Modelling social care complexity: the potential of System Dynamics

Douglas McKelvie
The School for Social Care Research

The School for Social Care Research is a partnership between the London School of Economics and Political Science, King’s College London and the Universities of Kent, Manchester and York, and is part of the National Institute for Health Research (NIHR) http://www.nihr.ac.uk/.

The School was set up by the NIHR to develop and improve the evidence base for adult social care practice in England. It conducts and commissions high-quality research.

About the author

Douglas McKelvie is a partner of The Symmetric Partnership, a consultancy specialising in using innovative approaches, such as simulation and modelling, to support thinking, planning and commissioning in the health and social care sectors. He has developed models for the NHS, government departments, other public bodies and third sector organisations, across a wide range of applications including delayed discharge, capacity planning, the impact of new treatments, many aspects of mental health, prison capacity, and workforce planning. He previously had a career in social work, including working as an adviser with the Central Council for Education and Training in Social Work and the Scottish Social Services Council. He is a member of the Policy Council of the UK chapter of the System Dynamics Society.

NIHR School for Social Care Research
London School of Economics and Political Science
Houghton Street
London
WC2A 2AE
Email: sscr@lse.ac.uk
Tel: +44 (0)20 7955 6238
Website: www.sscr.nihr.ac.uk

© School for Social Care Research, 2013
ISBN 978-0-85328-473-4

This report presents an independent review commissioned by the NIHR School for Social Care Research. The views expressed in this publication are those of the authors and not necessarily those of the NIHR School for Social Care Research, the Department of Health, NIHR or NHS.
ABSTRACT

Social care research requires understanding of complex issues surrounding human need and ageing, capacity and impact of services, sustainability, use of resources, and change over time. Social care and social work theory draw on systems thinking, and simulation is used in education and training. Yet use of modelling and simulation to aid understanding is rare. This review introduces one approach to modelling and simulation, System Dynamics, which can bring potential benefits and insights. Examples of its use in the health sector are cited, and its possible relevance to social care discussed. The review introduces some generic building blocks for representing populations, services, human resources, and finance, and outlines a possible methodology, group model building, as an approach that uses simulation to enhance thinking and learning.

RECOMMENDATIONS FOR RESEARCH ON ADULT SOCIAL CARE PRACTICE

• Most problems that include any of the following might benefit from an element of System Dynamics modelling: population change, ageing, long-term conditions, health and social care integration, resource allocation across a pathway, policy analysis, workforce planning, and options appraisals. There are, of course, other social care scenarios to which System Dynamics could be usefully applied.

• Development of a small System Dynamics model towards the start of an enquiry can clarify more precisely its focus and highlight data definitions/requirement.

• The field of System Dynamics is relatively small in the UK; it needs support to extend beyond a small number of business schools into mainstream social science.

• Social care researchers need a greater awareness of the potential of System Dynamics to enhance understanding of complex, dynamic problems, and to recognise when they are dealing with an issue that might benefit from this approach.

• Social care researchers who already have modelling and simulation expertise should be encouraged to incorporate this within current work.

• There should be greater dialogue between the modelling and simulation field and social researchers, for example, through participation in each other’s networking events and conferences.

• There is great potential for a conversation between health economists with an interest in social care and system dynamicists; each group is concerned with almost identical problems, especially at the aggregate level, but using a different approach.

• Improving the use of System Dynamics within social care research would benefit from:
  – the establishment of a centre of excellence that brings together social care researchers and system dynamicists;
Modelling social care complexity: the potential of System Dynamics

– commissioning of some small-scale projects involving collaboration between social care researchers, social care practitioners, service users and carers, and system dynamicists;

– training for social care researchers in modelling and simulation;

– more awareness amongst those working in social care practice and policy of the potential benefits of engaging with System Dynamics.

KEYWORDS
Social care, social work theory, modelling, simulation, System Dynamics, computer, learning, complexity, non-linear, change over time, group model building, ageing, limits

ACKNOWLEDGEMENTS
The author is hugely indebted to: Professor Eric Wolstenholme, whose thinking about capacity planning, sector boundaries, bottlenecks, archetypes and group model building permeates this paper; Steve Arnold, David Monk, Sarah Nield and David Todd of Symmetric; Gill Smith and OLM. The SSCR gave valued support, and Sarah Moncrieff brought a fresh look to the diagrams.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>When and why to consider using system dynamics</td>
<td>2</td>
</tr>
<tr>
<td>What kinds of problems are not amenable to a System Dynamics approach?</td>
<td>3</td>
</tr>
<tr>
<td>System dynamics in context</td>
<td>4</td>
</tr>
<tr>
<td>System dynamics and health</td>
<td>10</td>
</tr>
<tr>
<td>Examples of published studies of system dynamics modelling in social care</td>
<td>11</td>
</tr>
<tr>
<td>Systems thinking and simulation in social work theory</td>
<td>13</td>
</tr>
<tr>
<td>Elements of, and structures in, system dynamics models of services</td>
<td>14</td>
</tr>
<tr>
<td>Population and ageing</td>
<td>14</td>
</tr>
<tr>
<td>Service dynamics, capacity, length of stay and boundaries</td>
<td>18</td>
</tr>
<tr>
<td>Hospital and social care model with greater complexity</td>
<td>28</td>
</tr>
<tr>
<td>Including variation, for example when modelling capacity</td>
<td>29</td>
</tr>
<tr>
<td>Modelling need or dependency states</td>
<td>33</td>
</tr>
<tr>
<td>Other building blocks (finance, workforce)</td>
<td>36</td>
</tr>
<tr>
<td>Workforce planning/modelling</td>
<td>38</td>
</tr>
<tr>
<td>Putting these components together – studying the impact of interventions over time</td>
<td>39</td>
</tr>
<tr>
<td>Using a model – ageing and dementia</td>
<td>44</td>
</tr>
<tr>
<td>Software</td>
<td>47</td>
</tr>
<tr>
<td>Example of a model interface and model behaviour: dementia</td>
<td>49</td>
</tr>
<tr>
<td>The model building process and group model building as a particular approach</td>
<td>54</td>
</tr>
<tr>
<td>A word about data</td>
<td>58</td>
</tr>
<tr>
<td>Conclusion</td>
<td>60</td>
</tr>
<tr>
<td>References</td>
<td>61</td>
</tr>
</tbody>
</table>
INTRODUCTION

This Methods Review builds on the earlier School for NIHR School for Social Care Research's (SSCR) Review of mathematical modelling and its application to social care (Squires and Tappenden, 2011)*, which outlined a range of modelling approaches that might contribute to social care research. That paper touched on System Dynamics (SD) but with no scope to go into detail.

This review is intended to inform social care researchers and others about potential applications of SD. No prior knowledge is assumed. SD offers much to anyone trying to make sense of the behaviour of complex, interconnected systems.

The review considers, first, when and why such an approach might be considered before outlining what SD is, locating this discussion within a broader epistemological debate. Being concerned with modelling complex systems, SD arguably does not fit neatly within familiar notions of quantitative/qualitative approaches.

Then, a literature review starts by noting that SD is already used in the health sector; brief reference is made to some themes. But social care appears in the SD literature rarely in its own right, and mostly in relation to health. Some rare examples of SD applied specifically to social care or social work are outlined. Nevertheless, social work theorists have drawn extensively on “Systems Thinking” and educators use “simulation”, so might they be receptive to a discipline, Systems Dynamics, that brings these together?

The review then sets out a practical introduction to the building blocks of SD: stocks and flows. Starting with simple structures, some templates are outlined illustrating the variety of phenomena that could be modelled. Moving away from stock/flow diagrams, the idea of a user interface is introduced. A model is only worthwhile if it can be used to communicate to a wider audience.

Finally, one of the most common approaches to building and learning from a simulation is introduced: group model building. This reinforces the view that a model is primarily a tool to support thinking and learning, and not a magical device for forecasting the future.

So far, SD has been used in social care in a small way, mostly as a part of relatively brief, consulting engagements. Its potential to inform larger-scale research has not been fully exploited. This review aims to contribute to these debates and possible developments.

