4.669…
Evaluation & Planning

Deriving Value from Models: Appreciating Partial Truth, Simplification, and Multiple Uses

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

Two routes to slide deck
www.jamorell.com → blog → Program Theory and Program Logic → Deriving Value from Models: .Appreciating Partial Truth, Simplification, and Multiple Uses
http://jamorell.com/documents/Models presentation_AEA_2019.pdf

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

Respect data.
Trust judgement.

A few important concepts about models

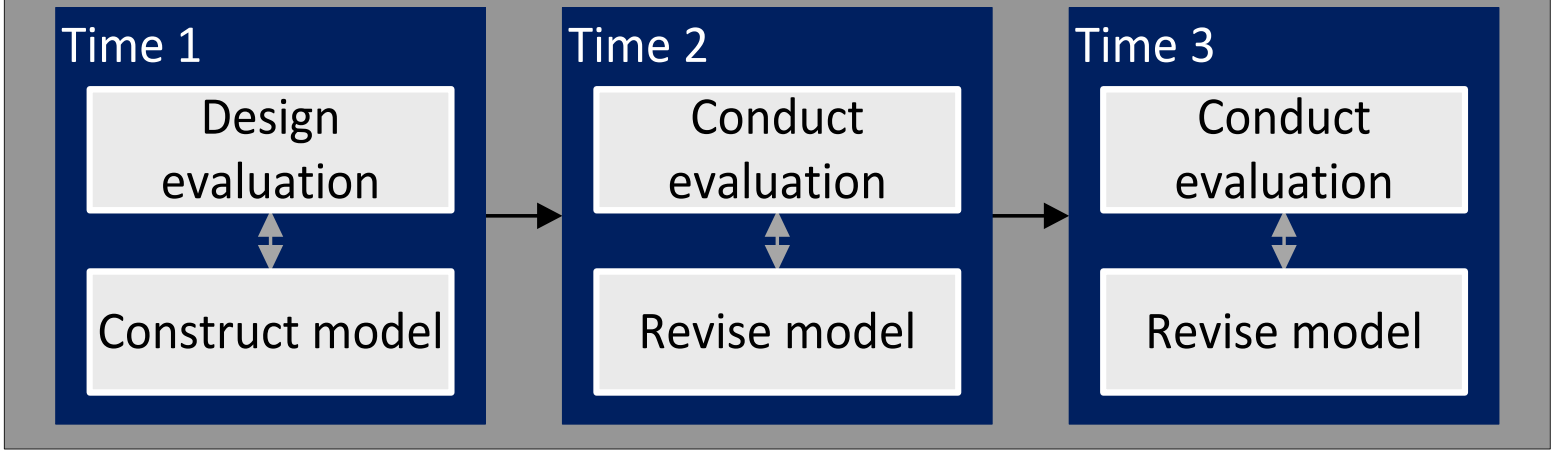
- All wrong but some are useful
- Starting with a model known to be false can be valuable
- Prediction is different from explanation
- The act of building a model has value
- Models can be help think differently
- Helping to think differently
- If systems are complex, traditional models may not depict reality
- Model content differs when used for prediction, explanation, or advocacy

Topics to cover depending on time and interest

- Change over time
- Need for modesty
- Local and global correct
- Information we could putt in but don't
- Nested models
- Emergent behavior
- Competing outcome models
- Sensitive dependence and attractors
- Recasting program theory into alternate frameworks (the case of evolutionary biology)

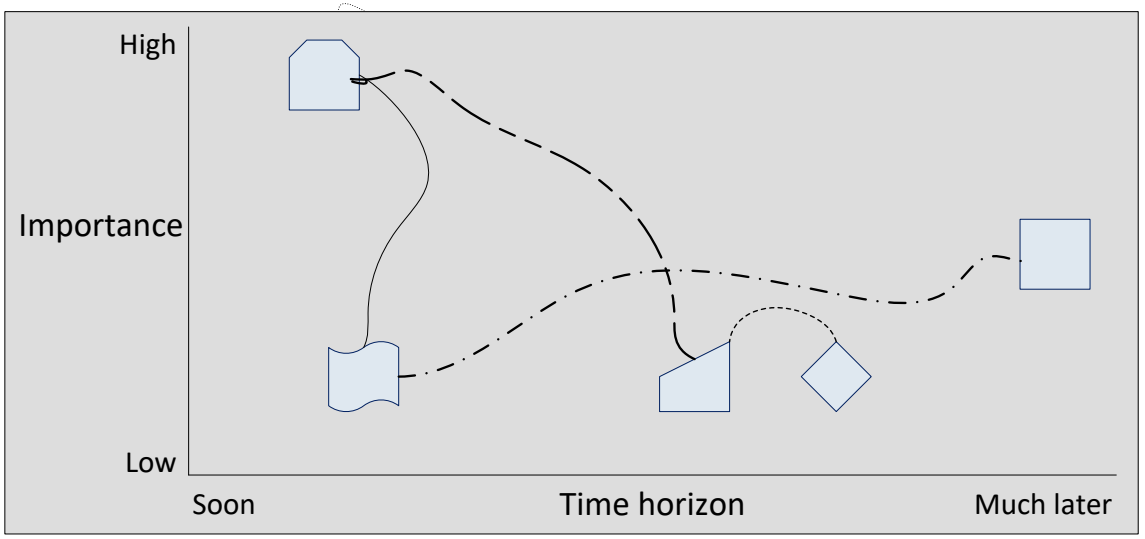
Data may show that the model is wrong or has changed.

Systematic iteration between data and model can 1) maximize time to adjust methodology and 2) keep program theory relevant.



