■ Research Paper

What Can System Dynamics Learn From the Public Policy Implementation Literature?

I. David Wheat*

University of Bergen, Bergen, Norway

More often than not, system dynamics model-based public policy analysis is limited to testing parameter changes instead of designing and testing new stock-and-flow policy structures. That is problematic because improvements in behaviour require improvements in structure. This paper considers how the public policy implementation literature could improve the operational thinking skills required for designing policy structure for public sector models. A familiar model of a public health problem is used to illustrate the recommended approach. And an instructional training strategy is offered for teaching and learning to think operationally during the policy-design stage of modelling. Copyright © 2010 John Wiley & Sons, Ltd.

Keywords feasibility; implementation; policy design; public policy; system dynamics

INTRODUCTION

During World War I, American humorist and social commentator Will Rogers was asked how the Allies should deal with the enemy submarine problem. He said the solution was simple, 'Boil the oceans'. When asked *how* to boil the oceans, he responded, 'I'm a policy man. I let others worry about implementation'.

Whatever its motivation, the remark is a fitting satiric reference to then-President Wilson and the prevailing political science paradigm that separated policy *making* from policy *administration*. The only social scientist ever elected to the

presidency, Wilson (1887, pp. 209–210) had already laid the intellectual foundation for thinking that 'administration lies outside the proper sphere of politics' and that 'administrative questions are not political questions'. One of the most visible advocates of civil service reform during America's Progressive Era, he viewed 'the field of administration [as] a field of business removed from the hurry and strife of politics'.

Nearly a century passed before erosion of the Wilsonian position became noticeable in the political science literature. Among the first critics, Hargrove (1975) called implementation the 'missing link' in public policy analysis. Minogue (1983) remarked that a distinction between making policy and executing it would be 'absurd' to modern social scientists. Hill (1993, p. 235)

* Correspondence to: I. David Wheat, System Dynamics Group, University of Bergen, Fosswinkelsgate 6, N-5007 Bergen, Norway.
E-mail: david.wheat@uib.no

articulated the new view that policy *ends* are often reshaped by administrative *means*: 'What [the Wilsonian view] was prone to disregard was the extent to which [public administration] would tend to transform policy, often fundamentally. Implementation must be seen as part of policy-making'. And Bardach (2005, p. 33) supports the normative implications of the new perspective: 'Even the best policy planners do not get all the details right in the design stage. They should therefore allow room for policy implementers to improve on the original design'.

The public policy implementation literature is an important resource for understanding the feasibility of policies aimed at solving problems in the public sector. In this paper, we consider how that literature might contribute to system dynamics models of public policy problems.¹

Origin of Implementation Research

Domestic policy developments in the U.S. during the late 1960s provided the early case material for the small but increasing number of social scientists and policy planners interested in implementation. Asians and Europeans may remember the war in Vietnam as the defining issue of Lyndon Johnson's presidency (November 1963–January 1969), but American memory of that era is divided along foreign and domestic lines. In the U.S., unpleasant recollection of the war years competes with memories of the lingering legacy of ambitious but discredited social policy legislation rhetorically known as the 'Great Society' programs. The performance of those programs generally fell short of the promises made by the president and his congressional coalition.² During the 1970s, some researchers highlighted a common

denominator in the disappointing policy outcomes: poor implementation (cf., Williams and Elmore, 1976; Nakamura and Smallwood, 1980).

Variations in research strategies have led to distinctions in the meaning of implementation. In the second edition of *Implementing Public Policy*, Hill and Hupe (2009) take several pages to develop a working definition. After reviewing variants in the public policy literature, they defer to Ferman (1990) and settle on wording that is indisputable but vague: implementation is 'what happens between policy expectations and perceived policy results'. Such a definition may be useful for the scope of their book, as it underscores the breadth and significance of the concept. For our purposes, however, more precision is needed, and we opt for the simple definition offered by Mazmanian and Sabatier (1983, p. 1): 'Implementation is the carrying out of a basic policy decision' made by government officials exercising formal authority.

Both the broad and narrow definitions of implementation imply a sequence in the policy process. Howlett, Ramesh and Perl (2009) refer to five stages in that process: agenda setting, formulation, decision-making, implementation and evaluation. The implementation stage occurs after a policy is considered, formulated, and authoritatively adopted.³ Studying the implementation stage *per se* has emerged as the province of academic political scientists. They identify and explain the origin of obstacles that lead to disappointing policy outcomes. In contrast, analysts trained in public policy graduate schools use an eclectic, multi-disciplinary research approach to implementation issues. The challenge for the policy analyst is to reduce the likelihood of problems during the implementation stage by better planning during the formulation stage.

¹Limiting the scope of this paper to the *public* policy literature is not intended to deny the potential of other disciplines (e.g. operations research) to contribute to improvements in SD public policy modeling.

²The gap between promise and performance is a theme in a special issue of *The Public Interest* (http://www.nationalaffairs.com/doclib/20060404_NUMBER034WINTER1974.pdf). A more sympathetic treatment of the Great Society is provided by one of its architects in the October 1999 issue of *The Washington Monthly*, (<http://www.washingtonmonthly.com/features/1999/9910.califano.html>).

³These processes are not one-way streets; iteration and feedback occur. The policy feedback perspective emerged at least 75 years ago when E.E. Schattschneider (1935) observed that 'new policies create new politics'. Policy feedback is now an important strand within the tapestry of public policy research (cf., Pierson, 1993; Angland 1999). The feedback perspective has obvious appeal to system dynamicists, but that is not the focus of this paper.

Implementation and System Dynamics

A decade before the emergence of the public policy implementation research program, system dynamics (SD) was developed as a method for designing policy solutions based on computer simulation of problematic endogenous feedback structures. Initially conceived as a business management tool (Forrester, 1958), the SD methodology was soon applied to public sector issues, beginning with Forrester's (1969) model of economic growth and decline in large American cities. SD has become a model-based policy-design discipline. It is natural to ask where the concept of implementation fits in the 50-year-old paradigm. We want to know if implementation considerations are part of the modelling process, in principle and in practice.

