4.669...
Evaluation & Planning

# The Logic in Logic Models Part 1: Extending Models in the Directions of Less, and More Specificity

Jonny Morell
jamorell@jamorell.com

This blog post is one of a series I hope to post about how we build models.[1] I'm not sure how many I will get to, but here is the first.

## What do evaluation models look like?

In what follows I will use the term "model" rather than "logic model" because "logic models" are really "models" as the term is used in the process of inquiry. "Models" have a deep history in epistemology, and my inclination is to orient evaluators toward that literature.[2]

Almost every model I have ever seen is made up of one of two types of logic. Those that:

1- establish specific relationships among two or more constructs (green in Figure One), or
2- admit uncertainty and do not include specifics (blue in Figure One).

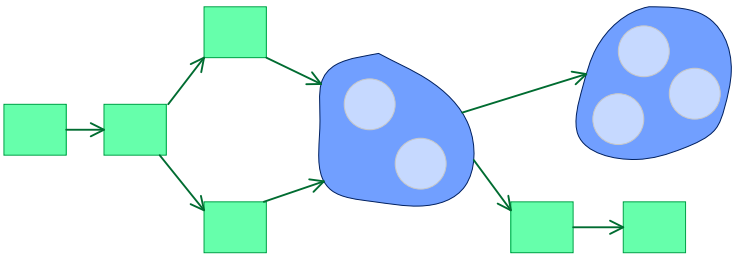

Figure 1: Combined Version of Common Models

I know the picture is trivially simple, but it does contain the essence of both logical forms. A more familiar view of the blue region is the Kellogg model (Figure 2). The logic is the same, "Do some things here and some things there will change".

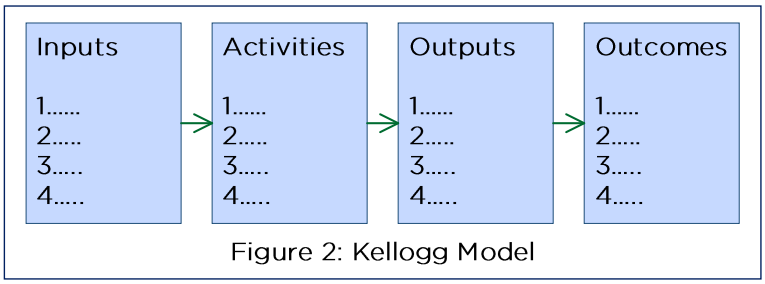


Figure 2: Kellogg Model

---

[1] This series is drawn from a report I did for Catholic Relief Services: *Revealing Implicit Assumptions: Why, Where, and How?* https://www.crs.org/sites/default/files/report_revealing_assumptions.pdf

[2] Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer & G. N. Wilkinson (Eds.), *Robustness in Statistics* (pp. 201–236). Academic Press.
Frigg, R., & Hartmann, S. (2018). Models in Science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford University. https://plato.stanford.edu/archives/sum2018/entries/models-science/
Rogers, K. (2012). Scientific modeling. *Encyclopædia Britannica*. https://www.britannica.com/science/scientific-modeling

## Why should we shy away from detail in our models?

There are good reasons to be biased in favor of sparse models.

Modesty: One reason to build a model is to explain how the program works and what outcomes it produces. Maybe it's just my personality, but I think we usually know less than we think we do about these topics.

Evaluating the noise: Focusing an evaluation on a part of a model is a statement to the effect that that part of the model matters, that it's worth testing, that it is worth the expenditure of evaluation resources. But evaluation findings can be wrong. It seems quite plausible to me that an evaluation will incorrectly find meaning in a part of a model, when in fact there is none. Why tempt fate when modesty should have given us a less elaborate model in the first place?

Evaluation design: The more elaborate the model, the more difficult (and expensive) it is to develop an evaluation that will test the model. This is not a matter of methodology, whether qualitative, quantitative, or mixed. The bigger the evaluation, the harder it is to do well.

Replication in diverse settings: If we can show that a program works, we want to replicate it in other settings. Doing evaluation that will guide replication is a major challenge, which is why Realist Evaluation is so important. One small part of the replication challenge rests on the models that drove the evaluation. The more specific the model, the lower the chances of it working in diverse settings. If we have good reason for believing that the detail matters, we should most certainly include it. But I am convinced that our "good reasons" are often not that good.

Modesty, noise, design and replication all point to a bias toward sparse models. But sometimes we have defensible arguments for including specifics. What then?

