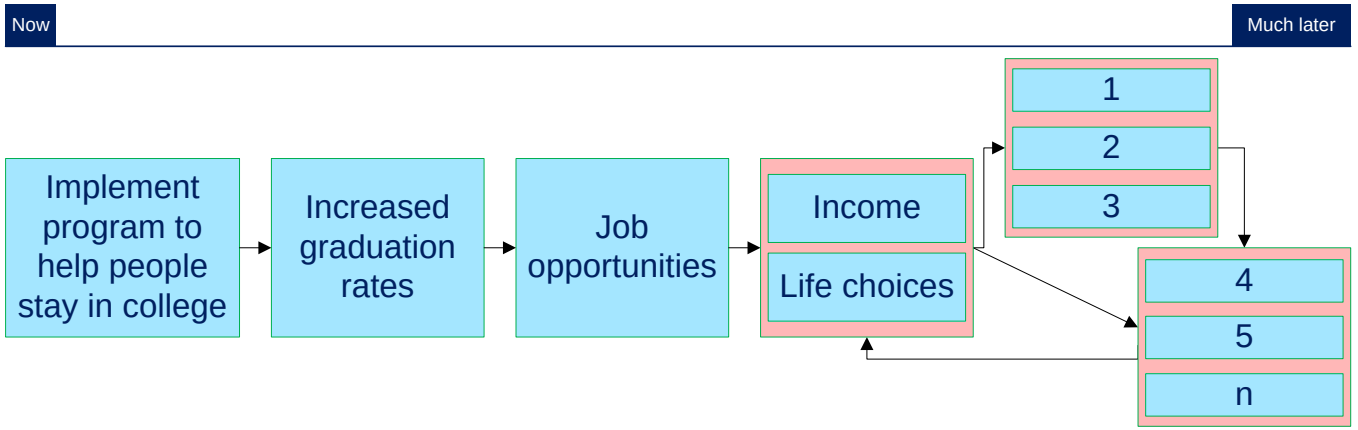
Later



Here is a better look at the data



But it's also possible for a single program to kick off profound impact



Emergence: Behavior a “whole” cannot be explained by the characteristics of its “parts”.

The whole is greater than the sum of its parts, but no emergence

*

Emergence on a non-human scale

**

Implications for

- How a program works
- Stakeholder comfort
- What needs to be measured

- Buyers, sellers, transactions interact to form Markets
- Relationships among people result in culture
- Mom, Dad, and the kids interact to form Family
- Lots of families interact to form a Community
- Communities interact to form a nation-state

Emergence on a human scale

* https://en.wikipedia.org/wiki/Cylinder_(engine)

** http://forum.canberrabees.com/t/natural-bee-hive-removal-in-narrabundah/50

Attractors

Different way of thinking about stability and predictability

If anything, our programs are too stable

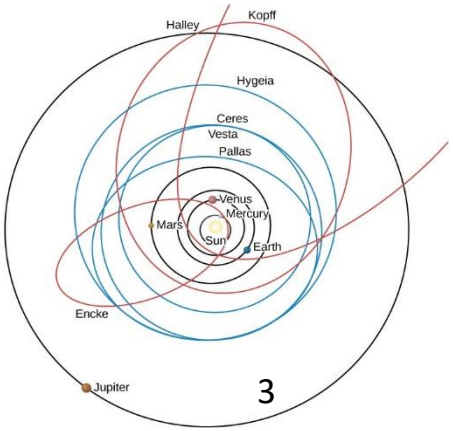
- Beat them over the head with data and they don't change

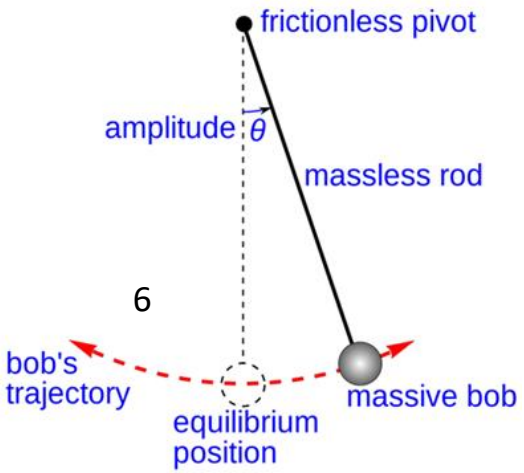
Why is it useful?