Program goals (and relevant models) may change over time

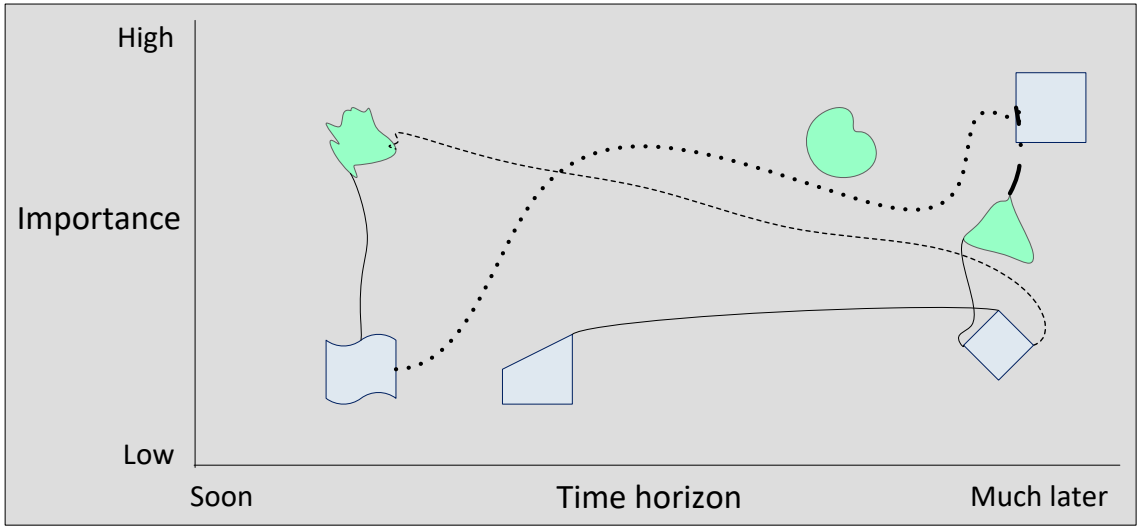
Now



Goals that drive models change

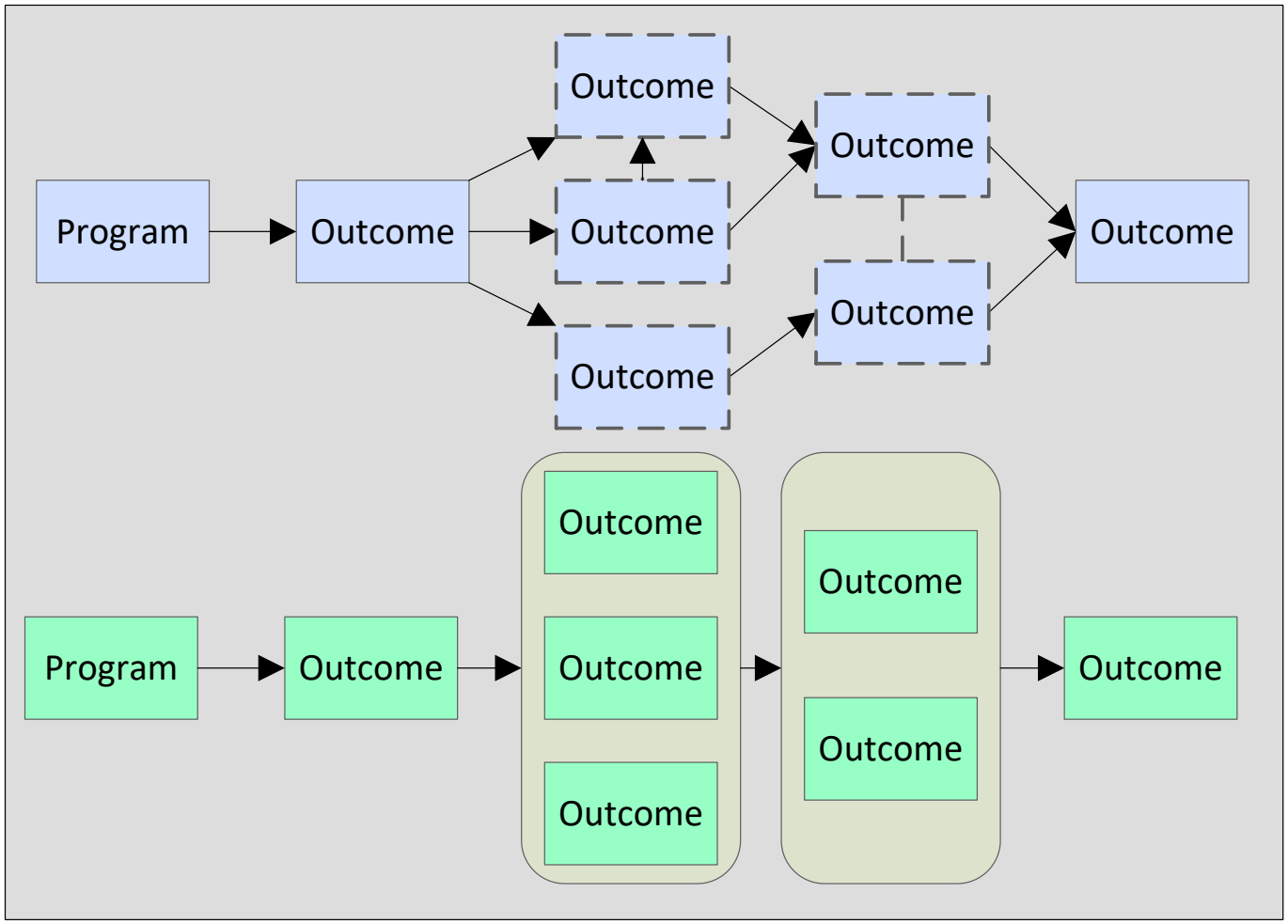
- Content
- Priorities
- Connections

Later



How much do we really know about a program?

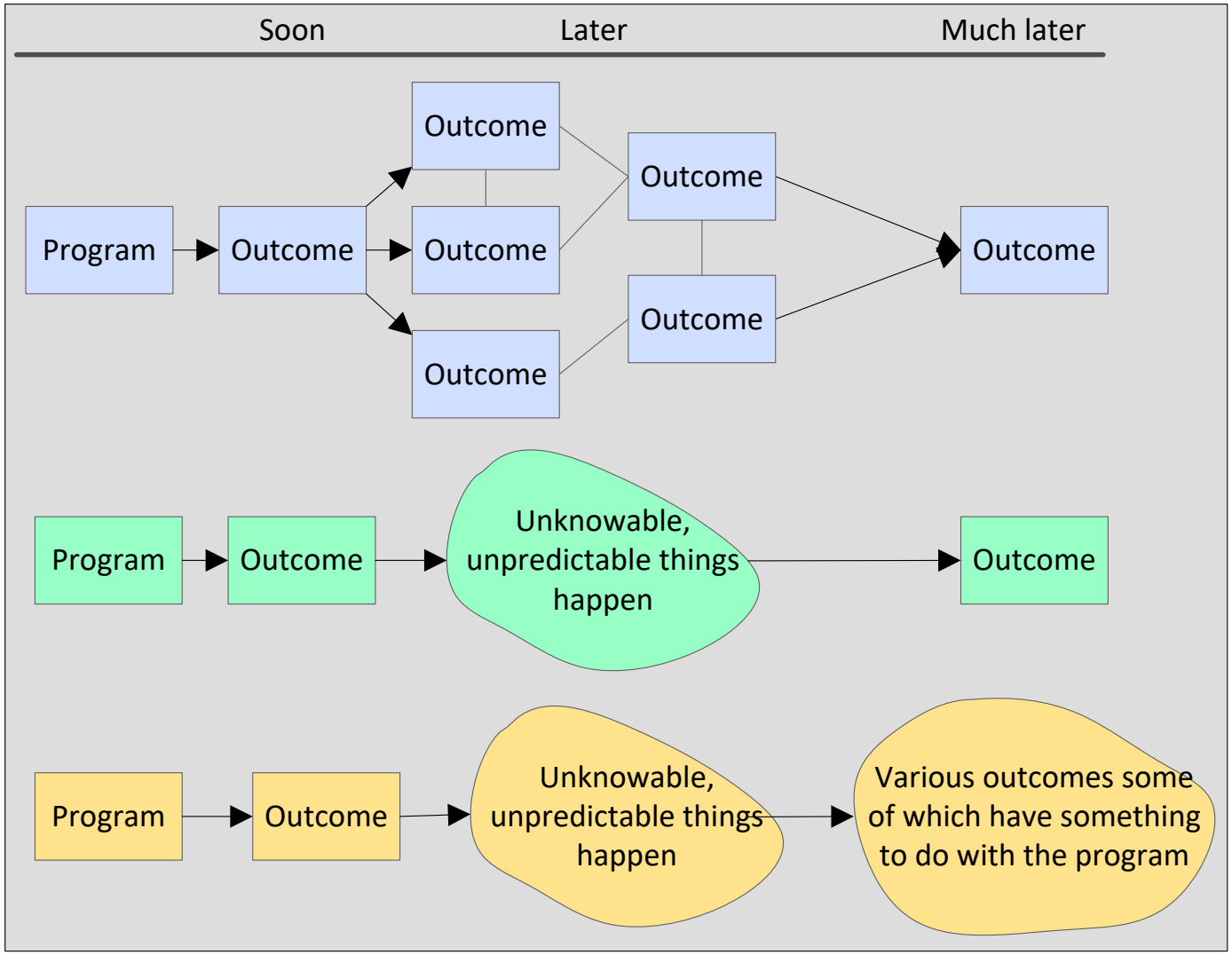
Maybe some modesty is called for.