## What if we have confidence in the specifics?

I think we trick ourselves into believing that if we are confident about relationships between all the adjacent elements in a model, then we can be confident in the model as a whole. But think about the models that are so common in evaluation literature. They are replete with 1:1, 1:many, many:1 and many:many relationships, both as forward directed and feedback relationships.

Models like this generally derive from collaboration and consensus among a group of stakeholders and evaluators. I can't speak for how other people develop these models, but when I do it, the process usually involves building up the model from dyadic relationships. Many elements are considered, two become the focus of discussion, and a consensus is reached as to whether they should be connected. Then on to the next set of connections.

A problem with these kinds of models is that while any single connection (or set of connections in a limited part of the model) may make sense, the model as a whole may not.[3] All of those elements, all of those connections among those elements, all of those feedback loops, all of those proximate and distal outcomes – all put together in one big model. Can that model be an accurate portrayal of the program we are evaluating? Of course it could be, especially if it is based on prior experience with similar programs.

But then again, we might have created a Frankenstein's monster. As a minimum, it's worth our while to check. How to do that check? I favor a three-step activity.

1- Develop the model in the usual way. This is a good start because so many people are familiar with the process.
2- After the whole model is developed ask its developers to take a step back, consider the entire model, and question whether it truly depicts reality.
3- Once they say "no" (which I am sure they will), take another pass. If nothing else, this exercise will force them into some system thinking.

I suspect that model redesign will feature:

1- Many more blue regions in Figure One and many fewer green regions. In other words, more admission of ignorance.
2- Fewer elements or connections among elements because we (well, I at any rate) tend to over-specify models. And over-specification increases the possibility of getting it wrong.

The simplified model will:

1 give us a better sense of what we do and do not know about our program, and
2- help construct an evaluation that will allow us to decompose those unspecified regions into their components.

[3] The reason why small domains of a model may be correct but the model as a whole many not be is rooted in the behaviors of complex systems, but the details are beyond the point I want to make in this blog post.

## If we believe the specifics, get more specific.

When we develop models we should take another step. We should specify AND OR relationships. See Figure 3.

The green version is our usual way of doing business. We do not announce our intentions, but my sense is that we are implicitly assuming "AND" relationships. The blue region is an explicit version of the green. It is immediately obvious that the program will almost certainly fail because unless all parts of the model work, the program cannot succeed.

Now consider the red version. Because the program has many paths to success, the probability of getting the desired outcomes is high.

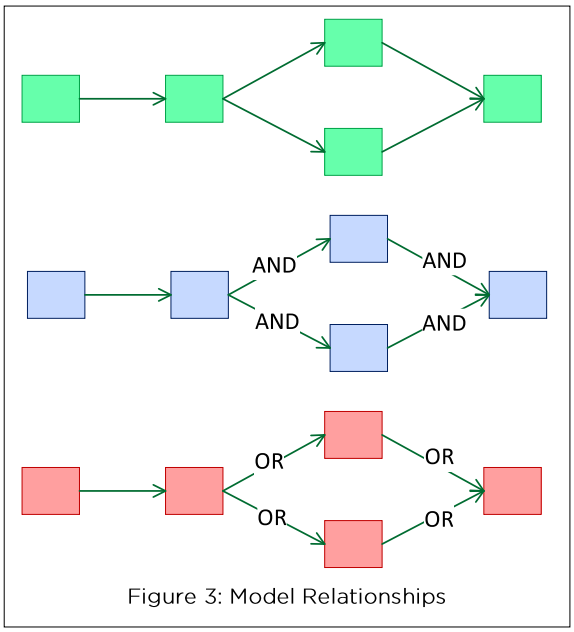

Figure 3: Model Relationships

Specifying AND or OR relationships is difficult because doing so requires a level of knowledge about a program that I don't think too many people have. But I do think that when we develop models, we should ask a few questions about whatever relationships we specify.

1- Can we specify *any* AND, or OR relationships?
2- Are there any parts of the model where such specification is particularly important?
3- Can we design an evaluation in which, as the evaluation proceeds, we can get data that will enlighten us on the ANDs and ORs?

## Overall Message

We know less about the details of how our programs work than we often think we do. Our models should reflect our ignorance. Sometimes we do know the details. When we do, we should specify AND OR relationships. We can think of models as containing a continuum that spans what we know about a program. We would do well to extend that continuum at both ends – toward less understanding of program operation, and toward more understanding. The result would give us models that would allow us to do more realistic planning and better evaluation.