- Program theory that captures historical trends
- Conceptualize change in terms of shape and stability
- Parsimonious way to classify a whole variety of settings.
- Insight about program behavior outside of stakeholder beliefs
- Promotes technological perspective: What will happen, not why
- Hint where to look for similar programs when designing evaluation

What are attractors?

- Values close enough remain close if slightly disturbed
- Properties toward which a system evolves, regardless of starting conditions





1 http://www.watershedconnect.com

2- https://unofficialnetworks.com/2019/02/15/colorado-avalanche-worst-on-record/

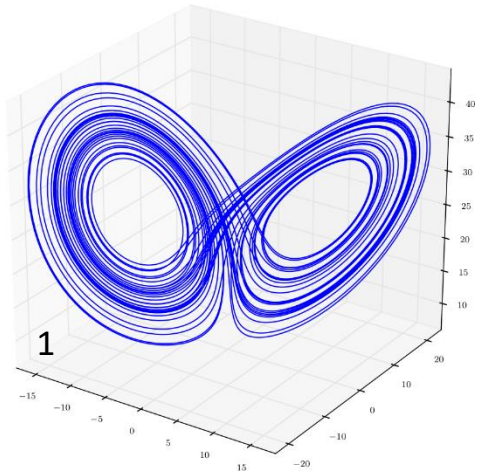
3- https://courses.lumenlearning.com/astronomy/chapter/orbits-in-the-solar-system/

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

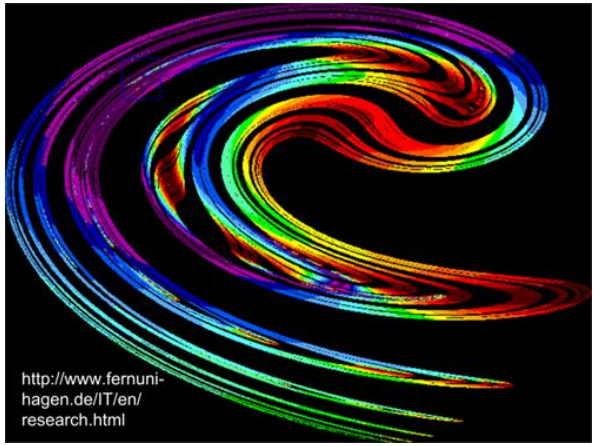
5 http://alpha.fdu.edu/~gl221015/

6- https://en.wikipedia.org/wiki/Pendulum

Strange attractors



- Impossible to know location from one instant to the next
- Know where it will not be
- Bounded chaotic movement
- Fractal



Insight

- Scale up
- Replication
- Realistic program theory
- Transfer to other settings
- Stakeholder expectations
- Identify program outcomes
- Methodology to touch range of outcomes

1- http://node99.org/tutorials/ar/

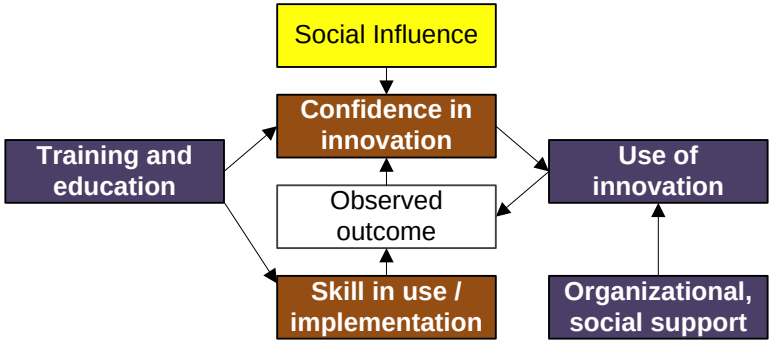
What is an Agent? What is an Agent Based Model?

An “agent” is an entity that can	▪ Sense its environment ▪ Respond to set rules
What can an agent be?	▪ Person ▪ Animal ▪ Wetland ▪ Hospital ▪ School room ▪ County government ▪ Etc.
What is in an agent-based model?	▪ Numerous agents (sometimes at different scales (e.g. teachers, schools) ▪ Decision rules (e.g. “Do this if your neighbor does it.”) ▪ Learning rules (e.g. If you do it twice in a row, continue to do it 90% of the time) ▪ Environment specifications (e.g. “New treatment is 25% better.”
▪ Why an agent-based approach?	▪ All variation, not just group means & variances ▪ Sensitive dependence ▪ Emergence

Traditional statistical versus agent-based epistemology

Traditional analytical models

- General linear model Y = XB +U.
- True score
- Error uncorrelated with true score, mean = 0
- With good data lets us discern group differences
- All that counts are means and variances
- Approach is powerful and useful. More is better.

