

System Behavior in the Programs We Evaluate: Implications for Theory, Measurement, and Methodology

Panel Presented at the 2003 Meeting of the American Evaluation Association

Furthering Systems Perspectives in Evaluation

This presentation is part of a series of workshops and panels at AEA 03 designed to further the application of systems perspectives in evaluation. As a post-conference follow-up to these activities, we have developed a listserv for ongoing conversation. To join, go to <http://evaluation.wmich.edu/archives/index.html>, and sign up to EVAL-SYS

Panel Members

Jane Davidson, discussant/moderator Jane.Davidson@wmich.edu

Jonathan A. Morell presenter/panel organizer jonny.morell@altarum.org

Patricia Rogers presenter patricia.rogers@rmit.edu.au

Bob Williams presenter bobwill@actrix.co.nz

Panel Description

There seems to be a growing sense among evaluators that while logic modeling and program theory are the foundation of good evaluation, our efforts in these areas are unsatisfying. No matter how extensive the model building exercise, and no matter how many stakeholders are involved, and no matter how much previous experience and data are brought to bear, models don't seem to help as much as they should. If we were to try harder, could we build better models and more powerful program theory? We believe the answer to this question is NO, IF by "trying harder" we mean putting more effort into doing things as we have always done them. We believe the nub of the problem is that the programs we evaluate often exhibit behaviors that are a function of system dynamics. These behaviors are quite different, (often in radical and surprising ways), from the behaviors we expect based on our usual logic models and program theory. Because those theories drive our measures, our methodologies, and our relationships with stakeholders, failure to appreciate system behavior is leading to our weak evaluations. We propose a panel to explore these issues, and point the way to solutions. To achieve these objectives we will combine multiple views of system behavior, an illustrative case study, and heavy audience participation. Our goal is for participants to leave the panel with new and useful insight about how to develop logic models and how to apply those models to make decisions about measurement, methodology, and interaction with stakeholders.

Presentation Abstracts

Jonathan A. Morell

Complex adaptive systems (CAS) consist of numerous autonomous elements who can sense their environment and act independently based on internal decision rules. A great deal of difficult to predict behavior can result from these interactions. Consider three examples. 1) High level ordered behavior can emerge from interactions among agents that are unable to collaborate with each other. 2) Systems can remain impervious to powerful forces, while stable conditions can change quickly and radically in response to seemingly minor perturbation. 3) Depending on their ecological niche, a system may fare better or worse by changing in small, incremental ways, or in radical jumps from its present state to a new state. This presentation will cover these and other

CAS behaviors, and illustrate (using the panel's common case study) how an evaluation can be improved by considering these dynamics when developing logic models.

Patricia Rogers

While logic models usually contain feedback loops, they essentially posit linear relationships among constructs. Interactions, resonance, and dampening effects are absent. Systems, however, are rich with non-linear behavior. This presentation will consider the value of including non-linear reasoning in the systems diagrams we use as evaluators. Would it make a difference? What can we as evaluators do about it? What are the implications for developing our methodologies and our interactions with stakeholders? The panel's common case study will be used to illustrate the application of non-linear models in evaluation.

Bob Williams

Soft systems methodology (SSM) was one of the first systems approaches to be developed specifically to explore complex human systems. Thirty years on, and with thousands of studies under its belt, the basic idea of gaining insights by exploring real life through a series of different but rigorously developed systems lenses remains vital and relevant. This presentation describes the key features of SSM, applies them to the case study and reflects on the implications, advantages and disadvantages of using the methodology in evaluation.

Procedure for the Panel and Post-conference Activity

Time limits preclude in-depth treatment the systems disciplines on which this panel is built. Therefore our approach is to:

- Present a common case from which all presentations will draw.
- Provide examples in each presentation of system behavior that has implications for thinking about logic modeling in ways that evaluators do not usually consider.
- Begin a dialogue among the panel and the audience.
- Establish an Internet based forum through which people can continue the discussion.

The Case: Evaluation of an “Information Prescription” Program in Health Care

Disclaimer:

This case is deliberately constructed for the purposes of illustrating specific aspects of systems perspectives in evaluation. While many of its aspects are based on real programs, the case itself is an amalgam of different programs and plausible but artificial detail. The case is not a faithful description of real world programs.