Some of the smaller models described in this review will shortly be made available in web-accessible format at www.symmetricpartnership.co.uk/sscr.

* http://www2.lse.ac.uk/LSEHealthAndSocialCare/pdf/SSCR%20Methods%20Review_7_web_2.pdf
WHEN AND WHY TO CONSIDER USING SYSTEM DYNAMICS

The reasons outlined by Squires and Tappenden (2011) for using mathematical models apply equally to SD, namely:

- by using a simulation one goes beyond reliance on intuition alone;
- a model provides a means of synthesising a wide range of diverse types of evidence;
- the ability to examine scenarios without the ethical impossibilities of experimenting on an actual population.

There is a considerable overlap between the potential areas of application of different modelling approaches. While it is usually obvious when a particular approach is not indicated, there are often areas where more than one type of model could be considered. So, when would one consider using SD? What kinds of problems does it tackle and what are its limits?

SD is concerned with how complex systems change over time, and it has a particular conceptual approach. The kinds of problem most amenable to a SD study would exhibit one or more of the following:

- **Complexity arising from feedback effects.** For example, in this review we consider ageing. There is a population of older people, some of whom are above the threshold of needing care at home. One consequence of providing care is that people are enabled to live longer at home perhaps before needing residential care, thereby increasing the number of people who need care at home. By providing care to this group we actually increase the number of people in that group (and perhaps reduce the number in a related group). This is what is meant by dynamics.

- **Behaviour over time,** especially where time spent in the system is significant. For example, in workforce planning there is normally a lengthy gap between the enrolment of students or trainees, the emergence of qualified staff, and their development into experienced practitioners. The challenge is to recruit sufficient students/trainees now to meet an imperfectly known demand many years hence. SD is an excellent method for tackling a range of workforce planning issues.

- **Complexity arising from the need to plan capacity across a whole system of care,** especially where pathways cross agency boundaries and the system contains multiple bottlenecks at transition points. The most obvious example here is the need for an integrated approach to planning health and social care services in a way that avoids/minimises the scope for delayed transfers of care. Key inputs to such a model are demand assumptions, capacities of differing kinds, rates of transition and lengths of stay in pathways.

- **Population dynamics,** such as where demographic change is having an impact on the
demand for services, whether as a result of the ageing population, the improved life expectancy of people with some disabilities, or the need to provide care for people with long-term conditions.

- **Service redesign**, i.e. once talking and thinking about new ways of working reaches the point where implementation can be envisaged, a model can provide a rich picture of something that existed only in the minds of its inventors. For example, the model might make explicit some difficult decisions about disinvestment, including the possibilities that some service systems will become more expensive for a while before savings are realised.

- **Innovation**, for example, the funding of services through social finance. A model can assist a group to assess the feasibility and sustainability of such new approaches. Because of its grounding in the real world of operations, its ability to simulate a range of different time delays (between investment and impact, between impact and the generation of realisable savings) such a model can provide a richer, more dynamic picture than a spreadsheet based cash-flow forecast.

- **Policy analysis**, for a good SD model will simulate the impact of different managerial policies.

### What kinds of problems are not amenable to a System Dynamics approach?

One definitely would not consider SD when considering a problem that had no time dimension (although such problems are rare in the care sector). Examples include linear allocation problems where the optimum combination of ingredients is sought in relation to some objective function (such as maximising profit or minimising cost). SD tends to inhabit the domain of “wicked” problems characterised by multiple, interdependent nonlinear effects.

Although SD models simulate flow, one probably would not use SD to simulate a discrete process in detail (for example, arrivals at and departures of people from a reception system, where an element of randomness is involved). But, SD certainly has much to offer when considering broader problems of queuing and waiting times within the context of a whole system. For example, the waiting time for elective surgery might be causally related to the availability of home care (considered later in this review). To the extent that is true, anyone trying to reduce elective wait times would experience limited success if they conducted an in-depth study of only the elective waiting list and service capacity, ignoring the wider system. The converse also holds. In a detailed study of an accident and emergency department, Lane et al. (2000) point to the need to take account of the whole system; it might be possible to reduce hospital bed capacity without impacting on A&E but only because the safety-valve of reduced elective admissions takes the strain.
SYSTEM DYNAMICS IN CONTEXT

What is SD and how might it contribute to social care research? Is SD a discipline, a field of study, a methodology or a world-view? How does modelling relate to research? Is modelling a form of research or does it belong to the domain of consultancy?

System Dynamics combines two familiar words that are already used extensively in social care, sometimes in combination (for example, when describing the dynamics of family systems), and uses them in a particular way. SD in our context describes a specific field of study that originated in the 1950s, invented by JW Forrester, an engineer and pioneer of digital computing, at Massachusetts Institute of Technology (MIT), and which features:

- An overarching concern with how complex systems change over time;
- A conceptual toolkit of stocks and flows;
- The identification of causal connections and feedback effects;
- System as cause, i.e. the idea that the source of a problem may be found within the system (endogenous) rather than imposed from outside (exogenous);
- The importance of computer simulation as a tool for learning.

SD practitioners tend to distinguish between Systems Thinking and System Dynamics. It is probable that social researchers will have encountered examples of the former, but perhaps not the latter. Systems Thinking covers a wide range of approaches utilising the language of systems (interconnectedness, processes, inputs, outputs, holism) typically leading to the production and discussion of diagrams. System Dynamics extends the diagrammatic approach into the realm of computer modelling and simulation.

SD has a very specific analytical toolkit, comprising stocks (also known as levels), flows (or rates), and feedback loops.

A stock represents a resource (noun), which might be people, money, or a physical resource (such as hospital beds). A flow (verb) represents the movement of material into, between, and out of stocks. A key insight is that the value of a stock can only change as a result of what flows into, or out of, it. For example, a bank balance is a stock. The only ways to change the amount of money in a bank account are to change its inflows or outflows, i.e. by earning or spending at a different rate.

SD employs two methods of drawing diagrams, Causal Loop Diagrams and Stock – Flow Diagrams, both demonstrated in Figure 1.

* A good example would be Peter Senge (1990) *The fifth discipline; the art and practice of the learning organization.*
Causal loop diagrams emphasise causal connections between variables. A connection marked with a plus sign denotes a positive connection; minus denotes negative. So, in the first diagram in Figure 1, there is a positive connection between births and population, and between population and births. The connections form a feedback loop, marked R for reinforcing. There is a positive relationship between population and deaths, but a negative relationship between deaths and population. This also forms a feedback loop, marked B for balancing. Most feedback loops involve more than two variables. The convention is to trace the links around a loop; an odd number of minus signs means the loop is balancing and an even number (including zero) of negative signs means the loop is reinforcing. In the case of births and deaths, these polarities are quite obvious:

- As population increases, births increase; as births increase, population increases (reinforcing);
- As population increases, deaths increase; as deaths increase, population reduces (balancing).

The stock/flow notation in Figure 1 shows the same structure using conventional SD notation. The stock and flow symbols are obvious. Note that in this diagram, the cloud signs denote either a source or destination outside the system boundary; in our model we...
are not interested in the metaphysics of either our pre-birth place of origin or post-death destination. The circular symbols represent converters or auxiliary variables. The curved arrows denote causal connections.

This is the kind of diagram that is generated when using SD software. Contained within each of the objects in the stock/flow diagram is either a number or an equation, enabling the model to be stepped forward in time increments. A stock variable must be given only an initial value (its value when the model starts); once the model is running, its value changes each time-step in line with its inflows and outflows. Variables into which connections have been made must be formulated in terms of the connected variables. Here, given the normal convention for birth and death rates (expressed per thousand of population per year), and given a model time step of one year, these expressions would be:

- Births = population times birth rate divided by 1,000
- Deaths = population times death rate divided by 1,000

To observe a system by freezing it at a moment in time, the parts that can be seen are the stocks. But in order to understand how a system changes (behaves), we must pay attention to its flows, and discover what drives them, if necessary, using ancillary variables such as fractional rates (for example, birth rates or lengths of stay).