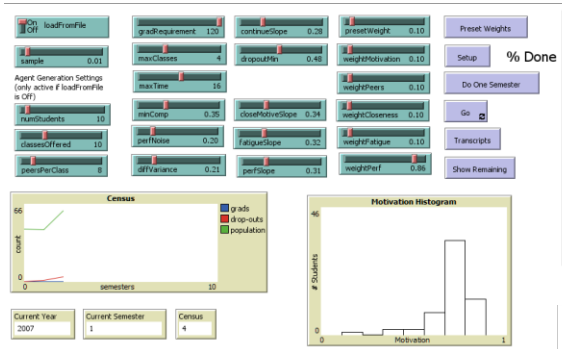


In the agent-based view *

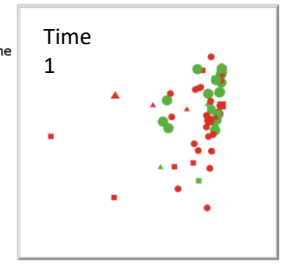
- Systems have history, small changes matter
- Means and variances are no longer how we view group behavior
- Evolutionary paths and attractor shapes are informative.
- Emergent behavior can be observed.

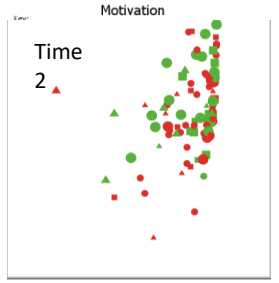
An "agent" is an entity that can

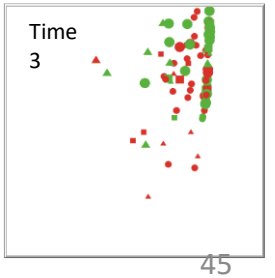
- Sense its environment
- Respond to set rules



Model animation





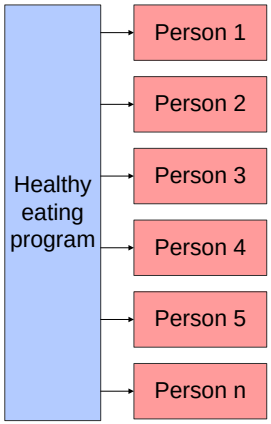


* Research supported by the Faster Forward Fund. Data based on agent-based simulation of data from an evaluation of the Arizona General Education Curriculum. Results summarized in three YouTube videos: Part 1 (7 min.) / Part 2 (20 min.) / Part 3 (20 min.)

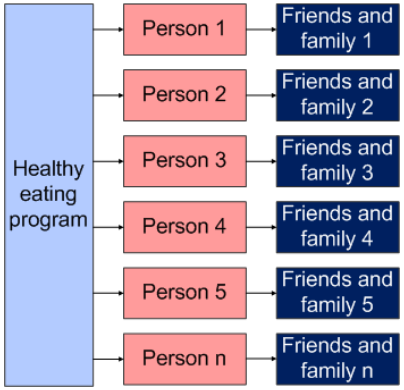
It's a choice whether to pay attention to complex behavior.

- What do evaluators need to consider when making this choice?
- What do our customers need to consider when deciding what evaluation to buy?

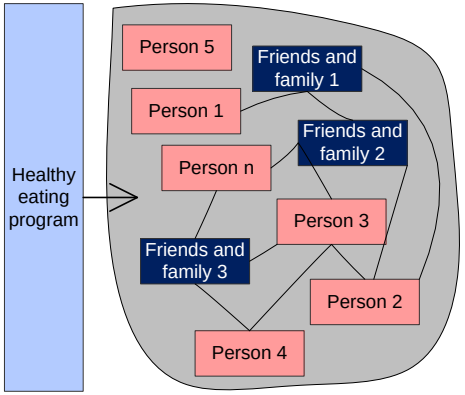
- Cheap and easy
- Results can be very useful to decision makers