To address that question, we begin by aggregating the SD modelling process into two high-level stages: Problem Explanation and Policy Design.⁴ The task in the first stage is to explain the reasons for the problematic dynamic behaviour of the system. That requires building and testing an *explanatory* model. In the second stage, the challenge is to design and test policies that alleviate the problem; that is improve the dynamic performance of the system. That requires building a *policy* structure (another 'model' if you will) that can be integrated with the explanatory model. Forrester made this distinction forty years ago and reiterated it recently. 'First... generate a model that creates the problem. Only if we understand the processes leading to the difficulties can we hope to [secondly] restructure the system so that the internal processes lead in a different direction' (Forrester, 1969, p. 113). In a December 2009

listserv posting, Forrester joined a discussion about climate change policy with a similar comment: 'A model should demonstrate how the symptoms are being generated... Only by clearly understanding what is causing the problem can one begin to see where [policy] attention should be focused'.⁵

For our purposes, the key term in Forrester's second-stage description is 'restructure'. At the policy-design stage, the explanatory model should be modified by adding stock-and-flow feedback structure that implements decision rules specifying when the new policy will become operational, how it will work, and what will happen over time to improve the performance of the system.⁶ The SD literature reiterates this point. Richardson and Pugh (1989, p. 332) stress that 'policy improvement... involves the addition of new feedback links'. Ford (1999, p. 177) makes clear that adding new stock-and-flow structure is necessary 'to describe the details of policy implementation'. And Sterman (2000, p. 104) emphasizes that 'policy design... includes the creation of entirely new... structures and decision rules'.

Therefore, an SD model containing new policy structure has implementation ('how it would work') assumptions embedded in the policy design, at least implicitly. That does not mean the embedded assumptions are always valid. There is no guarantee that the new decision rules reflect an adequate understanding of the implementation task or the obstacles in the way. If the model's new decision rules are based on naïve assumptions regarding policy feasibility, then the simulated behaviour generated by the model will raise false hopes about the efficacy of the policy. Yet, a naïve policy structure may be better than none at all. An explicit policy structure is ripe for improvement, and constructive criticism of underlying assumptions can lead to enhancements in the policy design.

In the absence of structural modifications to the explanatory model, policy analysis is limited to

⁴Descriptions of the SD modelling process can be found, for example in Randers (1980, ch. 6), Richardson and Pugh (1989, ch. 1), Ford (1999, ch. 15), Sterman (2000, ch. 2), Barlas (2002), Roberts (2007) and Morecroft (2007, ch. 5). Different authors have emphasized different details in the modelling process, but there is a consensus about the key steps when developing an *explanatory* model: (1) specify the dynamic problem, (2) develop a hypothesis in the form of a model that explains the dynamics of the problem, and (3) analyse the model with various validation tests. Of course, this is an iterative process; for example what the modeller learns during analysis often necessitates revising the model (and the hypothesis it represents) and further analysis. Guidelines regarding *policy* modelling distinguish between testing policy parameters and adding policy structure, with the latter required for assessment of a specific policy option.

⁵Email communication from Jay Forrester to System Dynamics K-12 Discussion listserv, 6 December 2009, 12:40 a.m. GMT. Used with permission.

⁶It may also be possible to improve systemic performance by *removing* structure that caused the problem.

'what if' questions stemming from changes in parameters that might be the target of a policy intervention. Ideally, feasibility issues would still be discussed, at least in general terms. However, assessment of a particular policy idea would not be informed by modelling 'how it would work'.

Concerns about the implementation component of SD models surfaced about the same time the 'missing link' issue was raised in the political science literature. From the beginning, the goal was similar to the aim of public policy analysts: improving prospects for effective implementation. Roberts' (1972) implementation strategies for SD models included practical suggestions for achieving model adoption and use by corporate clients, as well as criteria for evaluating policy options in terms of their organizational feasibility. Weil (1980, 1983) critiqued his own projects and identified conditions and consulting strategies that favoured effective implementation. More recently, Roberts (2007) reiterated the importance of taking implementation seriously throughout the modelling process and suggested specific steps for improvement, while Größler (2007) raised implementation questions about modelling projects that 'failed' in the business policy arena. In addition, SD textbooks address implementation issues (cf., Richardson and Pugh, 1989; Coyle, 1996; Ford, 1999; Maani and Cavana, 2000; Sterman, 2000; Morecroft, 2007), as does the Barlas (2002) SD contribution to the online *Encyclopedia of Life Support Systems*.

For SD modellers specializing in public policy analysis, implementation is sometimes elevated to the object of their research. Consider, for example *Simulating the Dynamics of Social Program Management: The Case of JOBS Implementation* (Ratanawijitrasin *et al.*, 1990) and *Using Group Model Building to Inform Public Policy Making and Implementation* (Zagonel and Rohrbaugh, 2008). Group model building specialists (cf., Vennix, 1999; Andersen *et al.*, 2003) have displayed a particular affinity for raising implementation issues in settings of multiple stakeholders.

Thus, implementation has not been a missing link in the SD literature. Implementation issues have had the attention of an influential part of the SD community from the earliest days of the

discipline. And the literature contains guidelines for effective implementation of SD models in both private and public settings. Ironically, however, there is little evidence that the guidelines are being *implemented*. Explicit policy structure – and the implementation assumptions embedded in stocks and flows – are hard to find in published SD models of public policy issues.

In the next section, we document this problem and frame it in terms of an operational thinking challenge. Then we draw on the public policy literature to improve operational thinking and policy-design modelling, illustrating with a simple model that needs a corrective policy.

THE POLICY DESIGN PROBLEM

A preliminary survey of the SD literature suggests that most models developed for public policy analysis do not include structure that indicates how policy options would be implemented. During its 25 year history, the *System Dynamics Review* (SDR) has published over 450 articles and notes, about one-third of which address public policy issues. Fifty-one contain some form of model-based public policy analysis, but only 14 (27 per cent) have models exhibiting new policy structures and decision rules. A recent example of a policy structure article is Wolstenholme *et al.* (2007), and an example from early SDR years is Saeed (1986).⁷

Caution is advised when interpreting these findings. By limiting the survey to articles published in the SDR, we have overlooked SD model-based articles published elsewhere. In particular, a survey of journals devoted to public sector issues might yield a higher percentage of articles containing policy structure. Another reason for caution is that mistakes could have been made during the categorization process. Yet it seems unlikely that the SDR articles are so unrepresentative or that categorization was so error-prone that the general finding would be overturned by a more thorough investigation. Further research is needed to be sure.

⁷The complete list of articles grouped into policy structure and policy parameter categories is available from the author upon request.