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

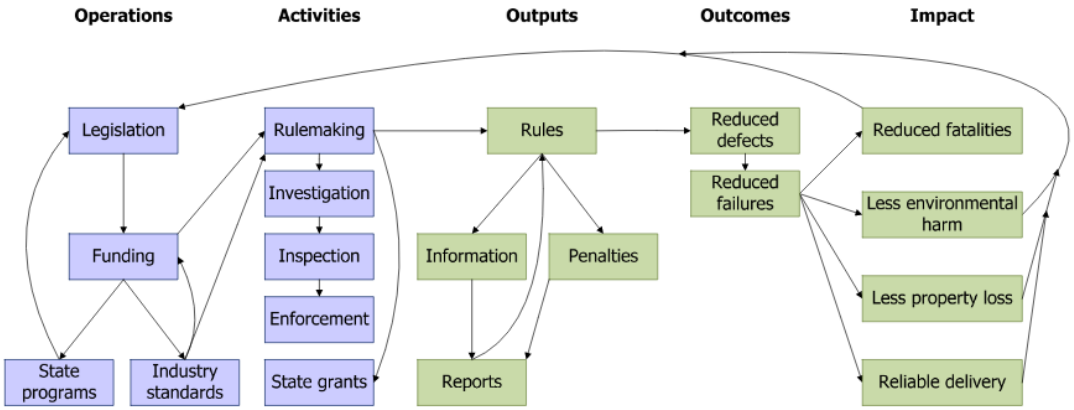
Three models for the same program.

- 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?



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.

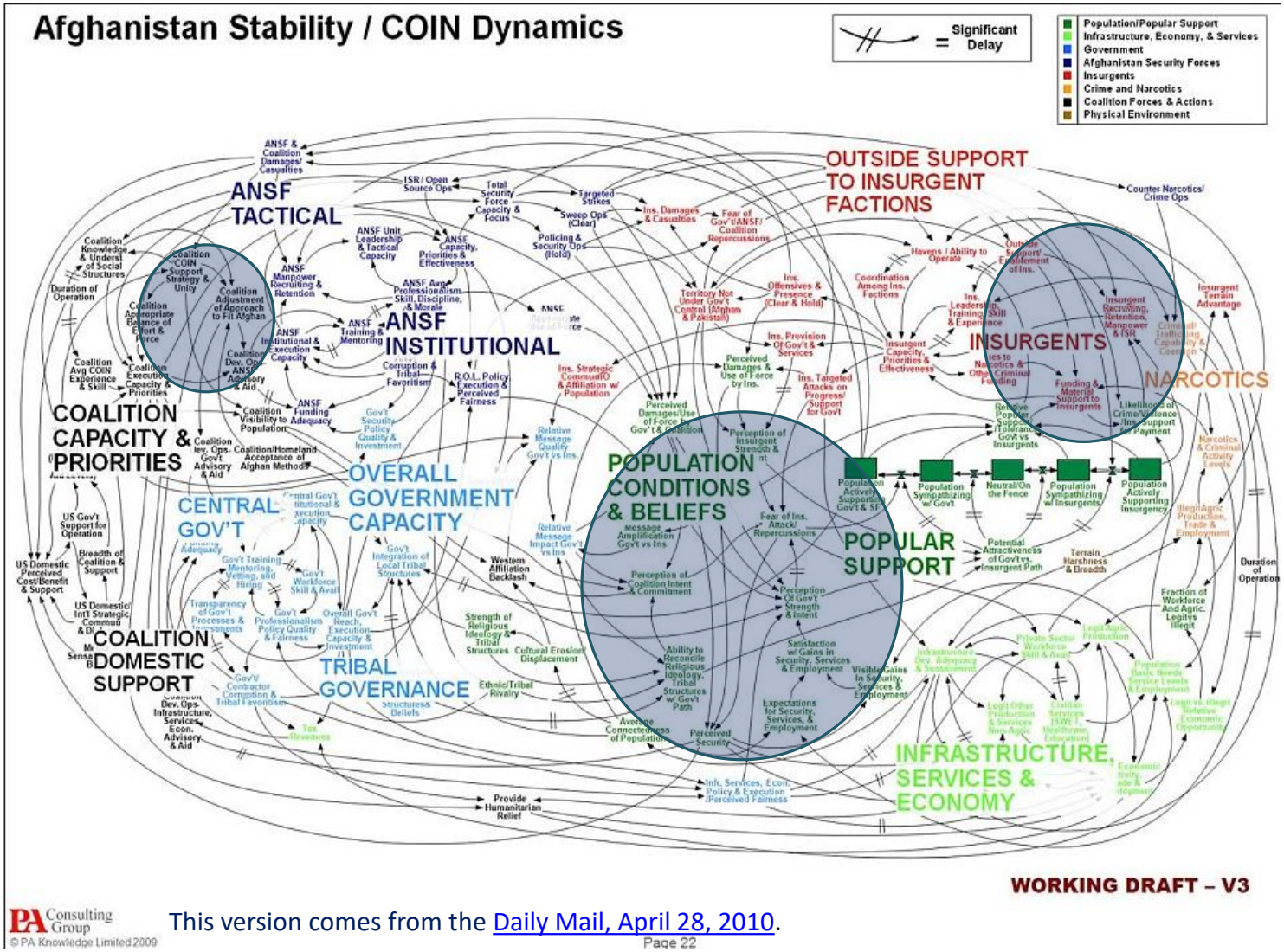
Inputs	Throughput	Outputs	Outcomes	Long Term
Legislation	Rulemaking	Rules	Reduced defects	Reduced fatalities
Funding	Inspection	Reports	Reduced failures	Less environmental harm
Industry	Enforcement	Penalties		Less property loss
Industry standards	Investigation	Information		Reliable delivery
State programs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

Things will happen here.

Models can be locally correct but not globally correct.