As we have seen above, feedback loops exist within systems, and are either reinforcing or balancing. Reinforcing behaviour exists where the growth of a variable is dependent on itself, for example the relationship between births and population, or compounding interest. Balancing behaviour results from the impact of a goal or limit, for example a service capacity (such as hospital beds) or a budgetary constraint.

The classic behaviour of a reinforcing loop is exponential growth, and of a balancing loop, is the tendency towards a limit. Since most systems consist of more than one feedback loop, of differing types, these combinations of feedback and delay-effects give rise to a number of distinctive patterns of behaviour. These patterns are sometimes characterised as archetypes (Wolstenholme 2003).

SD is explicitly a product of the computer age; a key premise is that the calculations required to represent how systems change over multiple time steps would be impossible to execute without a computer. Although diagrams can be helpful, one cannot deduce how a complex system will behave under given conditions from the exploration of a diagram alone. Equally, it may be impossible and/or unethical to conduct experiments on complex human systems, for example, to test the efficacy of an intervention.

In many domains, it would be commonplace to use a form of simulation as a thinking or learning tool. For example, a new aircraft design would be tested using a scale model in a wind tunnel. Someone learning to fly will first spend time in a flight simulator. Architects
routinely produce scale models of new buildings. For some reason, simulation is not normally considered as a part of social service design, even though this requires careful examination of interconnected services and balancing of scarce resources across different capacities.

Forrester’s (1961) early SD studies were in industry, serendipitously moving on to modelling urban problems (1969), always with the premise that the cause of problematic phenomena would lie within the system (endogenous explanation). The work that brought SD to a wider audience was the controversial Limits to Growth study undertaken for the Club of Rome in 1972. Limits to Growth relates the use of the earth’s natural resources to population dynamics. The resulting computer model appeared to show that, without corrective action, the world would no longer support a growing population. The most dramatic scenario was described as overshoot and collapse.

For the purposes of this review, the Limits to Growth controversy exemplifies how the use of computer simulations can be misunderstood, or misconstrued, as an attempt to predict the future. As the authors of Limits to Growth explained in their “30 year update” (Meadows et al. 2004), the purpose of modelling is not to make definitive predictions but to illustrate “broad sweeps” of the future. A model is a simplified view of reality, useful for some specific purpose: how else are we to grapple with understanding complex systems that encompass delays, accumulations, nonlinearities and feedback loops?

A major strength of SD is that it attempts to analyse the big picture, or whole system, but is rooted in the real world of operations conceived of as “things that are causally connected“ (for example, people, raw materials, service capacities, money) and processes. An early pioneer of the field (Richmond 1993) contrasted an econometric model of the American dairy industry with the kind of operational thinking that characterises SD: “The model contained a raft of macroeconomic variables woven together in a set of complex equations. But nowhere in that model did cows appear” (p.128).

SD has a large number of areas of application spanning multiple disciplines, from industrial processes, to use of natural resources (all of which are stocks, mostly with a negligible inflow, for example, non-renewable energy sources), financial modelling, population modelling, human services, workforce planning, and even biochemical processes. There is a very active international community of practitioners who are applying SD in the health sector. Being interdisciplinary in nature, many of those who use SD have had an earlier career in a particular domain.

In the UK, SD tends therefore to reside in serendipitously formed pockets of academia, mostly in business schools, as a part of management science or operations research. Unsurprisingly, the examples cited in its key texts tend to be business applications. Sterman’s (2000) book remains the most comprehensive textbook, locating SD within the broader context of systems thinking, with a strong focus on business but drawing on examples from health and public policy from a North American perspective. From the UK, Morecroft’s textbook (2007) includes a chapter on public sector applications, and Warren...
(2002) has applied SD extensively in the field of business strategy, mostly from the perspective of the firm operating in a competitive environment but with perhaps a more accessible treatment of the main SD building blocks. The definitive textbook outlining how to use SD in human services remains unwritten.

SD practitioners tend to be interested in the modelling process, and how it relates to cycles of learning. Whereas the popular image of a modeller might be someone whose main concern is predicting the future, the main preoccupations of system dynamicists are thinking and learning, and the relationship between the model generated in a project and individual or group mental models.

SD has attributes that can disturb the sensibilities of some social scientists. It may appear to be an uncritical application of a scientific method from the physical world and industrial processes, into social systems. It undoubtedly contains elements of determinism with its axiom that system structure drives behaviour, its interest in cause above correlation, and its insistence on the value of quantification. Forrester (1991), for example, has suggested that social systems are subject to the same determinants of behaviour as physical or engineering systems.

But to the extent that stocks and flows are ubiquitous in human systems, SD provides a way of understanding their behaviour. A stock comprising people (such as people with unmet needs, people on a waiting list, or who are using a service) has this in common with a stock of a physical resource; its level can only change if its inflows or outflows change, and that flow behaviour is subject to a complex network of resource constraints or policy decisions. The key difference is that this stock consists of people who are self-conscious and have agency.

Given the multi-disciplinary nature of SD, members of its community inevitably represent a range of epistemological positions. There is a risk that practitioners could unwittingly commit the reification fallacy of treating complex events as if they were natural objects governed by recurrent processes and universal laws, ignoring the role that human intention and agency play in their construction.

Lane (2001a, 2001b) tackles this head on, alert to the possible charge of positivism that could attach to a quantitative method: “system dynamics is strongly shaped by its engineering roots …. system dynamicists frequently do not speak a language that communicates well with social scientists“ (2001b). He acknowledges that in adopting a systems approach there is a danger of guilt by association with Parsonian functionalism. But “models do not presume to be a complete representation of human agency, merely a vehicle for rational debate involving a range of criteria and perspectives“ (2001b).

SD is an approach to understanding non-linear complex systems. By non-linear, we mean that the causal relationship between two variables is not proportional. In our home care example, there is a relationship between receiving home care and being able to remain at home for longer, but that relationship is not proportional. The effect of receiving 8 hours
of care per week is not necessarily twice that of receiving four hours. Thence, to model the impacts of social care services requires a non-linear methodology.

Halmi (2003) considers what social work theorists should make of the recent interest in non-linear approaches such as are found in chaos or complexity theory. Dismissing criticism that this exemplifies a naïve scientism uncritically applying new scientific discoveries, stemming from the popularisation of chaos theory, to social theory, he cites (Harvey and Reed, 1996) a continuum of modelling strategies from “predictive” to “historical narratives”, in which SD would probably fit within the category of “iconological modelling” using pictorial methods and visual correspondences to map complex systems of non-linear differential equations.
SYSTEM DYNAMICS AND HEALTH

SD has been extensively applied in the health sector, often with implications for social care at the boundaries of the health and social care systems (ageing, delayed discharge, addiction services, inter-agency working in mental health). Typical applications include:

- Management of long-term conditions such as diabetes;
- Public health policy, for example polio eradication;
- Whole system capacity planning to improve integration and minimise delays;
- Mental health population and service configurations;
- Substance abuse;
- Management of epidemics and infections;
- Workforce planning.

In the USA, Homer, Hirsch and Milstein (for example, Homer et al. 2007; Jones et al. 2006) have conducted an extensive range of SD studies of health, with a particular focus on the management of long-term conditions such as diabetes (for a good overview of this body of work see Homer 2012). A recurring theme is the impact of early detection and case management to slow down the rate of deterioration and thereby reduce the number of people with long-term, irreversible complications of the condition.

At a global level, a model constructed by Thompson and Tebbens (2008) proved to be influential in persuading the World Health Organization’s Global Polio Eradication Initiative to counteract the danger of what they described as “wavering”. As the success of vaccination initiatives led to a smaller incidence of the problem, or:

With success comes a perception that the high level of investment compared to the low incidence is no longer justified. If policymakers succumb to the resulting pressure to reduce vaccination spending, this creates a situation in which populations again become vulnerable to new outbreaks (op. cit. p.436).

In its study of healthcare associated infections in hospitals in England, the National Audit Office (2009) drew on a SD model to compare the effectiveness of a number of policies/interventions.