Most (73 per cent) of the model-based public policy articles in the *SDR* survey rest solely on an analysis of a model's sensitivity to policy parameter changes. And that is problematic since 'policies represented as parameter changes frequently tend not to be very effective in system dynamics analysis' (Richardson and Pugh, 1989, p. 332) and 'policy design is much more than changing the value of parameters' Sterman (2000, p. 104).

Although parameter testing alone is insufficient for policy design, it is necessary. It helps a modeller estimate the potential impact of a general strategy for influencing key feedback loops in a problematic system, what Richardson and Pugh (1989, p. 322) call finding a 'leverage point' in the system. For example, in an epidemic model, a key parameter is the probability of infection when a contact occurs between an infectious person and a susceptible person. Simulation experiments with different parameter assumptions provide estimates of the *potential* payoff from a strategy aimed at lowering the infection probability as a means of weakening the problematic reinforcing loop that drives the infection rate. It is quite another task to estimate and evaluate the *expected* payoff from a particular policy. This second task requires modelling the process by which the policy eventually has its expected impact on the infection probability. Parameter testing begs the question of *how* something represented by a parameter in a model gets changed in the real world. Specifically, we need to know what combination of tangible resources, perceptual and attitudinal adjustments, institutional capacity and coordination, authoritative and voluntary cooperation, and process time are needed to implement changes that would produce the desired parameter value.

Unless the word *design* is mere rhetoric, policy-design modelling requires a modification in the stock-and-flow structure of the base-case model (the 'explanatory model') that replicated the problematic dynamic behaviour. A fundamental SD principle is that a system's stock-and-flow feedback structure determines its endogenous dynamic behaviour. If endogenously induced behaviour was problematic in the past, then a

modification in structure is a pre-requisite for better behaviour in the future.

Designing a policy structure that is both effective and feasible requires what Richmond called 'operational thinking', by which he meant thinking in terms of 'how things really work' in the 'plumbing' of the problematic system (Richmond *et al.*, 1987; Richmond, 1993, 1994, 2000). But operational thinking must be contextualized. Modellers are not expected to know all the answers when assessing policy feasibility; contextual experts must be involved in the modelling process. What is more, modellers may not even know all the questions. Asking the right questions of others is a skill that must be learned and honed.

The premise of this paper is that the public policy implementation literature can help SD modellers ask questions that improve the operational thinking required for designing model-based policy structure. If that premise holds true, we can realistically anticipate the day when adding policy structure to explanatory models is the rule rather than the exception. A controlled experiment to test this two-stage hypothesis awaits further research, but the next section illustrates how such learning might occur and how it could contribute to policy-design modelling.

A POLICY DESIGN EXAMPLE

To illustrate the application of insights gleaned from the public policy implementation literature, we start with a pre-built, base-case model that replicates the problematic dynamic behaviour. Consider it the 'explanatory model' built during the Problem Explanation stage of the SD modelling process. That is, we assume the 'first step in modelling' (Forrester, 1969, p. 113) has already been taken. Our task now is to design and test new policies for improving model behaviour.

Base-Case Model

The base-case model is the familiar SEIR version of the epidemic diffusion model, containing stocks of Susceptible, Exposed, Infectious, and

Figure 1 Stock-and-flow structure of base-case SEIR model

Removed persons (Sterman, 2000, ch. 9). The stock-and-flow structure is displayed in Figure 1. The inflow shock is used to trigger a simulated epidemic by a one-time exogenous injection of ten infected persons.

SEIR modelling work typically yields quarantine and isolation policies that aim to reduce the daily contact between infectious and susceptible persons, as well as policies for reducing the probability of infection when a contact occurs. It is the isolation option (separation of infectious persons) that serves as our policy example. Quarantine (separation of exposed persons) is ignored because we will assume the infection can be transmitted only after the incubation period; that is by persons with symptoms of the disease. This simplifies our example with no loss of generality, and it happens to be a realistic assumption for some epidemics (e.g. SARS).

The causal loop version of the model in Figure 2 highlights the feedback structure. Counteracting loop C1 is responsible for the decline in the Susceptible Persons stock, while the single reinforcing loop R is responsible for growth in all the other stocks. The intervening loops C2 and C3 reflect first-order delay assumptions, causing the Exposed and Infectious persons stocks to fall after rising. The main story line is the battle for dominance between loops C1 and R. Structure-behaviour tests reveal that loop

R is dominant early in the simulation run, giving rise to exponential-like growth. Dominance eventually shifts to loop C1, however, and goal-seeking behaviour emerges.

Figure 3 displays the base-case simulation pattern of all four stocks over a time horizon of 120 days. The simulation begins with a fixed population of 1000 susceptible persons and zero exposed, infectious and recovered persons. Other parameter assumptions: the contact rate per infectious person is three per day, the probability of infection is 0.25, the incubation period is 5 days, and the recovery period is 7 days. The initial equilibrium is disturbed on day 15, when an exogenous shock injects 10 infected persons into the Infectious Persons stock. The resulting behaviour is s-shaped decline for the stock of Susceptible Persons and (almost) a mirror image s-shaped growth for the final destination stock: Removed Persons. The middle stocks – Exposed and Infectious – rise and fall in smaller s-shaped patterns, and their accumulations are responsible for delays that produce the slight asymmetry between the Susceptible and Removed stocks. (The Removed total eventually rises above 1000 due to the exogenous infected persons, but not all 1010 get infected.)

The problematic dynamic behaviour is the rapid spread of the epidemic, with the prospect of reaching almost everyone in the population in

Figure 2 Causal loop structure of base-case SEIR model

Figure 3 Behaviour of base-case SEIR epidemic model

just a few months if uncontrolled. The mechanism responsible for the growth of the epidemic is loop R, the only reinforcing feedback loop in the model. If that loop could be cut, the epidemic would end. The policy-design question, therefore, is 'How can loop R be eliminated, or at least weakened'? It is at this point that SEIR modellers begin to think in terms of reducing contacts between infectious and susceptible persons and/or making those contacts safer (i.e. reducing the probability of infection when a contact occurs).

This base-case model for the hypothetical epidemic example is assumed to operate without any socially responsive corrective policies; that is no public health warnings, no media coverage and no isolation or quarantine programs. Our task is to think operationally about the process of policy design. We separate that process into two steps: (1) identifying strategic leverage points in the model through policy parameter testing, and (2) designing stock-and-flow structure that represents a particular policy tactic. Along the way, we draw on the public policy implementation literature for assistance.