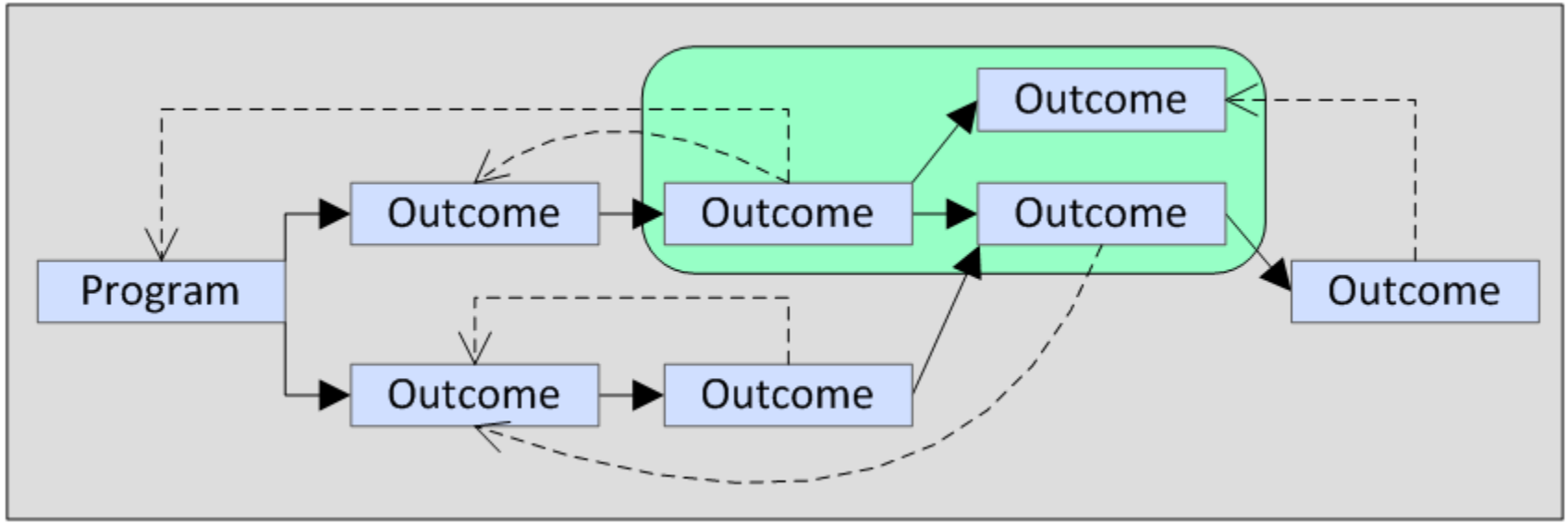
- Emergent behavior
- Adaptive network
- Sensitive dependence

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



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

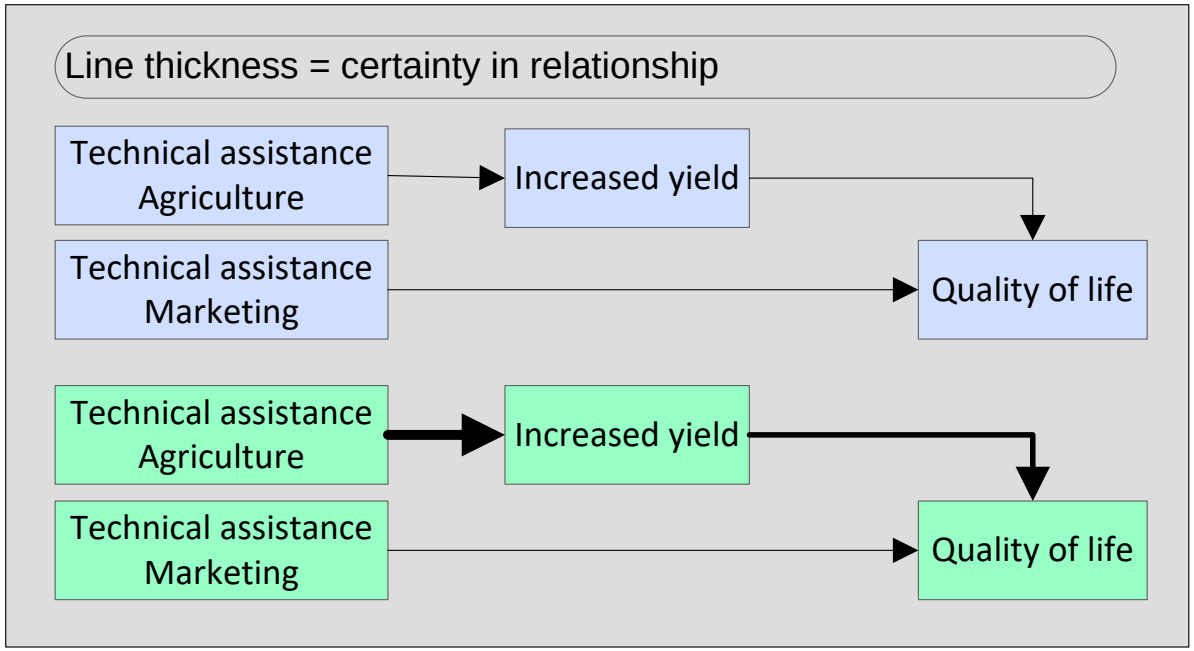
Because of feedback loops and sensitive dependence, a program theory may be everywhere locally correct, but never globally correct.



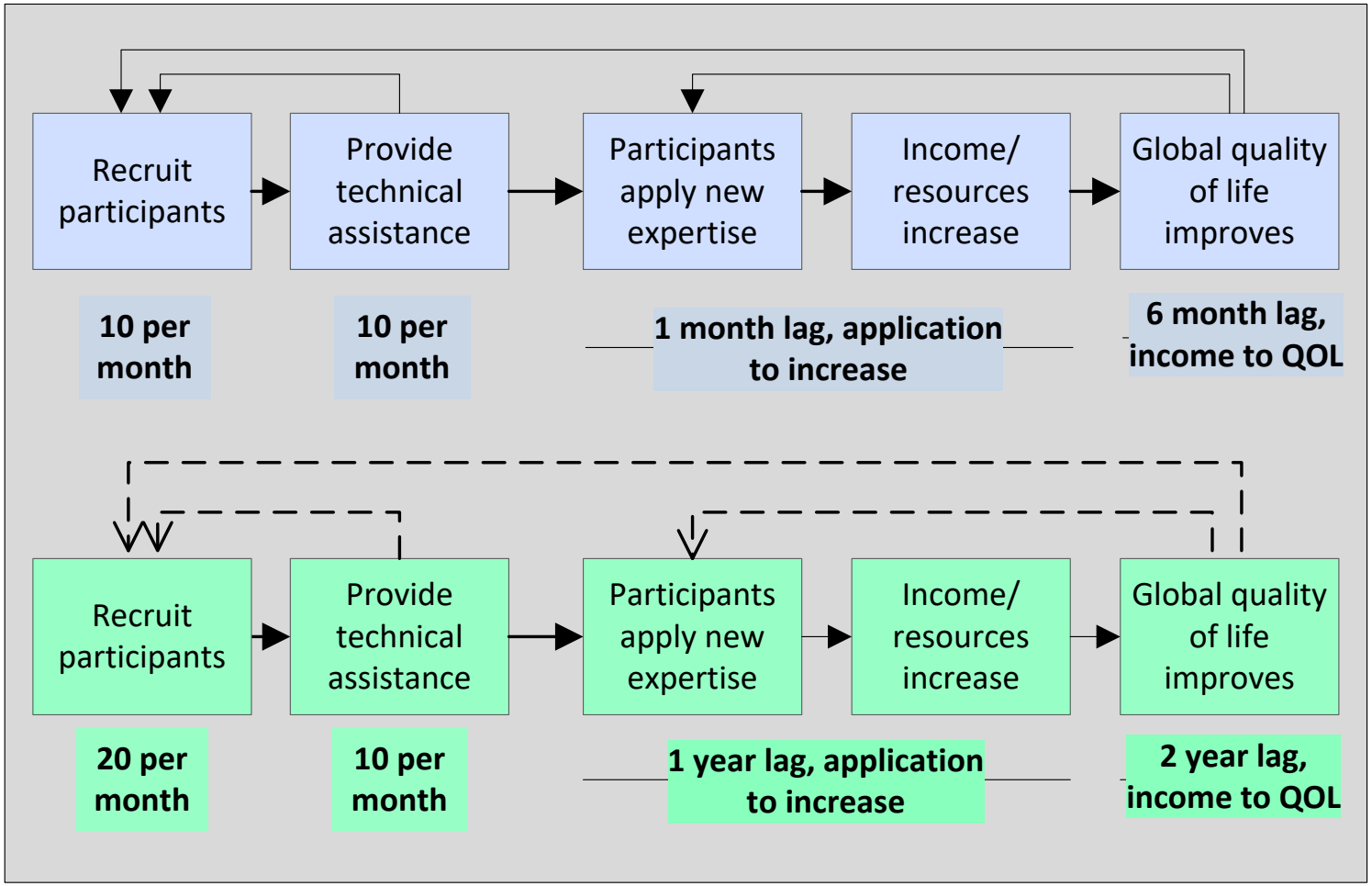
No matter how many specific outcomes stakeholders can specify, they cannot understand long term impact by combining the impacts they are sure of.