Whereas these models have simulated the dynamics of ill health, where the stocks and flows represent illness states, studies in the UK, probably because of the different role of the state as service provider, have focused more on the dynamics of services, and the delays that arise when respective capacities in consecutive services in a treatment chain are not appropriately calibrated, taking into account demand, progression rates and length of stay (Wolstenholme et al. 2007b, Eldabi et al. 2011).
EXAMPLES OF PUBLISHED STUDIES OF SYSTEM DYNAMICS MODELLING IN SOCIAL CARE

As already noted in the SSCR Methods Review of Mathematical Modelling (Squires and Tappenden 2011), there are few examples of mathematical modelling being used in social care research. In contrast to health, there are few examples of SD studies covering only social care. A search of Social Care Online (http://www.scie-socialcareonline.org.uk/) for “system dynamics” yielded nine records, not all of which use the term in our context. Not surprisingly, where social care does appear in the SD literature it is as one part of the whole system of care, almost always in relation to the health system in some way, most often in relation to the interface between community services and hospital. This is especially true of work in the UK.

Because of its suitability for modelling population dynamics in relation to service use, probably the most common SD studies relate in some way to the demands of an ageing population — and the impact of dementia.

In a good example of work describing an actual SD model, Bayer et al. (2007) modelled the impact of introducing telecare services for frail elderly people living in the community, under a range of assumptions about cost and efficacy (for example, reductions in hospital admissions or admissions to residential care). They noted that the time before impact might be discerned is perhaps longer than normal service planning horizons, and cautioned against overly-optimistic expectations about the demand for institutional care in the medium term. Over twenty years significant reductions might be realised.

Desai et al. (2008) built a SD model to assist Hampshire Social Services to consider the impact of the ageing population on demand for assessment and care-managed services. The model went into most detail about the stages involved in the referral and assessment processes, using array functions to represent a diverse range of care package types. The sole focus of the model was on demand for local authority services.

SD has a great deal to offer in developing the integration agenda between health and social care. In particular, the interface between hospital and community, including the potential use of intermediate care services either to reduce admission rates or speed up discharge processes. This scenario is an ideal candidate for SD modelling. Eldabi et al. (2011) report on a local authority-led project using group collaborative model building to explore capacity requirements under a range of scenarios.

Wolstenholme (for example, see Wolstenholme et al. 2007a) has written extensively about the interface between health and social care, examining how patients/service users flow through health and care systems, exploring how and why blockages emerge (such as delayed discharges), and modelling the informal coping strategies by which those operating the system attempt to keep it working when, from a design perspective (using some simple calculations based on inflow, length of stay and capacity and the proportions
in which people move on to other services) the capacity across the whole system should be rebalanced.

Ackerman et al. (2010) used an innovative approach to group model building in a Scottish study of social care services for people with dementia. The paper concentrates on the process of eliciting material from a diverse group as part of the model building process, and describes a high-level stock/flow structure. The model's use as a simulation tool was not described.

Perhaps the most prominent recent example of the contribution of SD to the field of social work (as opposed to adult social care) is Lane's *Analysis of impact of increased prescription in social work* published as an appendix to the Munro Review of Child Protection (Department for Education 2011). The analysis explores the unintended consequences arising from attempts to improve practice through standardisation; in particular the impacts on staff morale, time spent working directly with service users and the scope for dealing with the variety of needs.
SYSTEMS THINKING AND SIMULATION IN SOCIAL WORK THEORY

The field of social work is somewhat receptive to systems-based approaches and simulation. During social work’s move towards professionalisation in the post-war years, systems theory offered a unifying general theory that moved beyond an explanation of social problems in terms of individual pathology (Pincus and Minahan 1973, Goldstein 1973). Arguably, this approach was superseded by more radical sociologically based theories that emphasised social conflict and disadvantage above the apparent cohesion that might be implicit in systems based approaches.

For all the permeation of systems ideas, SD never made its way into the social work curriculum in the UK. In the USA, building on the link between social work theory and systems theory, Robards and Gillespie (1980) suggested that SD modelling was the next logical step. From the same university*, Hovmand and O’Sullivan (2008) describe a graduate-level interdisciplinary course that brings together social work and engineering students. Such examples are rare.

As well as being receptive to systems concepts, social work routinely uses simulation as a learning tool; for example, the use of role-play in training as a proxy for working directly with service users. There are also examples of management learning through simulation (Harvey et al. 2011), but not in the sense of computer modelling. Wastell et al. (2011) have used computer simulation to study decision-making. Their Microworld simulator replicates computerised information systems that are now prevalent, but is not a dynamic (behaviour over time) model. This resonates with Parton’s (2008) observations about the changing nature of knowledge in social work, characterised as a shift from the “narrative” to a “database” way of thinking, which he views as linear. His plea for a way of thinking that is nonlinear, recognising complexities and connectedness, inadvertently describes some of what SD can offer. Wastell and Parton are both concerned with child care social work, but the challenges of assessing and managing risk are just as much a feature of adult social care.

* Washington University in St Louis
ELEMENTS OF, AND STRUCTURES IN, SYSTEM DYNAMICS MODELS OF SERVICES

In this section, we start by looking at the building blocks of models and then assemble these into some templates (sometimes thought of as molecules) conveying some generic phenomena that are used in modelling social care.

At the heart of any SD model is its stock/flow structure, accompanied by the feedback loops that drive flows. Model behaviour is determined by structure. A model consists of a diagram, and a set of mathematical calculations.

The stock and flow symbols (see Figure 2) are obvious, using a bathtub analogy, where the stock is the bathtub, the inflow is the tap, and the outflow is the drain. The cloud symbols represent material entering and exiting the model from outside the system, represented in the model (although it might form part of another model – for example where does bathwater come from and go to?). The valve symbols on the flows are used to specify what drives each flow. In order to do this, we use ancillary variables (or converters) and connectors, for example, as shown in the population diagram in Figure 1.

Figure 2 Stock and flows

![Figure 2 Stock and flows](image)

Population and ageing

Now, Figure 3 shows that because connectors have been drawn from population and birth rate to births, births must be a function of these two variables; and similarly deaths is a function of population and mortality rate. Additionally, we specify that there is an initial population of 50 million.

Figure 3 Simple population model

![Figure 3 Simple population model](image)
To work with the model*, we build a simple interface, where relevant inputs (here birth rate and mortality rate) are put on input devices (sliders in this case) and the main output (population) is put on a behaviour over time graph. The model is run for 80 years. We can then run a number of scenarios with our model (see Figures 4–6).

**Figure 4 Birth Rate = Death Rate (equilibrium)**

In this run, the birth rate and mortality rate are equal to each other (17 per thousand per annum), and the population remains a flat line, in equilibrium, with two feedback loops (births and deaths) effectively balancing each other (inflow = outflow).

**Figure 5 No Births (balancing loop)**

In this run, there are no births and the mortality rate is high. This shows the typical behaviour of a negative feedback loop, tendency towards a limit (or target), in this case 0.

To experiment with a web-enabled version of this model, and some other models featured in this review, please go to www.symmetricpartnership.co.uk/sscr.
In this run, nobody dies, and we easily observe the characteristic behaviour of a single reinforcing feedback loop (exponential growth). We have to redraw the scale of the graph, and the population nearly quadruples in 80 years.

One would not generally use such a simplified structure to represent population, since clearly births only arise from one section of that population, and death rates vary enormously by age group. (For a very detailed discussion of this see Eberlein and Thompson 2012.) We would use an “ageing chain”, a sequence of stocks and flows representing the population as a series of age bands (see Figure 7).
Figure 7 Ageing Chain of a Population

Figure 7 shows:

- The movement of the population through different life stages;
- At each stage, people either die or progress to the next stage;
- There are different death rates for each stage (not obvious from the diagram);
- Births are a function of only one stock, not the whole population (obviously, we could go into more detail, for example differentiating gender or clarifying exactly which ages are covered by our stocks).

One great benefit of being able to model the population dynamically is that a population module can be built into a wide range of types of model.
Service dynamics, capacity, length of stay and boundaries

Modelling service stages is a powerful application of SD that gives insight into, for example:

- Relationship between demand, occupancy and length of stay;
- Capacity planning;
- Waiting lists and waiting times;
- Delayed discharges;
- How to balance capacity across a whole system, especially where this crosses service boundaries.