Policy Parameter Testing

In the base-case model, two parameters influence the epidemic growth feedback loop R: (1) the daily contact rate per infectious person, and (2) the probability of infection when there is contact

between infectious and susceptible persons. Both parameters are potentially influenced by health policy. For our illustrative example, the focus is on the daily contact parameter, initially assumed to be 3 per day. This parameter influences loop R by its effect on the total number of contacts by infectious persons.

Typical parameter sensitivity testing involves several simulation runs with different parameter settings. Whatever the deterministic setting, however, the parameter remains constant during the simulation. A stochastic setting adds a little more excitement to the sensitivity test results, but does not permit testing a specific hypothesized change over time in the parameter value. In the public policy literature, the work of Sabatier and Mazmanian (hereafter 'S&M', 1980, 1981) suggests an additional way to conduct policy parameter testing; namely, by using a dynamic implementation estimate of the parameter value.⁸

Consider, for example, S&M's 'Gradual Erosion Scenario' for policy implementation in Figure 4. It is a behaviour-over-time pattern describing the impact of policies that take time to get off the ground and reach desirable levels of achievement, only to settle out at a diminishing

⁸Paul Sabatier and Daniel Mazmanian have collaborated on numerous publications, and the 'lead author' has varied. In the list of references, the bibliographic format indicates the lead author. For simplicity in the text, however, all joint publications are cited as 'S&M' and the dates will serve to identify the particular reference.

Figure 4 Gradual erosion scenario. Source: Mazmanian and Sabatier (1981, p. 27)

level of performance. The pattern in Figure 4 suggests an underlying structure dominated early by reinforcing growth processes and later by eroding goal-seeking tendencies.

S&M offer this behaviour pattern – slow rise to a peak, followed by a decline – as a generic type, and suggest that it characterizes the dynamics of much public policy implementation, particularly in the regulatory domain. What happens later – continued erosion, steady state, or rejuvenation – is a function of the systemic structure responsible for the particular dynamic problem and its associated policies, buffeted by variable exogenous winds.

This dynamic policy impact assumption can be used in parameter testing. Figure 5 displays a modified version of the base-case model. There is now an exogenous policy link ('Dynamic Effect of Policy A'). This policy structure makes the daily contact rate vary as a nonlinear function of time, using the graphical table function in system dynamics software. The former parameter is now a variable with this equation:

dailycontactrateperinfectiousperson

$$= 3 * (1 - \text{Dynamic Effect of Policy A})$$

The Dynamic Effect of Policy A ('the Effect') can vary between 0 and 1. It represents the degree of conformity of policy outputs with statutory objectives (the vertical axis in Figure 4).

In Figure 5, when Policy A completely fails to reduce the contact rate, the Effect equals 0 and the daily contact rate remains 3 per day. When Policy A eliminates all contacts, the Effect equals 1 and

the daily contact rate falls to zero. An Effect value of 0.5 would cut the contact rate in half.⁹

Both the Effect and the contact rate change during the simulation run, according to the shape of the graphical function. There are, of course, an infinite number of nonlinear patterns that could unfold over time, but some are more likely than others. Here, we consider just two examples: a 'goal-seeking' pattern and an 'erosion' pattern, both displayed in Figure 6a.

A goal-seeking pattern has some intuitive appeal as a generic description of how implementation might unfold. One might hypothesize that implementation proceeds rapidly at first but eventually saturates at a level reflecting constraints imposed by relevant legal, political, bureaucratic, and socioeconomic structures. On the other hand, S&M (1980, 1981) argue that the erosion pattern is more typical, at least with regard to regulatory policies. In their view, after a slow start, implementation gains momentum, only to run out of steam and erode after reaching a short-lived peak. Their argument also depends on the assumed effects of pertinent legal, political, bureaucratic and socioeconomic structures.

For comparison with the standard method of sensitivity testing, the graph in Figure 6a also includes a static flat-line implementation assumption. In all three implementation estimates, the average value is a 50 per cent effect over a simulation time horizon of 120 days. Although they have the same *average value*, they differ considerably in the *patterns* that produce that average. We want to see how that difference in implementation pattern assumptions affects the behaviour of the epidemic generated by the modified SIER model.

Figure 6b displays the behaviour of the epidemic in the three scenarios stemming from the different implementation pattern assumptions. The *pattern* of the implementation impact – and not just its average impact – has important implications for the growth *pattern* of the epidemic. And transition patterns from one state

⁹The dynamic implementation parameter test is not intended to replace traditional static parameter sensitivity tests.

Figure 5 Base-case model modified by Dynamic Effect of Policy A

to another can be as socially and politically important as the final state achieved. A simulation based on a static impact assumption would forecast one epidemic pattern. If, in a subsequent real epidemic, the actual impact turned out to be time-dependent with a slow start, the epidemic pattern experienced would be worse than expected, measured in terms of the accelerating infection rate and the total number of people infected at the peak.

In nonlinear feedback systems, model-based policies tend to have variable rather than constant impacts on dynamic problems. Thus, traditional static parameter sensitivity tests may not provide sufficient insights about the potential impact of a policy strategy. Dynamic parameter tests, implicit in the public policy literature, could produce more realistic expectations regarding the impact of a policy strategy.

In the next section, we go beyond policy parameter testing and add new policy structure to achieve isolation of infectious persons.

New Policy Structure

The public policy literature emphasizes that every policy is based on some theory of cause and effect, and implementation is doomed without a valid causal theory (S&M, 1981, 1983; Hogwood and Gunn, 1984). Therefore, new policy ideas – no matter how clever or creative – have little chance to produce desired outcomes if merely grafted onto an inadequate causal theory. For example, a mandatory premarital AIDS testing policy in the state of Illinois relied on a theory that the mere existence of the test requirement would cause state residents to submit to the test. The policy was thwarted when couples held weddings in neighbouring states to avoid the test (Brandeau *et al.*, 1993, cited in Walt, 1994).