The Information Prescription Movement

In health care there is a movement afoot to provide specific, trustworthy, relevant information to patients to aid in their health care decision making. The catch phrase is: “the right information, to the right patient, at the right time”, otherwise known as an “information prescription”, Ix. The belief behind this effort is that relevant information will lead to better clinical care and lower costs.

The Evaluation Task

Aardvark Health Care System is a large health care system that is implementing an Ix program. The evaluation task is to determine whether the system has the desired outcomes. National context matters. The program is taking place in the United States. No national health insurance. Multiple (often competing) health care providers and payers – traditional insurance, health maintenance organizations, Medicare, Medicaid. Insurance is paid for as part of employment benefits. Hospitals compete for patients. Physicians bill separately from hospitals, even when in-patient services are provided. Costs going through the roof. Fundamental change in the system will not happen in the near future, so the emphasis is on smaller scale improvement.

Outcomes of interest are: clinical outcome, cost, patient satisfaction, and enrollment levels, and The timeline for the evaluation is three years. Data collection can begin about the time when the Ix program is implemented.

Data and Methodological Options

Aardvark can provide the evaluators with a large amount of historical data on patient characteristics, diagnosis, treatment plans, and outcomes. Evaluators can also get a great deal of historical data, and access to ongoing data post-program implementation. Also, there is good data on how much Aardvark pays to the third party service provider that developed and implemented the system. Aardvark lacks a good activity based costing system, however, so it has no good data on internal labor effort devoted the system. Aardvark does know how many FTEs in their IT department are working with the outside contractor, but that is about it.

The Ix program is implemented simultaneously (or very near it) in all the system's clinics. Thus opportunity for time series analysis across patient and diagnosis groups are rich, but opportunity for control – treatment comparison at the same point in time is nil. Also, there is no opportunity to compare Aardvark to another health care system (aka Aardvark's competition) that has not implemented Ix.

Evaluators can get good access to Aardvark's employees to conduct interviews. Random sampling of respondents is possible. There is also possibility to interview patients, but only with volunteers. In addition, the evaluator can get access to the routine patient satisfaction surveys that Aardvark conducts several times per year.

Doing this evaluation is the task presented to the evaluator. That's the job. That is what Aardvark cares about. As the evaluator does due diligence however, a variety of complications and possibly important contextual issues arise.

Ix Provides Information, but is not a Stand-Alone Activity

The Ix functionality is bundled with other functionality. Aardvark has established a system that works as follows. First, patients register with a password to assure secure communication over the Internet. With this system, the patient can then: 1- discuss health status with nurses, 2- access medical records (for instance, the patient can look at blood pressure readings over the past year), and 3- make appointments.

As part of the discussion, the health care provider may want the patient to read up on some particular topic. For instance, based on weight, physical activity level, family history, cholesterol levels and blood pressure, the patient may be on the verge of entering a high risk category. Prior to a doctor visit, the nurse may want the patient to read about the condition, learn medication options, understand the need to change diet and physical activity etc. To make this happen, the nurse, in an

email message to the patient, can embed a hyperlink that goes to a trusted source of information that was written specifically for patients with this condition.

Scientific Justification for Ix is Sound, but Incomplete

There is defensible research to show that Ix does lower cost and improve outcome with respect to particular conditions. The evaluator realizes, of course that it is one thing to demonstrate this finding in a controlled clinical trial for a particular disease, and quite something else to show that a generalized Ix system, working in a real world setting, has overall beneficial effects. Despite these studies of particular conditions, there is no in-depth, comprehensive review of the literature on Ix. We don't know the circumstances under which benefit shows up, what research can be trusted, and the like.

Theory and Logic Models are Sound but Incomplete

No single logic model is sufficient to explain the effectiveness of Ix because the reasons for its impact depend on specific medical conditions. For instance, in certain situations, adherence to physician's orders tends to a problem, and giving patients information increases people's willingness to follow those orders. (Figure 1).

Another situation is the case where misperceptions about treatment options and outcomes tends to be a problem, and the best evidence based medicine points to less intrusive, and less expensive, options. (Figure 2) Usually, this scenario applies when choices between surgical and medical treatment options are at play.