The main dynamics of service use start from the premise that the number of people using a service (a stock) is equal to the inflow times the average (mean) length of stay. In SD, this looks like Figure 8.

Figure 8 Simple service showing referrals and length of stay

This model describes a simple service where people are referred, all referrals are accepted, and they remain for a mean length of stay. The behaviour of this model is quite straightforward, yet often service planners are not confident about the way in which a stock of service users is a function of referrals and length of stay.
If the initial number of service users is 100, the referral rate is 20 per month and average length of stay is 3 months, when the model runs the number of service users moves from its initial 100 until it reaches 60, where it remains in equilibrium. It does this because, if the outflow is stock divided by mean length of stay, at the point where the stock reaches 60, the outflow will be 20 (60/3), the inflow and outflow balance, the stock level does not change and the calculation remains the same for each successive time step.

If we had set the initial value of the stock at this level (60) the picture would be as in Figure 10.
The process operates similarly where the initial state of the stock is less than its equilibrium value, for example rising from an initial point of 30 in Figure 11.

Figure 11 Simple service output (low initial value)

Obviously, few services behave in this ideal way, owing to, for example, daily, weekly and seasonal fluctuations in demand, staffing variations, delays caused by relationships with contiguous services and the implications of capacity constraints. The introduction of a capacity constraint means that not everyone can be admitted at the point of referral (see Figure 12).

Figure 12 Simple service with capacity constraint
Service capacity (often, but not in Figure 12, represented as a stock) creates a balancing loop where the number of service users cannot exceed the capacity. Note that:

\[ \text{Spare places (now)} = \text{capacity (constant)} - \text{stock (now)} + \text{outflow (in the last time step)} \]

\[ \text{Starting} = \text{the minimum of spare places or waiting} \]

We already know that the capacity required for a service with 20 referrals per month and average length of stay 3 months is 60. So, if we set capacity at 50 (see Figure 13), the behaviour of the model should not surprise.

**Figure 13 Capacity model run: insufficient capacity**

Occupancy rises from its initial 30 to 50 quite rapidly. At this point, the number starting equals the new referrals (to waiting flow).

Once the service reaches capacity, the number starting is less than the new referrals, the service users stock flat-lines at 50, and a waiting list emerges, growing constantly.

Note that the flat graph of service users should not be confused with equilibrium, even although it might be observed as such; it is the inevitable result of a lack of capacity.

Increasing capacity to 60 or 70 has the following effects on the stocks graph (see Figure 14).
With capacity at 60, the model is now in equilibrium.

With capacity at 70, occupancy reaches 60, then flat-lines, and there are always 10 unused places (the gap between service capacity and users).

This relatively straightforward template can be deployed as a building block of a more complex model, where users of one service flow into another, and each service has its own dynamics set by inflow, capacity and length of stay, compounded by the problem that the waiting list for the second service comprises people whose only reason for still using the first service is that they are waiting to move on (i.e. the delayed discharge phenomenon).
The following model (Figure 15) is a “Microworld” (representing a theoretical, simplified, system used for training purposes). The challenge for the user of the model is to imagine they are in charge of a completely pooled budget for hospital and social care, where there is only one kind of hospital bed and all social care transfers come from within the hospital. People are admitted to the hospital, most are discharged home, a fraction will need care and move (still within a hospital bed whilst waiting) to the stock Delayed Discharges, where they wait until there is capacity in social care. Social care service users also move on based on a very long length of stay.

**Figure 15 Two capacities model diagram**

The model's capacity calculations effectively give a user of the model the chance to allocate a different proportion of the total budget to hospital or social care capacity, using this model interface illustrated in Figure 16.
The model is run for one year. The top graph shows the hospital. At first, the hospital capacity copes with demand and there are no delayed discharges. From day 130, the hospital is full and a waiting list develops, growing rapidly. At around the same time, some delayed discharges appear. The bottom graph shows the care system, which reaches capacity a bit earlier. Note that the delayed discharge line is shown on both graphs.

The hospital is full and there are delayed discharges. Could the problem be resolved by increasing the hospital capacity? Instead of allocating 88% of the budget to the hospital, increase that to 90%. This is the result in Figure 17.
Figure 17 Two capacities model: allocate more capacity to hospital

Clearly, giving more capacity to hospital makes matters worse, because resources have been taken away from the care system, resulting in a delayed discharge problem occurring earlier.

Another possibility might be to increase the budget, from £20m to £22m, while continuing to allocate 88% to hospital (Figure 18).
As can be seen this merely postpones the problem.

As is probably obvious, the problem in this system arises from insufficient capacity in the care system, resulting in delayed discharges and a full hospital with a waiting list. With a finite budget and a hospital becoming blocked, it may take some boldness to disinvest in the hospital.

Eventually, we get round to taking capacity away from our blocked hospital by reducing its allocation of the budget to 85% and investing in more social care (Figure 19).
Now the system is in some kind of balance. Because there are no delays of people waiting to get into the social care system, the hospital occupancy remains (only just) within capacity.

In reality, we would not want to run services at that level of occupancy and there would be a greater level of complexity (not just two capacities and a single flow). The point of this modelling, however, is to illustrate how sequential capacity-constrained services require joint planning because a change in the dynamics of one will have a consequential impact on the other.
Hospital and social care model with greater complexity

The model structure shown in Figure 20, while still offering a high-level view of things, captures a more complex range of dynamics.

Figure 20 Hospital and social care capacity model with two types of admission
In this model, variants of which have been used in real applications (Wolstenholme et al. 2007a), hospital beds (and admission types) are split between emergency and elective. A fraction of the emergency admissions require discharge to social care. In this instance, if social care is blocked, and delayed discharges result in emergency beds being fully utilised, one informal coping strategy is to board out emergency patients in elective beds. This might work for a short time but, if elective capacity is fully utilised, the result is cancellation of elective procedures and an increase in the elective waiting list. Given the typically much longer lengths of stay of emergency patients, the problem is compounded as elective beds are filled for a much longer time per patient.

As with the simpler examples outlined in Figures 14 to 18 it is possible that the solution to an apparent capacity shortage in one place is to increase the capacity somewhere else in the system. The model provides a means by which those engaged in joint planning of services can explore, in a risk-free environment, the possible outcomes of apparently counterintuitive policies.

Including variation, for example when modelling capacity

The capacity graphs shown so far consist of mostly smooth lines and depend on inputs such as an average length of stay, a simplification that makes the relationship between system structure and behaviour clearer.

SD is probably more often deployed to model strategic or whole population problems, but, as already noted (in section headed What kinds of problems are not amenable to a SD approach?, page 3), there are some problems, such as capacity planning, perhaps more operational than strategic, where SD is one of several possible simulation approaches. A key question is how to model variation, such as the uneven distribution of new arrivals entering a process. Other simulation approaches typically deal with this by randomly sampling from a distribution.

In SD the questions are “where does this apparent ‘randomness’ originate?” and “might it derive from something happening within this system?” For example, the following (unpublished) model of a tertiary general hospital (Figure 21) was designed to facilitate better planning for winter pressures. It shows the interplay of some processes that introduce (non-random) variation into the whole system, along with an exogenous demand function.

There are three main flows:

- Emergency medical (mainly via A&E);
- Emergency surgical (mainly via A&E);
- Elective (from waiting list).
Modelling social care complexity: the potential of System Dynamics

Figure 21 Whole systems model of a general hospital for winter pressures planning
Each flow goes through different stages, each capacity-constrained:

- Emergency medical pathway goes through admissions ward, short stay ward then medical ward; some patients then require discharge to a social care placement or community hospital; patients can be discharged at any stage and some potential medical admissions go straight to community hospital (admission avoidance);
- Emergency surgical pathway through surgical admissions ward then surgical ward;
- Elective patients are admitted directly into surgical or specialist wards.