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

How sure are we of relationships that we depict?



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

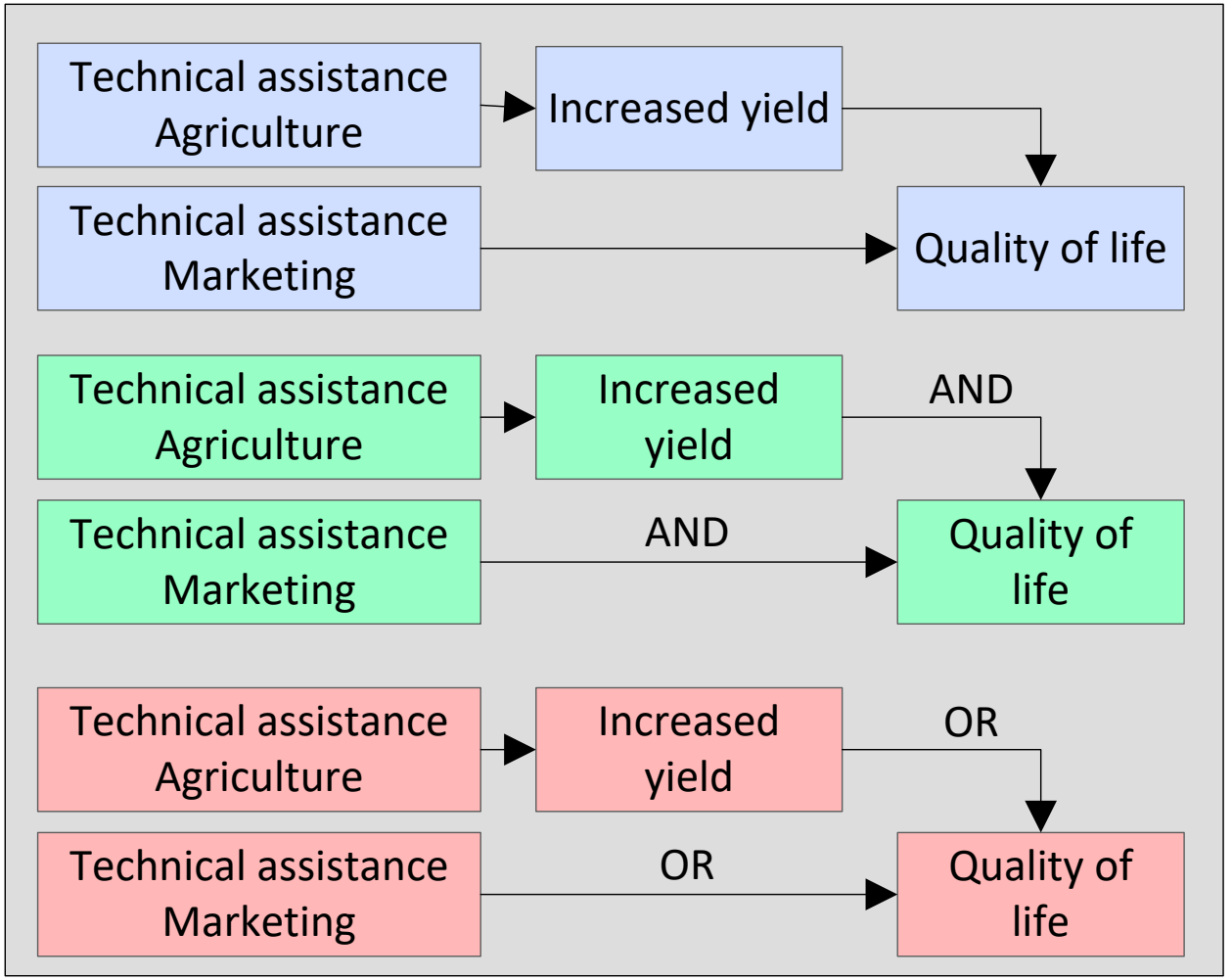


It is entirely possible that the different latencies of these feedback loops will result in very different performance of the same logic.