Moreover, S&M (1981) emphasize that those whose cooperation is needed to implement the policy must also subscribe to its underlying causal theory. In our example, assuming high

Figure 6 Implementation pattern (left) determines epidemic pattern (right). (a) Different implementation patterns but same average reduction in contact rate (50%). (b) Labels refer to implementation pattern assumptions at left

levels of cooperation and effort from health workers presumes they have confidence in the underlying epidemiological theory. In this example, that is not difficult to imagine. In other cases, there may be a serious credibility gap that hinders acceptance of a theory represented in an SD model. When designing policy, therefore, attention to the explanatory model is first and foremost. This should resonate with those who adhere to an SD modelling process that includes a battery of tests for building justifiable confidence in a model, as well as communication tools for conveying essential features of the model to a non-technical audience.

For our example involving a stock-and-flow feedback representation of an epidemic theory, Richmond's 'plumbing' metaphor is particularly apt for an isolation policy structure. We want to reroute people flowing through a system in much the same way that a plumber might want to redirect water through a different pipe.

We define our Isolation Policy as relocating infectious persons to a hospital and limiting their ability to have contact with susceptible persons while infectious; then releasing them to go home when they are no longer infectious. The policy is designed to weaken the problematic reinforcing loop. The growth rate of the epidemic could be slowed if the level of the infectious persons stock could be lowered, even if the contact rate did not change for those who remained in that stock. In

Figure 7a, structure for implementing the Isolation Policy has been added to the base-case model.

With the structure in Figure 7a, infectious persons would move into an Isolated Persons stock (e.g. in a hospital) after some delay (time until isolation), where they would remain until they were no longer infectious. If the time until isolation were shorter than the removal time, there would be some benefits from this policy, and if the time until isolation were very short (e.g. 2 days), the epidemic would have minimal adverse effects.

The simulated impact of this policy is displayed in Figure 7b, under three assumptions about how quickly people could be relocated from the Infectious Persons stock to the Isolated Persons stock ('time until isolation'): 2, 4 or 6 days. The base-case scenario (no isolation policy) is also displayed. As anticipated, the quicker the relocation occurs, the greater the expected impact of this policy. The impact dissipates as the time until isolation approaches the normal removal time in the base-case model.

Is the Isolation Policy likely to work as indicated in Figure 7b? If, for example, we assume that the average 'time until isolation' is 4 days, is it reasonable to expect the Isolation Policy to cut the peak number of infections in half, compared to the base-case ('no policy') scenario?

Figure 7 Isolation policy structure and simulation behaviour. (a) Simplified view of model after adding isolation policy structure. (b) Simulation behaviour with assumptions for 'time until isolation'

Answering that question requires an explicit recognition and critique of the assumptions embedded in the policy structure. We have assumed that our causal theory is valid and that we have 'buy in' from those who must carry out the policy. In addition, utilization of an implementation framework developed by S&M (1981) would reveal other implicit assumptions; for example that infectious persons would, in fact, go to the hospital soon after symptoms appeared, and that health workers would know how to identify infectious persons who suddenly start appearing at the hospital, would know how to isolate them in various facilities, would know how to prevent hospital personnel from becoming infected, would have sufficient hospital capacity to handle the surge in patients, and would have the skills and motivation plus materials and technology to make it happen with a high degree of certainty.

If any of these assumptions lack credibility, we should consider modifying the policy structure in a more realistic (i.e. 'operational') way. For example, if we had serious doubts about the protection of health care workers in contact with infectious persons, the policy structure would need a feedback loop from the Isolated Persons

stock to the total contacts variable. That revised structure would reduce somewhat the expected benefits of the isolation policy.

To illustrate further the extension of policy structure, we consider the possibility of inadequate hospital capacity during an epidemic. If it is doubtful that hospital facilities are adequate to handle the expected number of patients, we would want to know about expansion possibilities and delays that could be expected. If hospital capacity constraints were expected to slow the inflow of infectious persons to the hospitals, and if the expected rate of change in hospital patient capacity was of concern, more realistic policy structure would be added, such as in Figure 8a, which we have called the Constrained Isolation Policy.

The simulated behaviour of the Constrained Isolation Policy is displayed in Figure 8b with three scenarios. Again, the base case is 'no policy' at all, the same as in Figure 7b. The 'no capacity constraints' scenario is the same as the original '4 days' Isolation Policy scenario in Figure 7b. The 'capacity constraints scenario' reflects the inability of the hospital to handle the patient flow in a timely fashion. That makes the 'time until isolation' effectively longer than four days.

Figure 8 Constrained isolation policy structure and simulation behaviour. (a) Simplified view of model after adding constrained isolation policy structure. (b) Simulation behaviour with and without capacity constraints

Figure 9 Implementation framework. Source: Adapted from Mazmanian and Sabatier (1981, p. 7)

And the number of infectious persons peaks at a level that is about 60 per cent higher than the unconstrained '4 days' scenario.¹⁰

By acknowledging the impediments to policy implementation and building them into the policy structure, the simulated behaviour is more likely to resemble the eventual real-world behaviour resulting from the policy.

Public Policy Literature Contribution

As the policy structure was being developed in the previous section, flags were raised along the way – flags that reminded the modeller to think operationally and visualize what really happens during the implementation stage. Those reminders came from the public policy implementation literature, particularly from S&M (1981, 1983). After their own comprehensive review of that literature, S&M (1981) developed the implementation framework displayed in Figure 9.

The S&M framework 'is the best-known and most frequently used top-down analysis frame-

work' (Winter, 2006). It has been utilized in implementation assessments of policies covering a wide range of issues in diverse national settings. Six years after the framework first appeared in the literature, Sabatier (1986) could cite 21 empirical applications. At that time, the policy issues included environmental protection, education reform, urban renewal, civil rights protection and social welfare programs, with work performed by a variety of scholars in France (2), Germany, Norway (2), Poland, Sweden, the United Kingdom, and the United States (13). Since then, the range of policy issues has grown, along with the international research network (cf., Sabatier, 2009).

The value of the implementation framework is that it can be modified into a checklist of questions ('flags') useful to modellers seeking to represent implementation impediments in policy-design structure. Recall, for example that in the previous section emphasis was placed on (a) the importance of a valid causal theory and (b) the skills, resources and technology available in the health system. In this context, consider items #1 and #15 in Figure 9. For the epidemic problem, 'availability of a valid causal theory and technology' relates to the implementing agencies

¹⁰The Constrained Isolation Policy structure also contains parameter assumptions – capacity adjustment time and the slope of the nonlinear effect of patient density on the time until isolation – that would need sensitivity testing.

having a valid epidemiological theory (e.g. as represented by the SEIR model) and having the requisite technical means to act in accordance with that theory in specific situations (e.g. they can identify and isolate infectious persons). In our example, 'Resources of constituency groups' is relevant to hospital capacity for isolating infectious persons. The epidemic example in this paper draws on only a small part of a framework that has broad potential for helping modellers ask the right questions of contextual experts during the policy-design stage of modelling.