There are multiple sources of variation, most of which are not visible in the diagram:

- Arrivals at A&E vary using an external but patterned random function;
- Surgical and specialism capacity, complexity arising from the same beds accommodating multiple flows, including elective and emergency surgical/specialist admissions and boarders when medical beds are full;
- Elective surgical and specialist admissions are based on a 5-day per week working pattern;
- Outbreaks of infection, such as norovirus, can occur, in which case the hospital adopts a policy of closing affected wards to new admissions (the model includes an embedded SIR model, susceptible-infected-recovered as explained in the next section);
- Winter pressures, where emergency admissions increase for several days at a time owing to cold weather, influenza outbreaks, or both;
- Transfers to social care are batched to some extent, because they depend on decisions made by a fortnightly funding panel (meanwhile, people wait in hospital).

In combination, these phenomena have an apparently random effect on the performance of the whole system, but illustrate, for example, why elective admissions, which one might expect would exhibit the least variation (being buffered by a waiting list) actually show considerable variation, as elective capacity is taken up by emergency medical boarders at times of pressure. The model also shows that modelling hospital capacity based on a simple average daily demand times length of stay function might not work.

Typical behaviour of this model is shown in Figure 22.
“Target occupancy” is 85% of bed capacity, above this level a hospital does not cope well. A subset of occupancy is delayed discharges which follows a fortnightly pattern as decisions about funding are made every fortnight.

In the initial quarter, occupancy fluctuates around the 85% mark. The variability is due to a combination of factors, a variable number of admissions per day combined with some of the delayed discharges.

Total occupancy of surgical beds is the sum of elective admissions, emergency admissions and boarders (medical patients overspilling into surgical beds).

In the initial quarter the total reflects the weekly fluctuation resulting from elective admissions five-day working. With occupancy approaching 100%, note that elective admissions decrease when emergency admissions rise.

Medical Patients

At time A, there is an outbreak of norovirus in a medical ward, which spreads to another ward ten days later. Note that the occupancy line goes down (because beds are not being filled in these wards – once empty, wards are deep cleaned and capacity is restored).

At time B, there are winter pressures; cold weather precipitates ten days with higher admissions. Medical beds, no longer affected by norovirus, fill to capacity and medical boarders spill over into surgical beds.

Towards the end of the second quarter, normal occupancy is restored. But the number of delayed discharges has been creeping up. This eventually leads to...

At times C and D, medical capacity is once again stretched, this time because of the impact of delayed discharges combined with the “normal” fluctuation in admissions. This leads to two more periods of full medical beds and boarders placed in surgical wards.

Surgical Patients

At time A, because medical beds have been closed, medical boarders spill over into surgical beds. Note that elective occupancy falls as operations are cancelled because of medical boarders occupying surgical beds.

At time B, following a brief drop in boarders (all norovirus cases having ended) before the line rises, this time because of winter pressures.

Normality is restored.

At times C and D, once again, boarders overspill into surgical beds and for two distinct periods in the third and fourth quarters elective admissions are reduced.
Modelling need or dependency states

Perhaps the most powerful applications of SD use stock/flow chains to represent different states of the population in relation to some kind of need. Most examples of this in the literature cover states of ill health, with different types of structure for acute, long-term, and degenerative conditions, often also taking ageing into account.

The best examples of modelling an acute state are infections and epidemics using Susceptible, Infected, Recovered (SIR) approaches. A typical SIR model consists of the stocks and flows in Figure 23.

**Figure 23 SIR Stocks and Flows**

The model represents the whole population (of a locality, region or nation). Initially, everyone is in the susceptible stock. As people move towards being infected, so a reinforcing feedback loop comes into play; as more people are infected there is a greater probability of susceptible people coming into contact with them. This positive feedback is balanced by the recovery rate. The challenge is to manage the epidemic in a way that reinforces the balancing loop and minimises the reinforcing effect, as shown when the feedback loops/ancillary variables are added to the diagram (Figure 24).

**Figure 24 More detail in the SIR model**
The typical behaviour of this model can be shown in the sequence of graphs in Figure 25.

Figure 25 Runs from the SIR model

The higher the level of infectivity or frequency of contacts between infected and susceptible persons, the more likely it is that everyone will become unwell.

By reducing the contacts per day we realise an outcome where the spread of infection stops before everyone gets infected.

In this run, a strict quarantine policy was implemented ten days into the epidemic which, in this case, results in only half of the population being affected.
However, social care is probably more significantly affected by chronic or long-term conditions. A commonly-used template (owing much to the work of Homer et al. 2007) shows the whole population in relation to one condition (Figure 26).

**Figure 26 Long-term conditions stock/flow template**

In Figure 26, new incidence is a flow of people from the healthy population to the simple stage. In this case, the condition goes through two stages and is degenerative (i.e. it is impossible to flow back in the direction of recovery). At first the condition is undiagnosed, probably because the person is asymptomatic or has not sought help. If diagnosed at the simple stage, the illness can be managed and degeneration to the with complications stage can be delayed (but not stopped). The model shows that the challenge is, either through screening for asymptomatic cases or case-finding of those who feel unwell but have not sought help, to get more people diagnosed at the simple stage. An additional challenge might be to encourage the adoption of healthier lifestyle choices and slow down the incidence flow.

Typically, this type of stock/flow chain will be at the heart of a more complicated model that includes a detailed representation of a range of treatments, treatment capacity constraints, and the typical use of acute services arising from people with the condition. The behaviour of such a model is unpredictable. For example, maintaining more people at the simple stage can reduce the number of unplanned hospital admissions. Equally, slowing down the rate of degeneration will also increase the overall unwell population resulting in a greater number of people in need. One cannot always assume that early detection and intervention will produce a saving.

Other examples of long-term conditions that can be modelled include aspects of mental health (Todd et al. 2006).
Other building blocks (finance, workforce)

SD models can readily incorporate financial modelling because money is just one more thing that flows over time.

At a very simple level, the cashflow for an individual or an organisation can be represented as in Figure 27.

**Figure 27 Simple cashflow**

![Simple cashflow diagram](image)

Financial calculations can be built into a model by attributing costs, or prices, to elements of the model as it runs (with compounding or discounting calculations where appropriate). Spending is a flow and stocks can be used to track expenditure. For example, flowing spending through a stock with a length of stay of one year enables tracking of a year’s worth of expenditure, while flowing into a stock with no outflow gives a measure of cumulative spending throughout the duration of a model run.

With the move towards payment by results (PbR) as a basis for paying for health and some other services, a model can show how some cash flow depends on a stock, and other cash flow depends on a flow. From the point of view of a service provider, expenditure is largely determined by its stocks, especially workforce. Whereas its revenues might previously be made on the basis of some kind of block contract, under payment by results revenue will depend more on flows of service users.

In very simplified form, a PbR model from the perspective of a provider looks like Figure 28.

This simple model has three types of flow: service users, staff and cash. In order to operate, the provider must employ a stock of staff. Its service capacity (the total number who can use the service at a given time) is a function of its staff multiplied by the maximum caseload. Given that capacity, service users can flow in, use the service and flow out again.

Coming to cash flow, in a pure PbR model, payment would be made on completion of treatment (discharges). Income is therefore a function of completed treatments multiplied by the payment per treatment. Expenditure is a function of staff in post multiplied by average salary (plus overheads).
Note also that, in formulating a model, one would ensure that most types of stock cannot go negative; for example, it is impossible to have a negative number of people. In many of the templates previously explored, the outflow will be a function of the stock (at its simplest, stock divided by length of stay). But money represents a type of stock that can take a negative value and it can be seen in Figure 28 that expenditure is not a function of the stock from which it is paid (unless the overdraft limit is reached, but that is not shown).

This kind of model structure is useful to providers (and commissioners) implementing PbR schemes, which are now being applied to services, such as community mental health services, that can be characterised as having very low rates of flow in relation to the size
of their stock of service users (as people tend to use services for a long time).

**Workforce planning/modelling**

Workforce planning is an obvious area of application for SD modelling, which has hitherto often been done using large spreadsheet models*.

The challenge in workforce planning, especially for occupations such as social work requiring professional training, arises from the long lead-time between the recruitment of students and graduation (which might mean graduation into a particular career structure).