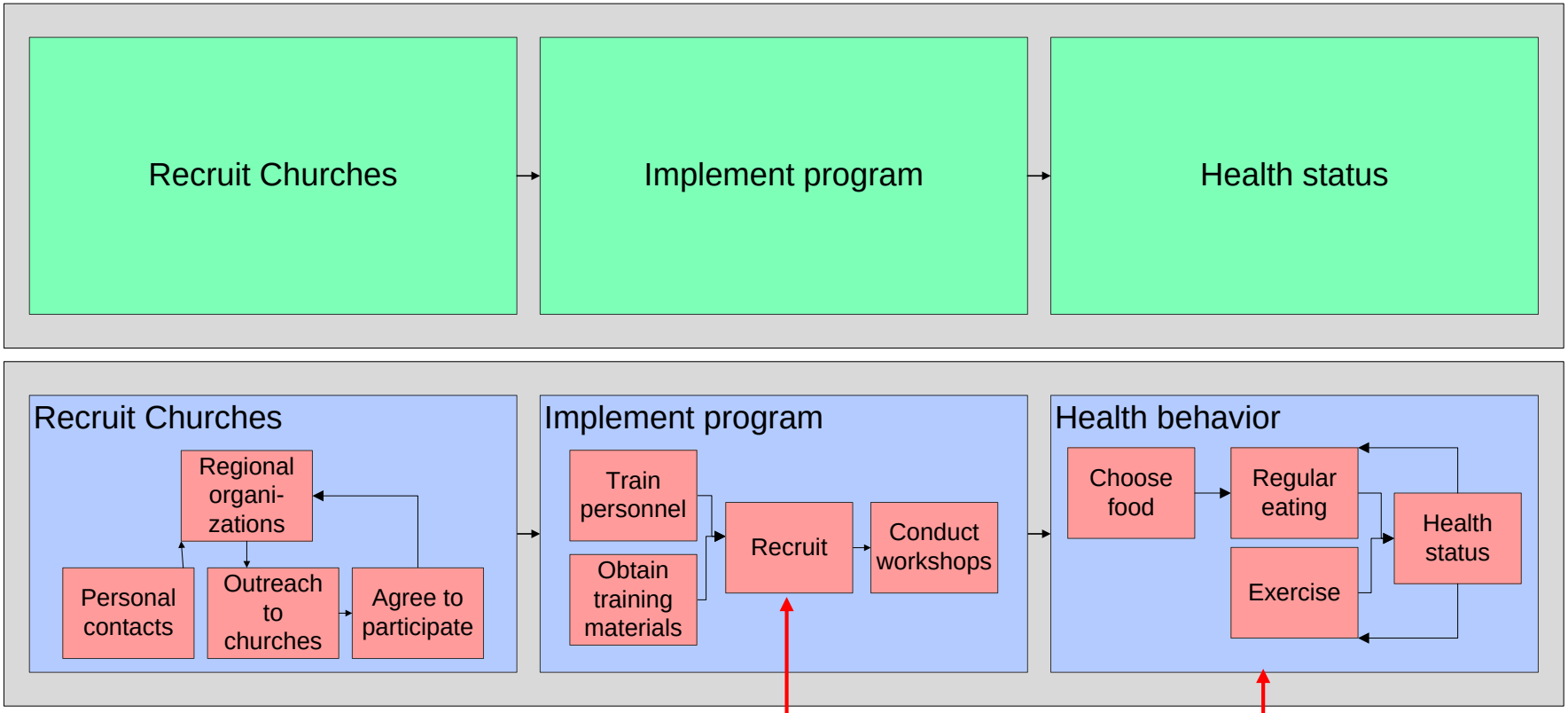
Shouldn’t we care about “and/or” relationships?

This program will probably fail

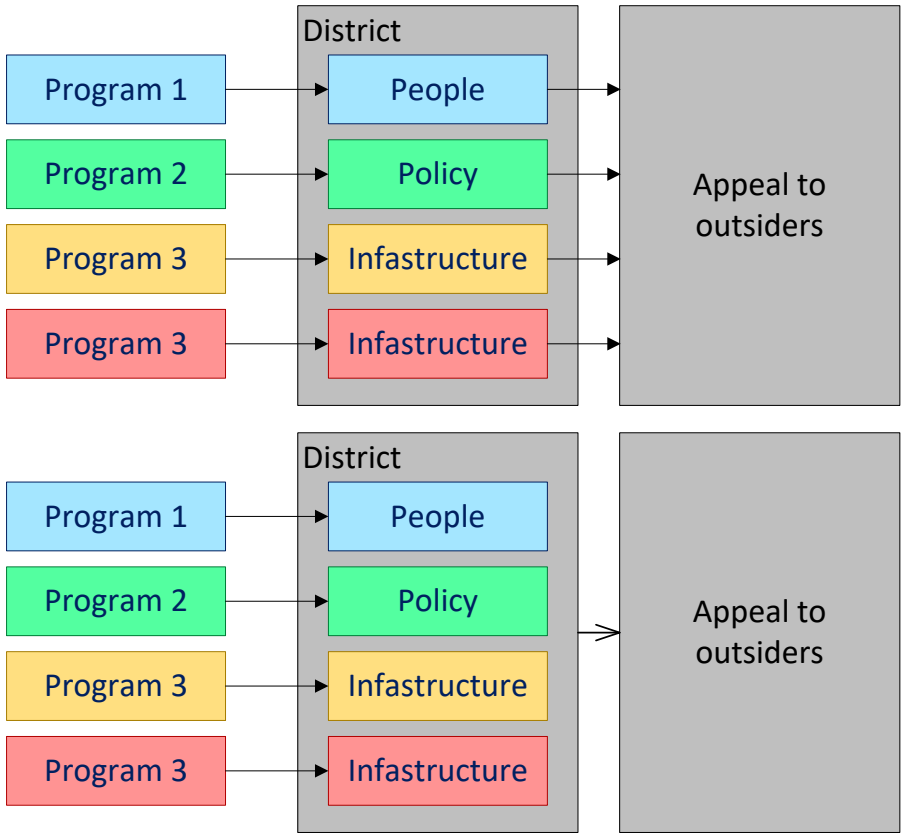
This program stands a chance



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



Emergent behavior may preclude decomposing reasons for an effect.



Why might we not be able to test this model?

1. We are not good enough methodologists
2. It is impossible because the effect is an emergent consequence of its inputs

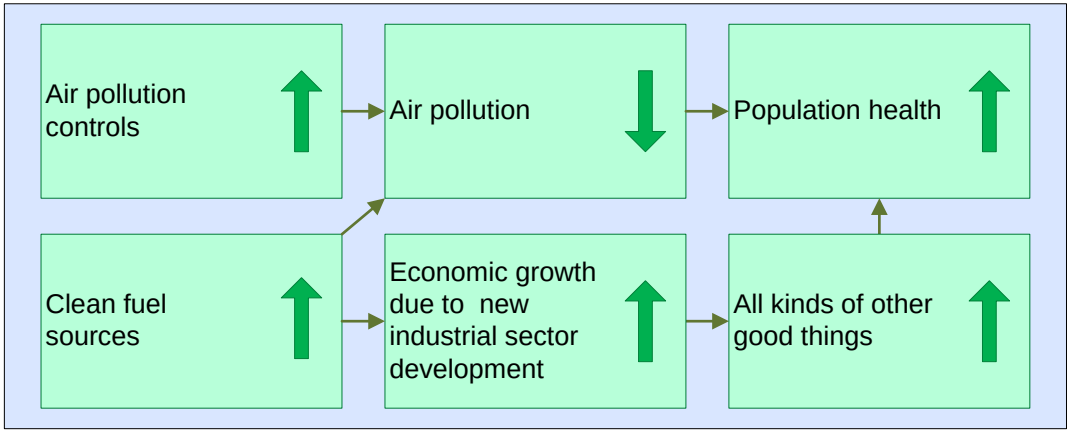
It may not be possible to predict which competing program theory will be correct