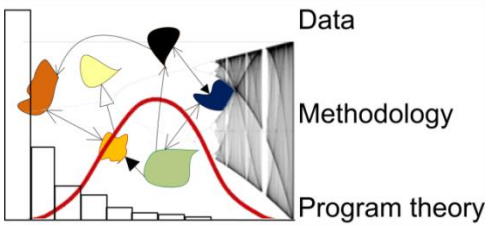
- Wider range of benefits
- Still a straightforward methodology
- More knowledge of community level benefits



- More work for qualitative efforts
- More likely to encounter unexpected outcomes
- Effort needed to assess community level effects
- Harder to find comparison groups because setting plays more of a role
- Unpredictable outcome chain, limits information for planning and advocacy
- State change behavior – might need more frequent data collection to identify inflection point.



Disclaimer



- I'm not trying to convert anyone into complexity devotees
- Much of the evaluation we do works just fine with traditional approaches

But

- Complex behavior is afoot in the world
- Taking complex behavior seriously may be useful
- Better program design
- Better evaluation
- Respect the good reasons to ignore complex behavior

Part 6
Parking Lot

https://premierpsychology.com.au/2016/06/the-traffic-jam-in-my-brain/

Part 7
Resources and further reading
(Nothing systematic. Just books and articles that I like.)

Further Reading on Complexity and Related Topics – Sources Outside of the Evaluation Literature Note: There is nothing systematic about this list. It is comprised of books and articles that I like.	
Ball, P. (2004). Critical Mass: How One Thing Leads to Another. New York: Farrar, Straus and Giroux.	Takes a complexity-informed view of a host of "big" social phenomena-
Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: Addressing the paradox of sustainment amid ongoing change. Implementation Science, 8(1)	Sustainability in terms of continual adaptation
Ferguson, N. (2017). The Square and the Tower: Networks and Power, from the Freemasons to Facebook. New York: Penguin Press.	Network effects in history from the printing press to the present day.
Frigg, R., & Hartmann, S. (2018). Models in Science. The Stanford Encyclopedia of Philosophy. Retrieved from https://plato.stanford.edu/archives/sum2018/entries/models-science/	Deep dive into the nature of models.
Hannan, M. T., & Freeman, J. (1989). Organizational Ecology. Cambridge MA: Harvard University Press.	Very technical and very interesting. Takes a population biology perspective to the rise and fall of types of organizations, treating them as species.
Kauffman, S. (1995). *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Okford UK: Oxford University Press.	Complexity and self-organization
Kellert, S. H. (1993). *In thge Wake of Chaos*. Chicago: University of Chicago Press	Shows how a complexity concept can be used to good advantage without strict adherence to its science and technical foundations
Lawlor, J. A., & McGirr, S. (2017). Agent-based modeling as a tool for program design and evaluation. Evaluation and Program Planning,	Agent-based modeling in evaluation

Further Reading on Complexity and Related Topics: Sources Outside of the Evaluation Literature Note: There is nothing systematic about this list. It is comprised of books and articles that I like.	
Marion, R. (1999). *The Edge of Organization: Chaos and Complexity Theories of Formal Social Systems*. Thousand Oaks CA: Sage.	Reinterpretation of organizational theories through a lens of complexity
Orrell, D. (2007). The Future of Everything: The Science of Prediction. New York: Thunder's Mouth Press.	Explanation of the inherent problems of prediction across a wide range of activity – weather, health, and more
Page, S. (2011). Diversity and Complexity. Princeton, NJ: Princeton University Press.	In-depth treatment of how diversity in populations can affect adaptability and evolution.
Silver, N. (2012). The Signal and the Noise. New York: Penguin.	Excellent discussion of randomness and what makes for effective prediction
Tetlock, P. E., & Gardner, D. (2015). Superforecasting: The Art and Science of Prediction. New York: Crown.	Great discussion of nurturing forecasters, but interesting from a complexity point of view because of how it compares forecasting to rare events and planning horizons.
Watts, D. J. (2011). Everything is Obvious Once You Know the Answer -- How Common Sense Fails Us. New York: Crown Business / Random House.	Conflict between common sense understanding of causation and the workings of complexity
Weisberg, H. I. (2014). Willful Ignorance: The Mismeasure of Uncertainty. New York: Wiley.	Need for analysis to deliberately ignore known salient information.
West, G. (2017). Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies. New York: Penguin.	Ubiquity of log-linear relationships and the reasons for them across a wide scope.