Workforce stock/flow chains will tend to be variations on a theme, whereby the supply chain of new entrants (minus those who drop out) feeds into a career chain characterised by some combination of ageing alongside career progression.

In social care this can be particularly messy, comprising, as it does, some job roles with specific pre-entry qualification requirements (such as social work), other positions that may be held by people with one of a range of qualifications, and others where the imperative is to secure, by secondment or on-the-job training, qualifications for a largely unqualified staff group. Moreover, social care employs a range of staff who may have qualified in a different sector (for example, nurses).

A major benefit of SD is that it can cover both supply and demand issues within the same model. There is a structure of posts (not people) that may change as service delivery, or regulatory requirements, change and there are stocks of people moving through qualification or career stages.

A good example of this from the health sector concerns (unpublished) work undertaken for the Department of Health and the Royal College of Paediatrics and Child Health to model the paediatrician workforce in the light of the requirements of the European Working Time Directive, combined with a possible reconfiguration of hospital/delivery units of paediatric services (see Figure 29). In this example, the left hand sector represents trainees, and the right hand sector is consultants. ST refers to the three stages of specialist training, implicitly taking 8 years but in reality taking an average of around 11 years. After ST5, trainees go into three branches (actually, the specialists category covers a wide variety of possibilities). The consultant sector is modelled in relation to age groups. At any stage, people can drop out. One of the main aims of the work was to examine the implications of the European working time directive, which would limit the hours that could be worked by middle-grade doctors (ST 4 – 8 in the model), thereby increasing the demand. As is obvious from the model, because consultants are the product of the training supply chain, it is not possible to increase the number of middle grades without also, in due course, increasing the number of consultants.

The recently established Centre for Workforce Intelligence, whose remit extends to social care, proposes to make significant use of System Dynamics in its workforce modelling.
Putting these components together – studying the impact of interventions over time

We have seen that SD is a good method for examining different types of flow, such as:

- Population in need;
- Workforce;
- Other service capacities (for example, beds or care packages);
- Money.

A modelling project should start with a problem, a question, or an issue that can be framed in dynamic (impact over time) terms. A good example is the Alcohol Systems Model (McKelvie et al. 2011), which explored the impact of brief intervention types of service on alcohol-related harm, measured in terms of alcohol-related hospital admissions in England. At the time of writing, the model is available to download* (after a simple registration to the system).

* www.alcohollearningcentre.org.uk/Topics/Browse/Data/Datatools/alcoholSystemsModel/
The presenting problem is the high number of alcohol-attributable hospital admissions of all kinds. The hypothesis is that by providing more brief interventions to people either in primary care (for example, opportunistic advice giving by GPs) or secondary care (for example, following a hospital attendance or admission which might be alcohol related), people can be persuaded to moderate their consumption of alcohol (note that these interventions are not primarily concerned with the higher-end problems of addiction or anti-social behaviour).

The model is summarised in Figure 30. At the heart is the population consumption chain, where an entire population (typically a local authority area) is in one of four consumption stocks, from abstinent to high consumption. As the model runs, there is a natural, or untreated, rate of flow between each stock. People attend hospital with alcohol-attributable problems in proportion to their consumption. If, as a result of brief interventions, a proportion of people in the two high consumption stocks move to the lower consumption stocks, after a time lag (because health improvement does not immediately follow a change in consumption), there will be a reduction in hospital attendances.

**Figure 30 Overview of alcohol systems model**
So, are these services affordable? Do the reductions in hospital attendance produce savings that are greater than the cost of the new service? A main finding of the model, because of the time lag between treatment and health improvement, highlights that the cost of the whole system (hospital admissions plus brief advice services) will initially be greater, as illustrated (Figure 31).

Each graph in Figure 31 plots a single variable (change in admissions, change in costs) over four scenarios (1 do nothing, 2 brief advice in primary care, 3 brief interventions in secondary care, 4 combine 2 and 3). Note the position of 0 on the Y axis. The model runs for 20 years and the intervention point is at 5 years.

**Figure 31 Alcohol model output showing changes in admissions and cost**

The top graph shows that there is a delay of around six months before hospital admissions begin to fall (eventually reaching a new equilibrium).

The bottom graph shows that from the intervention point there is an immediate rise in costs which peaks after about one year, and becomes negative (overall current saving) within two years. It actually takes about three and a half years before the initial investment in the new service is recovered by the saving (there is a cumulative change in expenditure graph in the report, not shown here).
The relevance of this model to social care is that it provides a generic template for applying SD to a wide range of human service cost-effectiveness problems (see Figure 32).

**Figure 32 Generic population – intervention – impact modelling**

- There is a population of concern (people in need for a particular reason, ageing, dementia, disability, health, substance misuse);
- There is a presenting problem (either overuse of a mainstream service, which costs money, or perhaps underuse, because they have not been connected/referred);
- A new intervention is proposed; it also costs money;
- The new intervention has an impact on the population, so they now use mainstream services differently (perhaps less, in the case of the alcohol model, perhaps more, as we shall see in relation to dementia);
- It is possible to measure the quality of life of a population, for example, by multiplying each stock/state by a quality of life index;
- It is possible to measure changes in quality of life in relation to changes in spending.

System dynamicists would argue that this approach resembles the creation of a glass box model, within which the causal links effecting change are visible and explicit, by contrast with a stereotypical econometric black box approach where the correlation and regression calculations are more impenetrable.
From the point of view of a social researcher, the difference is between perhaps the more familiar world of correlations and statistical significance, to a representation that is operational and, importantly, causal.

In this section, we have moved from:

• Looking at SD building blocks, to
• Considering some simple structures (populations, services), to
• Making some simple models using these structures as building blocks, to
• Exploring the dynamics of populations, interventions and costs, to
• Considering a generic framework.

In the next section, we will consider another SD model, of dementia, paying particular attention to its interface (i.e. how it can be used).
USING A MODEL – AGEING AND DEMENTIA

Typically, models of the ageing population will consider the impact of policies designed to support people living at home and delay admission to residential care. A simplified stock/flow structure is shown in Figure 33.

As people become frailer, they reach a stage of needing support to continue living at home. If home care is provided, this increases coping capacity, reducing the rate at which people become no longer able to live at home (alternatively, increasing the length of time to deteriorate).

The various feedback effects can be shown in a series of causal loops (Figure 34).
This reinforcing loop shows that if there is more home care service capacity and more people who need care, this results in more service users. The impact of service use is to increase the deterioration time (length of time for which people manage at home before needing residential care). Therefore there will be more people who need care at home.

Note that people who need care at home denotes people who need home care, whether or not they receive it.

Mortality itself is a balancing loop.

But it will be influenced by the following reinforcing loop.
Another way in which providing home care might increase the number of people who use it, is its impact on the mortality rate of service users. Assuming that the service has a negative impact on mortality rates (for example, by reducing service users’ propensity to fall, or have accidents, or other acute episodes) people receiving care will live for longer, thereby increasing the stock of people who need care at home.

The transition from needing care at home to needing long-term care, for example in a care home, is another balancing loop. As people who need care increase, the number moving on to the next stage (needing residential care) rises. But as this number rises, so the number needing care at home falls.

The final diagram in this sequence (Figure 35) puts all those causal loops together in one diagram.
Ultimately, what distinguishes System Dynamics from Systems Thinking is its insistence on the value of exploring complex problems by using computer simulation.

**Software**

A number of commercially available software packages have been developed by the SD community. To a large extent, they all do the same thing (using mathematical integration to calculate the changes in stocks, as they are determined by flow rates, moving forward...
in small time-steps). The choice of software package will be determined by a range of circumstances, most powerfully perhaps the context in which a user first encounters SD. Each also offers a range of options for making models available to people who want to use them, without needing to know how to build a model.

It is not possible to provide an exhaustive list of software options, but the field is still limited to a relatively few providers. The picture is always changing; the interested reader should put “system dynamics software” into a search engine and explore. Alternatively, the System Dynamics Society website (www.systemdynamics.org/) has links with the main software vendors.