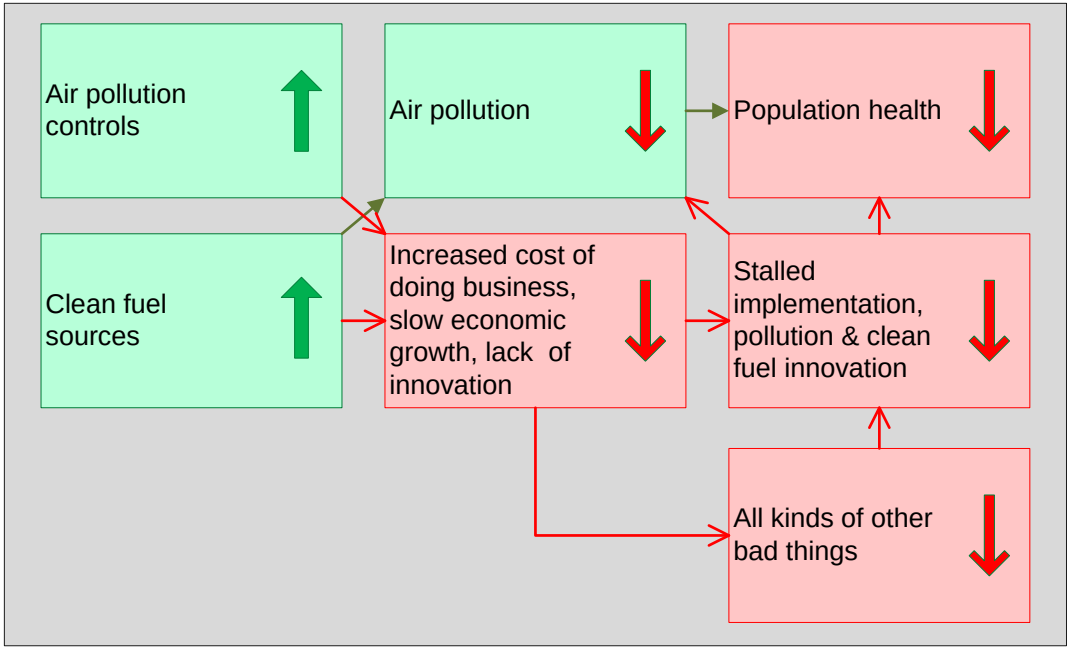
How many factors

- 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?





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

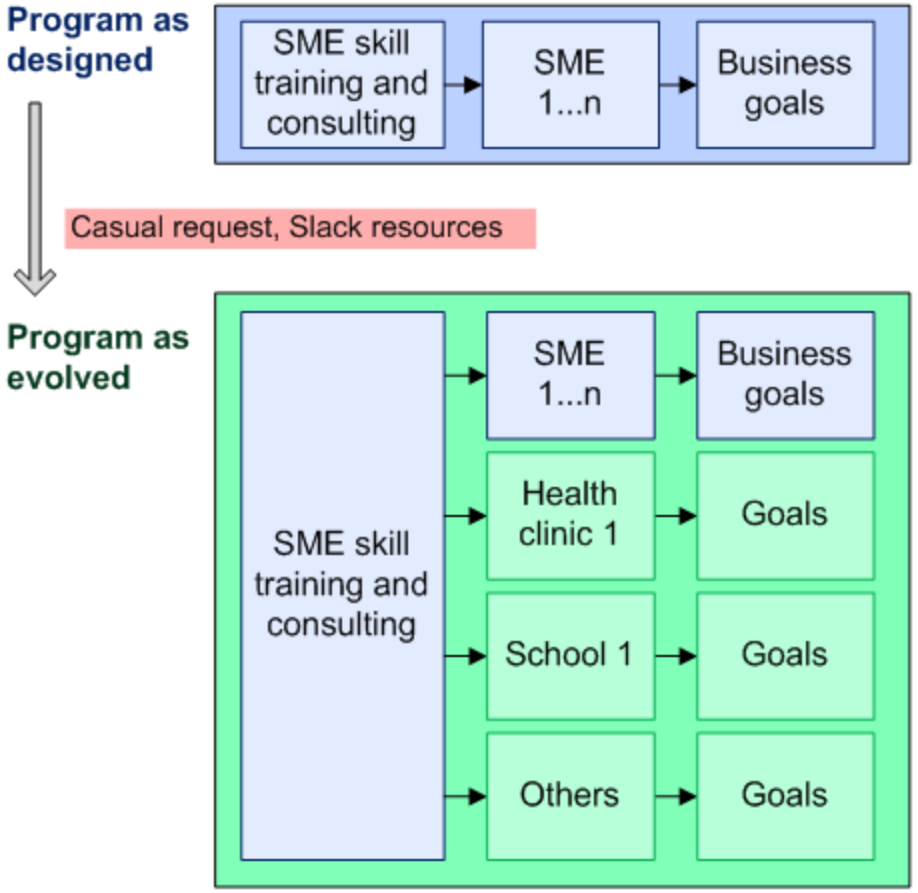
Example 1: historical accident where conditions line and result in program evolution

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



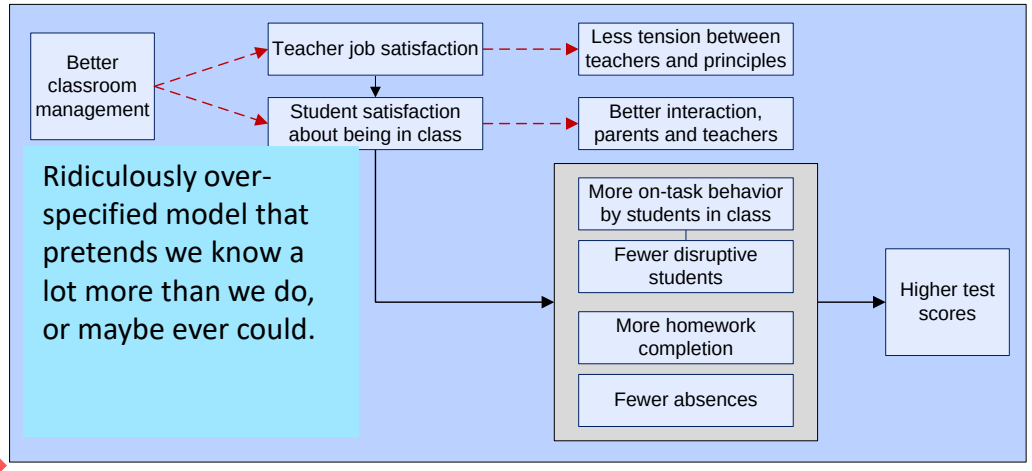
Because of sensitive dependence and attractor behavior, it may be impossible to specify an outcome chain even if the outcome can be predicted

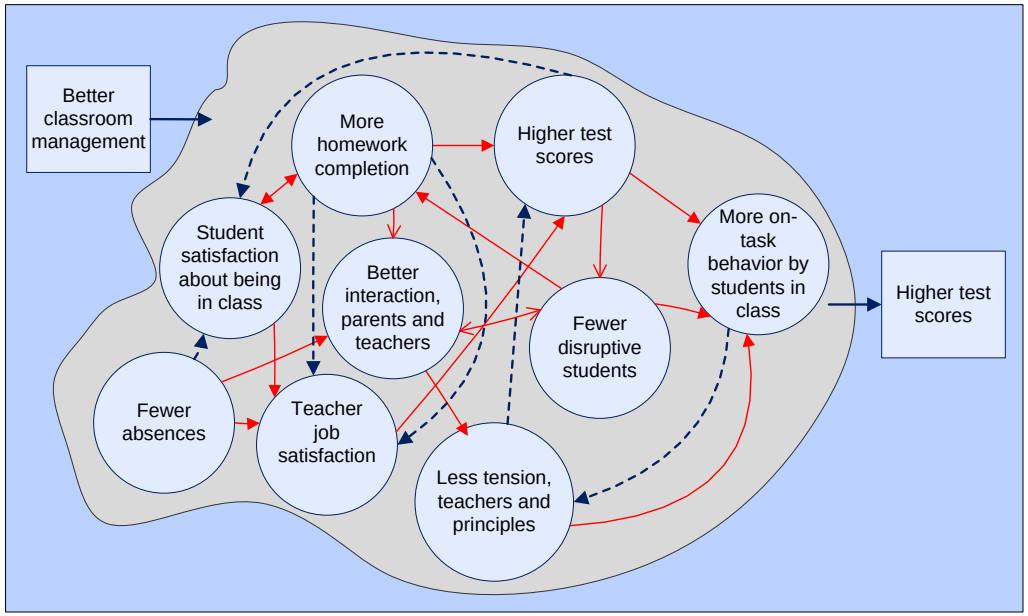
We and our stakeholder are comfortable with models like this. We commit to them with ease.