Further Reading on Complexity and Related Topics – Sources Inside the Evaluation Literature Note: There is a very large literature in evaluation on the topic of "systems". This list contains only the sources I like that deal with "complexity"	
Britt, H. (2018). Complexity-Aware Monitoring. https://usaidlearninglab.org/sites/default/files/resource/files/cleared_dn_complexity-aware_monitoring.pdf	Application of complexity in monitoring
Gates, E. F. (2016). Making sense of the emerging conversation in evaluation about systems thinking and complexity science. Evaluation and Program Planning, 59, 62-73	Interdisciplinary literature review on systems and complexity as it relates to evaluation.
Lusseau, D., & Mancini, F. (2019). Income-based variation in Sustainable Development Goal interaction networks. Nature Sustainability, Volume 2(March), 242 - 247	Empirical data on networking effects among the MDGs.
Mintrop, R., Pryor, L., & Ordenes, M. (2018). A complex adaptive system approach to evaluation: application to a pay-for-performance program in the USA. Educational Assessment, Evaluation and Accountability,	Specific application of complexity for a particular program.
Morell, J. A. (2010). Evaluation in the Face of Uncertainty: Anticipating Surprise and Responding to the Inevitable. New York: Guilford.	Jonny's most excellent book on evaluating unintended consequences.
Morell, J.A. Surprises in Programs and their Evaluations http://evaluationuncertainty.com/	Jonny's most excellent blog on complexity in evaluation
Morell, J.A. https://www.youtube.com/channel/UCqRIJjhqmy3ngSB1AF9ZKLg	Jonny's most excellent YouTube channel on complexity in evaluation
Morell, J. A. (2017). From Firefighting to Systematic Action: Toward A Research Agenda for Better Evaluation of Unintended Consequences. Paper presented at the The Unintended Effects of International Cooperation: An academic & policy cross-over conference, The Hague. http://www.ru.nl/publish/pages/814787/morell_2017_10.pdf	Research agenda for work that will improve evaluators' ability to deal with unintended consequences

Further Reading on Complexity and Related Topics – Sources Inside the Evaluation Literature Note: There is a very large literature in evaluation on the topic of "systems". This list contains only the sources I like that deal with "complexity"	
Morell, J. A. (2018). Linking Management and Evaluation: Project Schedules as Program Models. American Journal of Evaluation, 1 - 18.	Much discussion on the use, value, and limitation of models
Morell, J. A. (2019). Revealing Implicit Assumptions: Why, Where, and How? https://www.crs.org/sites/default/files/report_revealing_assumptions.pdf	About implicit assumptions however they arise, but many result from complex behavior
Parunak, H. V. D., & Morell, J. A. (2014). Emergent Consequences: Unexpected Behaviors in a Simple Model to Support Innovation Adoption, Planning, and Evaluation. Paper presented at the Social Computing,Behavioral-Cultural Modeling, and Prediction 7th International Conference, SBP 2014 , April 1-4, 2014, Washington, DC, USA.	Use of agent based modeling in conjunction with traditional evaluation logic modeling and program theory methods.
Patton, M. Q. (2011). Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation New York: Guilford Press.	Insightful discussion of how complexity plays a part in the evolution and adaptability of programs.
Ton, G.; Mayne, J.; Delahais, T.; Morell, J.; Befani, B.; Apgar, M. and O'Flynn, P. (2019) Contribution Analysis and Estimating the Size of Effects: Can We Reconcile the Possible with the Impossible?	Part of a discussion of contribution analysis is a discussion of squaring the methodology with complex behavior.
Walton, M. (2014). Applying complexity theory: A review to inform evaluation design. Evaluation and Program Planning	Literature review outside evaluation with explanation of application to evaluation
Williams, B., & Imam, I. (2007). Systems Concepts in Evaluation. Point Reyes, CA: EdgePress of Inverness.	Review of of systems in evaluation, some of which deals with complexity
Walton, M. (2016). Expert views on applying complexity theory in evaluation: Opportunities and barriers. Evaluation,	Interviews with people who have experience applying systems and complexity in evaluation
Wolf-Branigan, M. (2013). Using Complexity Theory for Research and	Social work perspective. Examples of useful familiar