To further appreciate the relevance of the S&M framework, consider Table 1, which contains excerpts from a SARS epidemic checklist that correlate with items #1 and #15 in the S&M framework. Jointly prepared by the U.S. Centers for Disease Control and Prevention and the Association of State & Territorial Health Officials (hereafter, CDC-STHO), this checklist is striking evidence that successful street-level implementation of epidemic control policies involves an extraordinarily high degree of preparation and coordination by multiple agencies at different hierarchical levels. Even the partial checklist in Table 1 illustrates that the issue concepts in the S&M framework have counterparts in real-world institutional settings. The full list, containing 53 criteria for assessing emergency preparedness, reveals even more.¹¹

The S&M framework is certainly not the only source of operational insight questions that can probe implementation issues. Also useful is Hogwood and Gunn's (1984) list of ten 'perfect implementation' criteria. No real-world policy would ever score 'perfect' on any one of the criteria, much less all ten. However, the list is a handy reminder of the potential obstacles in the way of implementation.

Nor is the S&M framework the only conceptual apparatus in the public policy literature that SD policy modellers will find useful. When engaged in operational thinking at the policy-design stage, an analyst must make feasibility estimates. And Allison (1971) has argued that such estimates

hinge on the analyst's mental model of how governments function in a particular policy context; that is the extent to which the analyst considers policy implementation (i) a choice by a rational unitary decision-maker, (ii) an output of bureaucratic organizational procedures, (iii) a result of political conflict and compromise, or some combination thereof. The mental model will influence the aggregation level in the policy structure of the SD computer model and, ultimately, the assessment of a policy's feasibility. Therefore, an SD modeller of public policy problems needs to develop an understanding of the relevant governmental policy structure in addition to the structure responsible for the historical dynamic problem.

Along these same lines, Linder and Peters (1989) focus on the 'instruments of government', the choice of which facilitates or impedes implementation. Extending that line of research normatively, Bardach (2005) provides an extensive list of 'what governments do' accompanied by suggestions to help policy planners consider what instruments might be used to effectively address various types of public problems. SD modellers who are familiar with the literature on the instruments of government may be more creative when brainstorming policy options. Later, when reviewing the list of options, SD modellers with better mental models of governmental policy structures may be more likely to recommend policies that are feasible.

SD modellers looking for field guides to the implementation research literature, including data collection requirements and strategies, should consult Irwin (2003) and Werner (2004). Beyond these specialized volumes, Sætren (2005) documents the explosive growth in the implementation research area, including its spillover into disciplines beyond its political science origin. His article, along with the Hill and Hupe (2009) textbook, provide the essential map for further exploration of the public policy implementation literature.

DISCUSSION

Implementation has never been a missing link in the SD literature, but it may be the weakest link in

¹¹See the website of U.S. Centers for Disease Control and Prevention, accessed 30 June 2009 (<http://www.cdc.gov/ncidod/sars/guidance/a/app1.htm>).

Table 1 SARS epidemic checklist items exemplify Sabatier and Mazmanian framework concepts (source: U.S. Centers for Disease Control and Prevention; <http://www.cdc.gov/ncidod/sars/guidance/a/app1.htm>)

S&M Framework Concept: Availability of Valid Causal Theory and Technology

19. I know how to access current recommendations on treatment of cases and prevention of transmission in the hospital, long-term care and home care settings
20. My jurisdiction's emergency response planning has involved health-care product and service providers to determine how to best prevent and control disease spread and manage the health care of the population during an epidemic
28. My jurisdiction has a plan for ensuring that appropriate personal protective equipment, including N-95 or higher level respirators, is made available for persons whose job requires exposure to people with SARS, and that needed training and fit testing are provided
29. My jurisdiction has a plan for dealing with mass mortality, including transportation and burial of bodies
41. In the event of a SARS epidemic, I will have available daily counts of key community health indicators, such as numbers of emergency department visits, hospital admissions, deaths, available hospital beds and staff, facility closings, numbers of contacts being traced and numbers under quarantine
42. The public health laboratory that serves my jurisdiction can test for the SARS-associated coronavirus by serology and/or PCR
44. The public health laboratory that serves my jurisdiction has linked to clinical laboratories and provided training on the use of SARS tests, biosafety, specimen collection, packing and shipping, and rule-out testing
45. Public health laboratories in my state have computerized record keeping to help with data transmission, tracking, reporting of results to patients and facilities, and analysis during an epidemic
46. My jurisdiction has determined how to assess and document the spread and impact of disease throughout the population, including special populations at risk (such as health care workers and first responders), during a SARS epidemic, including enhancements to routine surveillance
48. My jurisdiction's epidemiology staff, in cooperation with other public health agencies, has the capacity to investigate clusters of SARS cases, to determine how disease is being transmitted, to trace and monitor contacts, to implement and monitor quarantine measures, and to determine whether control measures are working
49. My jurisdiction has plans for educating health-care providers about recognition and reporting of SARS, about the current case definition, and about sources of current information on all aspects of SARS

S&M Framework Concept: Resources of Constituency Groups

5. I know whether my state allows hospitals and other licensed healthcare institutions to use temporary facilities for provision of medical care in the event of a public health emergency
11. My jurisdiction has identified key stakeholders responsible for development and implementation of specific components of the SARS epidemic plan, including enforcement of isolation, quarantine, and closure and decontamination of premises
16. My jurisdiction has identified an overall authority in charge of coordinating different medical personnel groups during an epidemic
21. I am familiar with the required protocol for securing needed emergency health-care services and supplies during a public health emergency
22. My jurisdiction has identified ways to augment medical, nursing, and other health-care staffing to maintain appropriate standards of care during an epidemic
23. My jurisdiction has identified ways to augment public health laboratory, epidemiology and disease control staffing to meet emergency needs and in the event public health workers are affected by an epidemic
24. My jurisdiction has a process to recruit and train medical volunteers for provision of care and vaccine administration during a public health emergency
25. My jurisdiction has identified alternate facilities where overflow cases from hospitals and well persons needing quarantine away from home can be cared for and has developed processes with Emergency Medical Services to assess, communicate, and direct patients to available beds

Table 1 (Continued)

-
26.	My jurisdiction has identified facilities for outpatient and inpatient care of children with SARS and their families
27.	My jurisdiction's epidemic plan addresses the mechanics of how isolation and quarantine will be carried out, such as providing support services for people who are isolated or quarantined to their homes or temporary infirmity facilities and protection for workers providing these services
52.	Community partners such as hospitals, EMS services, law enforcement agencies, health-care practitioners, environmental hygiene/remediation services, news media, schools and colleges know what part they are expected to play during an epidemic and are prepared to do so.
-

SD models of public policy problems. Implementation assumptions should be discernible in stock-and-flow policy structure. However, policy structure in public policy models is the exception rather than the rule. More often than not, public policy analysis is limited to the study of parameter changes in a model. And that is problematic because improvements in systemic behaviour require improvements in systemic structure.