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?





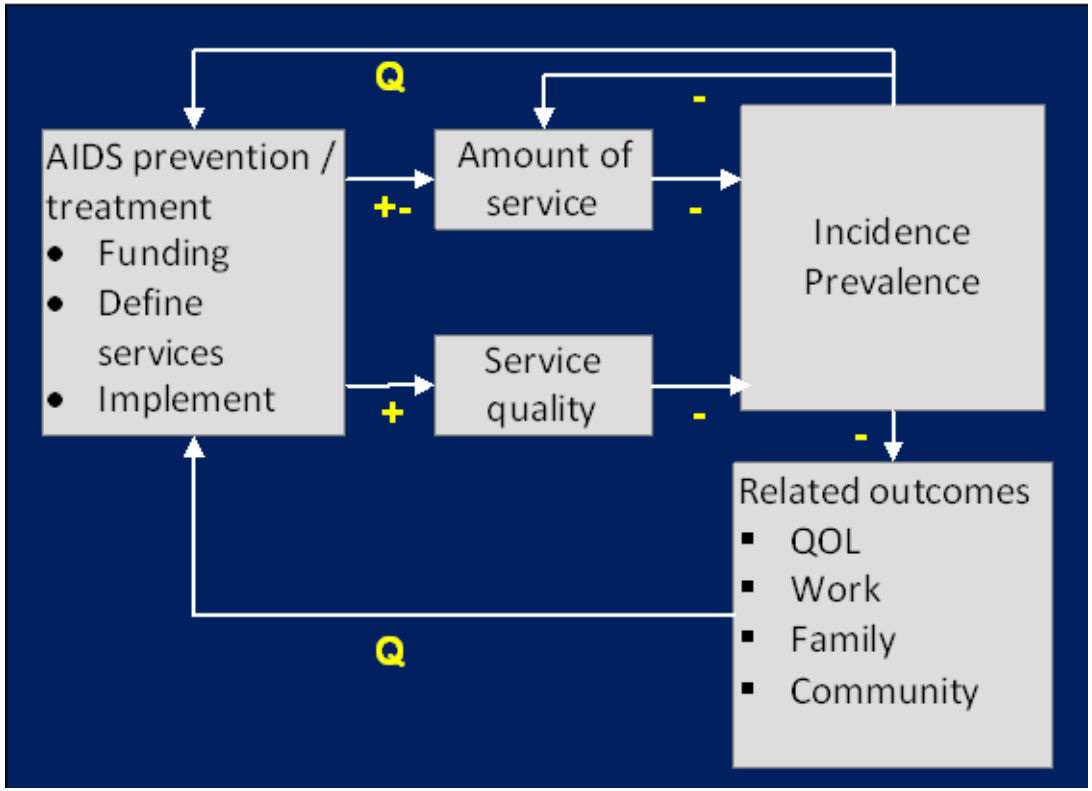
Thinking of theories of change in terms of evolutionary biology and ecology

A nice, traditional, comfortable model.

All outcomes are highly correlated

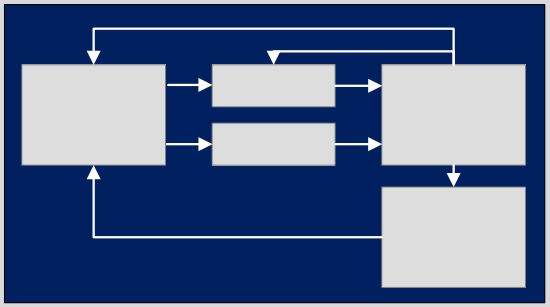
This is a fine program theory. I'd love a chance to do this work.

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



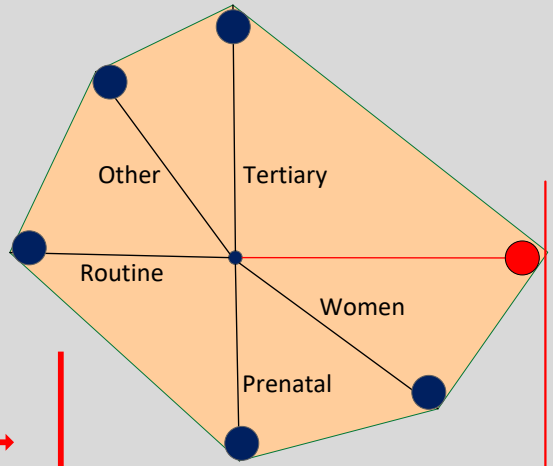
Neither program theory is inherently good or bad.

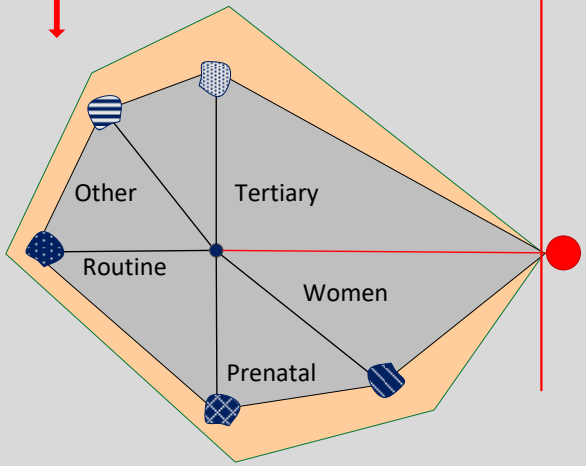
Outcome Maximization Program Theory



	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?

Some interesting reading about models. Not comprehensive, just what I happen to like and have been reading lately.

The Future of Everything: The Science of Prediction Orrell, D. (2007). New York: Thunder's Mouth Press.	Explanation of the inherent problems of using models for prediction across a wide range of activity – weather, health, and more
Linking Management and Evaluation: Project Schedules as Program Models. Morell, J. A. (2018). American Journal of Evaluation, 1 - 18.	Much discussion on the use, value, and limitation of models
Models in Science. The Stanford Encyclopedia of Philosophy. Frigg, R., & Hartmann, S. (2018).	Deep dive into the nature of models.
Purposeful Program Theory: Effective Use of Theories of Change and Logic Models Patricia J. Rogers and Sue C. Funnell	Best book I knows on this topic coming from within the field of evaluation
Revealing Implicit Assumptions: Why, Where, and How? Morell, J. A. (2019).	About implicit assumptions, with an emphasis on depicting assumptions by using models
Self-organised criticality—what it is and what it isn't Roman Frigg Stud. Hist. Phil. Sci. 34 (2003) 613–632	More about self-organized criticality than you ever want to know, but Section 5 is a very perceptive discussion of the use of models.
Weisberg, H. I. (2014). Willful Ignorance: The Mismeasure of Uncertainty. New York: Wiley.	Need for analysis to deliberately ignore known salient information.