The evaluator concludes that this mélange of specific logic models can be grouped into three categories:

- Information really can lead to more optimal decisions. The critical factor is having good information at the right time.
- Information can facilitate collaboration and joint decision making (e.g. between doctor and patient), and this has value over and above the actual decision that is made. Essentially, the information serves as a joint frame of reference for interaction among the parties.
- Information allows adjustment of action in light of changing circumstance. Here, the key concepts are: 1- ability to detect change, and 2- length of feedback loops between event and having information about it.

Figure 1: Ix & Adherence

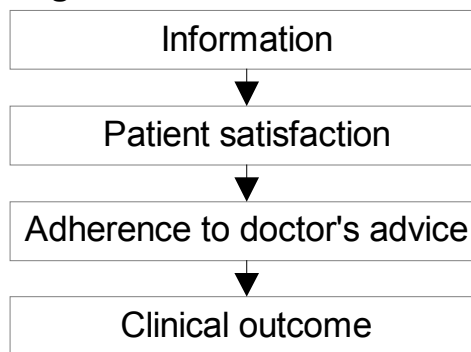


Table 1: Information Provided from Evidence Based Medicine

	<u>Common Belief</u>		<u>Reality</u>
	<u>Option 1</u>	<u>Option 2</u>	<u>Option 2</u>
<u>Risk of no action</u>	High	High	Low
<u>Clinical outcome</u>	High	Low	High
<u>Cost</u>	High	Low	Low
<u>Pain / discomfort</u>	High	Low	Low
<u>Post treatment QOL</u>	Low	High	High

Although there is no hard data on the subject, expert opinion has it that the scenario shown in Figure 1 is important for health promotion and disease prevention programs where there is a long lead time (possibly measured in years) from the beginning of the intervention and observation of clinically significant events. The scenario depicted in Table 1 usually deals with choices between various surgical and medical alternatives.

The evaluator realizes that all involved have a strong inclination to believe that Ix really can improve quality and decrease cost, and that in some cases, this belief is correct. However, the evaluator also knows full well that other possibilities must be considered.

- Ix might improve either cost or quality, and leave the other unchanged.
- Ix might improve cost or quality, and change the other for the worse.
- Ix may have no impact at all.

Knowledge and Information

More is not Always Better

We have to allow for the possibility that more information is not always better. This is particularly true when a course of action is not clear cut. One issue to consider is what it means to have “too much” information, and when trust in expertise and wisdom has to take over from rational decision making.

Specific vs. General Knowledge

Ix is seen in terms of very specific information. It assumes a narrow focus on the “right information to the right patient, at the right time”. But the evaluator realizes that its also important to think in terms of people’s general knowledge of health and related topics. As a simple example, imagine two patients with similar symptoms, both of whom are told they have a viral infection and do not need antibiotics. “A” has a pretty good knowledge about the public health problem of over use of antibiotics, and a general sense that viral and microbial infections are different. “B” knows none of that, but has access to the same quality information as “A” concerning the value of antibiotics for his particular set of symptoms. The evaluator concludes that even though they have the same condition-specific trustworthy information, “A” and “B” might react very differently to the doctor’s advice.

Appreciating Quality

An assumption is that there is in fact good information to provide. This is certainly true in a large number of situations. It is probably true in most of the situations where some good information can make a big difference, e.g. managing diabetes, or diet and life style factors that affect cardiovascular health. On the other hand, there are many important cases where it is not true. For instance, consider what was known about HRT a few years ago. At that time it was not at all obvious what the best course of action was. Most of the advice that made it the newspapers seemed to be something like: “Here is an ambiguous finding, it may affect your decision, talk to your doctor”. In cases like these, using the information requires having not only the results of the findings, but a sense of the scientific value of the information. Obviously, we can’t all be turned into biochemists and epidemiologists. And even if we were, they are unsure as well. Still, specific information in these situations would be more helpful if it were embedded in some general knowledge about methodology. For instance: Long term studies are better than short term studies. Large numbers of participants are better than small numbers of participants. Control groups are good. Random assignment is good. Replication is good. Early results of a study really can be wrong. New revelations should not be trusted, especially if they challenge accepted knowledge. Also, the windows of opportunity for providing specific and general information are different. When someone is in pain and hears about a possible new treatment on the news, that’s not the time to explain why early findings should not be taken at face value.