Designing policy structure for public sector models is not easy. It requires a capacity for operational thinking in what may be unfamiliar political and bureaucratic contexts. One way to build capacity involves learning to ask feasibility questions that the public policy literature suggests are important. To illustrate that approach, we designed a policy for an epidemic model with help from the S&M framework.

The example was selected for the simplicity of both the base-case model and new policy structure. There were no political complications. And there was no dissonance between legislative policy mandate and organizational mission or procedures, a situation often fraught with conflict and delay. In the absence of political and bureaucratic impediments, suggestions from other disciplines (e.g. operations research) might be adequate for designing the 'plumbing' of the policy structure.

¹²AHRQ is an agency of the U.S. Department of Health and Human Services. In research results cited by AHRQ, only one-fourth of the hospitals in the state of Utah were willing to be designated as isolation facilities. AHRQ cites the Federal Emergency Medical Treatment and Active Labor Act as 'problematic' for isolation policies because of required medical screening examinations during specified emergencies. See <http://www.ahrq.gov/research/altsites/altsite6.htm>, accessed 5 December 2009.

Imagine, however, cases where there is reason to doubt cooperation from hospitals or where there are conflicting government mandates. For example, information from the U.S. Agency for Healthcare Research and Quality (AHRQ) indicates the unwillingness of some hospitals to become designated isolation facilities for certain diseases. In addition, AHRQ offers a problematic example of a federal emergency medical treatment law that could actually hinder implementation of a local isolation policy.¹² Despite the simple example offered in this paper, prospects for authorization, funding, and effective operational capacity for epidemic policies could pose implementation obstacles in some jurisdictions. The argument here is that, in such cases, policy design and assessment could benefit from insights that may be unique to the public policy literature.

Getting Started

Exploratory pedagogic tests of the efficacy of this approach could be carried out in an advanced modelling process course for students or in a policy-modelling workshop for experienced modellers. Thorough training could take the better part of an entire course, but it could be squeezed into 12–15 h of lecture time (or an intense 3-day workshop). The outline below is for a series of course lectures organized into three distinct modules.

The first module should acquaint participants with the policy-making process regarding a public policy issue in a jurisdictional context of relevance to their studies. The overall aim would be to sensitize students to policy feasibility issues

and provide a 'field guide' to inform policy design. Allison's (1971) models would provide conceptual lenses through which students could begin to visualize governments-at-work in various national and international policy contexts. A checklist of implementation hurdles should emerge from comparing Allison (1971), Mazmanian and Sabatier (1981) and Hogwood and Gunn (1984). The checklist should be supplemented by considering policy instruments available to governments, as illustrated in Linder and Peters (1989) and Bardach (2005).

The second module should provide participants with opportunities to apply their policy feasibility insights to pre-built explanatory models. They should practice operational thinking when conducting (static and dynamic) policy parameter sensitivity tests. Students should be required to highlight implementation hurdles for specific policies and utilize their insights when interpreting the design and outcome of the sensitivity tests.

The final module should engage participants in policy design that requires adding stock-and-flow structure to pre-built explanatory models. Feasibility insights and implementation assumptions discussed during the second module should be discernible in the new policy structure, and simulation runs should facilitate comparing feasibility-constrained policy alternatives.

Extensions

After informal evidence is gained from hands-on training experiences, controlled experiments could be designed to test formally whether operational thinking skills can be improved with the aid of lessons learned from the public policy literature.

To be successful, instructional training programs will have to be accompanied by parallel efforts to motivate wider practice of structural policy design among experienced modellers. Along the way, some may discover what Richardson and Pugh (1989) found: 'The process of implementing policy change is a feedback problem in and of itself. In any given situation, it

may be as interesting as the original problem that stimulated the modelling study'.

Finally, despite our focus on *public* policy modelling, we are well aware that implementation problems are not confined to the public arena. When Größler (2007) examined *business* policy modelling projects that failed to make an impact, he called for an 'intensive discussion of how implementation can actually be achieved'. This paper aims to contribute to that broader discussion.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the helpful comments and suggestions of the special edition editors, David Lane, Carmine Bianchi and Enzo Bivona; workshop discussant Jim Duggan; and two anonymous referees.

REFERENCES

- Allison G. 1971. *Essence of Decision*. Little, Brown & Company: Boston.
- Andersen D, Rohrbaugh J, Richardson G, Lee T, Zagonel A. 2003. "Using Group Model-Building to Support Public Policy Development." Paper presented at the Conference of the U.K. Chapter of the System Dynamics Society, Harrogate.
- Angland SM. 1999. Policy Feedback: the Comparison Effect and Small Business Procurement Policy. *Policy Studies Journal* 27: 11–27.
- Bardach E. 2005. *A Practical Guide for Policy Analysis*. CQ Press: Washington, DC.
- Barlas Y. 2002. System dynamics: Systemic feedback modeling for policy analysis. In *Knowledge for Sustainable Development: An Insight Into the Encyclopedia of Life Support Systems*. UNESCO Publishing: Oxford.
- Brandeau M, Lee H, Owens D, Wachter R. 1993. Policy Analysis of Human Immunodeficiency Virus Screening and Intervention: an Overview of Modeling Approaches. *AIDS & Public Policy Journal* 5: 119–131.
- Coyle RG. 1996. *System Dynamics Modelling: A Practical Approach*. Chapman & Hall: London.
- Ferman B. 1990. When failure is success: implementation and Madisonian Government. In Palumbo DJ, Calista DJ, (eds.). *Implementation and the Policy Process: Opening Up the Black Box*. Greenwood Press: New York.

- Ford, A. 1999. *Modeling the Environment: An Introduction to System Dynamics Modeling of Environmental Systems*. Island Press: Washington, DC.
- Forrester JW. 1958. Industrial dynamics: a major breakthrough for decision makers. *Harvard Business Review* 36(4): 37–66.
- Forrester JW. 1969. *Urban Dynamics*. Productivity Press: Waltham, MA.
- Größler A. 2007. System dynamics projects that failed to make an impact. *System Dynamics Review* 23(4): 437–452.
- Hargrove EC. 1975. *The Missing Link: The Study of the Implementation of Social Policy*. The Urban Institute: Washington DC.
- Hill M. 1993. *The Policy Process: A Reader*. Harvester Wheatsheaf: Hertfordshire, U.K.
- Hill M, Hupe P. 2009. *Implementing Public Policy*. SAGE Publications: London.
- Hogwood B, Gunn L. 1984. *Policy Analysis for the Real World*. Oxford University Press: Oxford.
- Howlett M, Ramesh M, Perl A. 2009. *Studying Public Policy: Policy Cycles And Policy Subsystems* (3rd ed.). Oxford University Press: Ontario.
- Irwin LG. 2003. *The Policy Analyst's Handbook: Rational Problem Solving in a Political World*. M.E. Sharpe: Armonk, NY.
- Linder SH, Peters BG. 1989. Instruments of government: perceptions and contexts. *Journal of Public Policy* 9(1): 35–58.
- Maani KE, Cavana RY. 2000. *Systems Thinking and Modelling: Understanding Change and Complexity*. Prentice-Hall: Auckland, NZ.
- Mazmanian D, Sabatier P (eds.) 1981. *Effective Policy Implementation*. Lexington Books: Lexington, MA.
- Mazmanian D, Sabatier P. 1983. *Implementation and Public Policy*. Scott Foresman: Chicago.
- Minogue M. 1983. Theory and practice in public policy and administration. *Policy and Politics* 11: 63–85.
- Morecroft J. 2007. *Strategic Modelling and Business Dynamics: A Feedback Systems Approach*. John Wiley & Sons: West Sussex.
- Nakamura RT, Smallwood F. 1980. *The Politics of Policy Implementation*. St. Martin's Press: New York.
- Pierson P. 1993. When effect becomes cause: policy feedback and political change. *World Politics* 45: 595–628.
- Randers J. 1980. *Elements of the System Dynamics Method*. MIT Press: Cambridge, MA.
- Ratanawijitrasin S, Andersen D, Richardson G, Lurie I. 1990. "Simulating the Dynamics of Social Program Management: the Case of JOBS Implementation." Paper presented at the Annual Research Conference of the Association for Public Policy Analysis and Management, San Francisco.
- Richardson GP, Pugh AL. 1989. *Introduction to System Dynamics Modeling*. Pegasus Communications: Waltham, MA.
- Richmond B, Peterson S, Vescuso P. 1987. *An Academic User's Guide to STELLA*. High Performance Systems: Hanover, NH.
- Richmond B. 1993. Systems thinking: critical thinking skills for the 1990s and beyond. *System Dynamics Review* 9(2): 113–133.
- Richmond B. 1994. Systems thinking/system dynamics: let's just get on with it. *System Dynamics Review* 10(2–3): 135–157.
- Richmond B. 2000. *The "Thinking" in Systems Thinking*. Pegasus: Waltham, MA.
- Roberts EB. 1972. Strategies for Effective Implementation of Complex Corporate Models. Paper presented at the International Symposium on Model- and Computer-Based Corporate Planning, University of Cologne, Germany. Reprinted in Roberts (1978).
- Roberts EB. 1978. Strategies for effective implementation of complex corporate models. In *Managerial Applications of System Dynamics*, EB Roberts, (ed.), MIT Press: Cambridge, MA.
- Roberts EB. 2007. Making system dynamics useful. *System Dynamics Review* 23(2–3): 119–136.
- Sabatier P, Mazmanian D. 1980. The implementation of public policy: a framework for analysis. *Policy Studies Journal* 8(4): 538–560.
- Sabatier P. 1986. Top-down and bottom-up approaches to implementation research: a critical analysis and suggested synthesis. *Journal of Public Policy* 6: 23–48.
- Sabatier P. 2009. <http://www.des.ucdavis.edu/faculty/Sabatier/Sabatier.htm> (30 June 2009).
- Saeed K. 1986. The dynamics of economic growth and political instability in developing countries. *System Dynamics Review* 1(2): 20–35.
- Schattschneider EE. 1935. *Politics, Pressures, and the Tariff*. Prentice-Hall: New York.
- Sterman J. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill: New York.
- Sætren H. 2005. Facts and myths about research on public policy implementation: out-of-fashion, allegedly dead, but still very much alive and relevant. *The Policy Studies Journal* 33(4): 559–582.
- Vennix J. 1999. *Group Model Building: Facilitating Team Learning Using System Dynamics*. John Wiley, Sons: West Sussex.
- Walt G. 1994. *Health Policy: An Introduction to Process and Power*. Witwatersrand University Press: Johannesburg.
- Weil HB. 1980. The evolution of an approach for achieving implemented results from system dynamics projects. In *Elements of the System Dynamics Method*, 1980. Randers J, (ed.), MIT Press: Cambridge, MA.
- Weil HB 1983 The dynamics of strategy implementation. *Dynamica* 9(1): 43–47.
- Werner, A., 2004. *A Guide to Implementation Research*. The Urban Institute Press: Washington.

- Williams W, Elmore RF. 1976. *Social Program Implementation*. Academic Press: New York.
- Wilson W. 1887. The study of administration, *Political Science Quarterly* 2(2): 197–222.
- Winter S. 2006. Implementation. In BG Peters, J Pierre (eds.) *Handbook of Public Policy*. Sage: London.
- Wolstenholme E, Monk D, McKelvie D, Arnold S. 2007. Coping but not coping in health and social care: masking the reality of running organizations beyond safe design capacity. *System Dynamics Review* 23(4): 371–389.
- Zagonel A, Rohrbaugh J. 2008. Using group model building to inform public policy making and implementation. In *Complex Decision Making*. Springer: Berlin.