Social-Psychological Considerations

Information does not stand alone. Its embedded in a context of doctor – patient relationships. For instance, the evaluator posits that if trust between patient and health care provider is low, and a patient gets information from an independent source, distrust of the doctor by the patient will increase. This is because in any given situation, its likely that what the doctor says and the patient reads will not be (or appear to be) totally congruent. The evaluator realizes that the situation is

Non-Clinical Reasons for Ix

In addition to providing patients with information as a clinical tool, there is a motivation to provide information as a marketing tool. Patients like getting relevant information from their doctors, and will choose doctors and health care systems that provide such information. This is the case regardless of whether Ix has clinical consequences or not. Thus health care systems may come to view Ix as a tool to increase brand loyalty and market share. Aardvark shares this belief and has worked hard to be the first to offer a comprehensive Ix service to their members. As a result of their decision, and the time it takes to allocate funds, determine requirements, choose vendors, and implement systems, Aardvark has at least a 12 to 18 month lead time on its competition. For that period of time at least, Aardvark will be the only health system in the community offering an Ix system that is integrated into its other Web based services to members.

Its not unreasonable to assume that regardless of clinical value, health care systems will come to view Ix systems as a “must have”. This attitude might serve to drive up the overall costs of health care even though costs for particular conditions may go down. The reason is that because even if there are circumstances where Ix can drive costs down, those situations are probably constrained. And in any case, for the cost savings to show, there is a need for very carefully crafted Ix programs. Counterbalancing the benefits is the drive to provide good information to patients in order to compete for market share. If enough of this happens, a lot of patients will get a lot better information, BUT, the better information will not be the kind that could actually lead to cost justifiable outcomes. We then end up with health care systems implementing expensive Ix programs as a “must have” to keep patients, even though the systems will serve only to drive up costs.

On the other hand, It is by no means certain that Ix will come be seen as a “must have”. Aardvark’s competitors know there are very good reasons not to implement such systems. First, their contribution to lower cost and higher quality is by no means certain. Second, they are very expensive. Third, Ix systems have interfaces to numerous other information systems (e.g. scheduling, clinical records, accounting). Thus an Ix system can’t simply be dropped in to an existing IT infrastructure without significant work done on the other components of that infrastructure. Finally, using Ix systems to good advantage requires difficult changes in organizational process. For instance, the way in which physicians and nurses deal with patients’ calls will change. Money notwithstanding, process change can be difficult and painful.

Ix May Have Desirable and Undesirable Consequences for Different Parts of the Health Care System

From an organizational point of view, whether or not Ix is desirable depends on the incentive system under which an organization is operating. Capitated, and fixed fee services may benefit from Ix because their income is fixed, and financial viability comes from lowering costs while providing desirable services to customers (aka patients). On the other hand, if Ix works, it will decrease hospital admissions and decrease expensive medical and surgical treatments. Such a decrease will lower income for hospitals and other fee for service elements of the health care system.

Industrial Structure

It is more than a service to patients. It represents a rich, and interlinked a set of businesses connected by strategic partner, and buyer/supplier relationships. For instance where does the information come from that Aardvark supplies to its patients? Not from Aardvark's physicians and nurses, who have nothing better to do than compose articles on health care. Rather, there are several specialty companies in the business of generating those articles and selling them to others. To complicate matters even further, it is entirely possible that Aardvark is not a direct customer of one of those specialty companies. Rather, there is another set of companies that specialize in doctor/patient/clinic management software. That company is Aardvark's prime contractor, and the health information company is a supplier to that company, i.e. it is a second tier supplier to Aardvark. And, there is a lot of competition among all these companies for Aardvark's business. probably much more complicated, and that a literature review on the topic may be worth the effort.

Communication Among Consumers

Aardvark Health Systems operates in a community that has rich channels of communication among consumers of health care services. Several community groups are dedicated to improving the quality of health care in the region. They meet periodically, publish newsletters, and exchange information about health care issues and services via a listserv. Members of these groups include patients in the Aardvark system, patients in other health care systems, the privately insured, and the uninsured. In addition to consumers of health care, discussions are joined by service providers, community leaders, and state and local politicians. Topics of conversation encompass inquiries treatments and services, news about health and health services research, discussions about the cost of health care, and political action.