The main packages, in alphabetical order, are iThink/Stella*, Powersim, and Vensim. In addition, there are packages, such as AnyLogic, which are capable of running different types of simulation (for example, agent based modelling, or discrete event simulation). There are also packages developed in a particular context, such as Simile (environment and land use). Until recently, there were few examples of SD models being made available on the web, probably because practitioners are mainly concerned with building models with and for an identifiable user. More recently, Forio has emerged as a platform on which models originally developed on one of the main applications can be accessed using a web browser. Insight Maker and Sysdea offer the ability to build models via a web browser.

*Most of the model diagrams in this review have been developed on iThink.
EXAMPLE OF A MODEL INTERFACE AND MODEL BEHAVIOUR: DEMENTIA

A System Dynamic model has three levels or components:

- A stock/flow diagram, in which every element is a model variable and connections between variables are made explicit;
- A list of variables and equations (compiled by the modeller on the model diagram);
- An interface, being a set of computer screens on which:
  - Inputs to key values are made using a range of devices, such as sliders;
  - Outputs from model runs are displayed, mainly as behaviour over time graphs, or tables;
  - Other buttons enable a user to run, pause, stop, save, or view the model, navigate between screens, perhaps import variables from or export model outputs to, a spreadsheet;
  - Depending on the software package, there are a variety of additional functions to enrich the user experience, such as switching on or off particular modules, or storytelling.

Generally, it is possible to make a runtime version of a model where a user has access to the interface but cannot change the underlying model structure.

As an example of a model interface covering an issue of central concern to the social care sector, we draw on a model of dementia (as yet unpublished) developed by a collaboration between health and social service providers in the East Midlands of England (see Figure 36).

The purpose of the model was to consider some of the health and social care capacity requirements of implementing the England National Dementia Strategy for older people with dementia. The model was built using a Group Model Building approach (outlined in the next section), with a group drawn mainly from the NHS, local authority and voluntary sectors. A particular focus was to consider the staffing requirement to run a local memory service and assess the impact of that service on a range of other factors, principally the implications for other services.

The three main stocks represent the entire population of an area (typically local authority) of older people with dementia (prevalence). Unlike most treatments of dementia, which use the health defined classification of severity (mild, moderate, severe), this model used increasing dependency from independent (living at home not in need of mainstream services), at home needing support (living at home, needing mainstream services, such as care package) to in care home. Whereas in the mild, moderate and severe chain all incidence is initially mild, with the dependency chain, incidence could occur to someone already quite dependent for reasons unconnected with dementia. Similarly, mortality can
happen at any of the three stages, i.e. people die suddenly without ever becoming dependent, but there is an inevitable movement towards greater dependence.

As people use a range of services, so their trajectory along this degeneration chain is affected and there are various feedback effects. For example:

- By having better diagnostic services (memory services), more people might be diagnosed at an earlier stage;
- Memory services provide access to medication, which might extend the period of time someone can manage living at home;
- Diagnosis at a memory service might increase the rate at which people are referred to mainstream services, such as home care packages, or might improve the nature of the package provided (if it is know that a service user has dementia);
• This greater uptake of mainstream services might also extend the length of time spent living at home;
• Some other services, such as admission avoidance services for people living at home or care home support, can reduce acute hospital admission rates of older people with dementia.

As can be seen from Figure 34 (in which a reinforcing causal link is shown by a curved arrow and a balancing link by a curved arrow with a dotted line), each part of this model is connected to each other part. This interaction between population need, service use and impact of services on the population or on the use of related services generates some complexity, and the behaviour of the system cannot be deduced from the diagram alone.

The model interface allows us to vary a wide range of assumptions about incidence, initial prevalence, service capacity and impact, and envisage multiple scenarios (generally by either switching services on or off, or varying their capacities and impacts). In addition, every service shown in this model is costed, so the impacts of different scenarios on health and local authority expenditure can also be shown.

The interface for this model takes the form of a dashboard or home screen and a range of other screens for inputting assumptions about different services (see Figure 37).

Figure 37 Dementia model interface: the dashboard
The dashboard includes output graphs (with two on display at any time, but there are multiple pages of graphs available, accessed by clicking on the bottom left corner of each graph pad), buttons for running, pausing and stopping the model and, at the bottom, a series of buttons that link to other screens on which service settings and assumptions are entered along with switches that enable a user to control whether or not to deploy these services. Therefore, a range of possible scenarios can be constructed by changing the data in the service screen inputs, to be run individually or in combination with others.

The main outputs of note are the total local authority spend (care packages, plus care homes) and total health spend on older people with dementia. In both cases, this is rising in line with population increase. The breakdown of population is shown on the bottom left graph (a user of the model can click on view detail to get access to more detail); the small graph on display here shows the total number of older people with dementia, underneath broken down by dependency group, all of which are growing.

An obvious starting point is to switch on memory services and medication, as two components of the service that ought to be at the heart of diagnosis, treatment and signposting. The graphical outputs from this model run are as shown in Figure 38.

**Figure 38 Dementia model run with memory services and medication switched on**

![Graph showing local authority and NHS spending over years.](#)

The scenario is shown on each graph in red; blue is the base case. Not surprisingly, there is an increase in health expenditure because an additional service has been commissioned. It also appears that total local authority spending is unchanged (lines show little difference).

However, drilling down into the detail of local authority spending, there are a further two graphs showing spend on care packages and on care homes (Figure 39).
Although overall local authority spending is unchanged, there are now more people receiving care packages at home and fewer people in care homes.

This is for a variety of reasons. Firstly, people diagnosed at memory services will have a higher uptake of care services. Secondly, the increase in people receiving medication and using home care mean that the length of time spent living at home is extended, i.e. admission to care homes is delayed.

This powerfully makes the case for joint planning. Expenditure in one sector (health) can have repercussions for another (local authority social care). These kinds of dynamics are replicated through a wide range of scenarios that can be run on this model.
THE MODEL BUILDING PROCESS AND GROUP MODEL BUILDING AS A PARTICULAR APPROACH

Forrester (1991), the founder of System Dynamics, outlined concisely how models should be built:

Model building should be a circular process of creating a model structure, testing behaviour of the model, comparing that behaviour with knowledge about the real world being represented, and reconsidering structure. During the process of modelling, the system dynamicist should always be alert to new discoveries about behaviour. The new discoveries may relate either to the particular system being studied or to the general nature of systems (p.26).

Just as there is no standard length for a research study, there is no ideal duration for a model building project. The key principle is that the purpose of a model is always to improve our ability to think about a complex issue. Modelling is always an iterative process involving a dialogue between the model and an expert group, including a cycle of building, testing and improving. It is possible to build a serviceable high-level model in a single session, or to encounter a pre-prepared Microworld-type model as part of a training event.

System dynamicists have devoted a great deal of attention to methodology and the idea of group model building provides a practical way forward. Proponents of group model building are aware of both the advantages but also the dangers of group dynamics and decision-making (Vennix 1999).

Richardson and Andersen (2010) go into some detail about the roles that are needed to achieve successful facilitation of group model building, including the facilitator/ knowledge elicitor, modeller/reflector, process coach, recorder and gatekeeper, while noting that experienced group modellers can normally get along with just two facilitators.

A typical group model building project requires:

- The formation of an expert group comprising people who are familiar with the subject matter. For a typical project involving health and social care services that will mean a group containing people who have an operational understanding of the range of services/interventions. The expert group should have a familiarity with the nature of the population and issue of concern (whether that is ageing, deprivation, mental health);

- At least one member, or access to someone, who is familiar with the range of data sources/information systems in which data might reside. Although it is important to emphasise that the process of model-building does not begin and/or end with data. The model will probably draw on both data from academic research, as well as secondary analysis of local datasets;
Facilitators who, between them, have both some understanding of the domain and, obviously, knowledge and skills in model-building.

There should be scope for a series of iterations – around five is a minimum but it is difficult to be too prescriptive. Not least, the building of a model can trigger awareness that a larger enquiry is necessary; as the work progresses the model boundaries can be extended. Nevertheless, a group must resist the temptation to model everything in great detail. The best models focus on a specific problem and work out from it.

A typical sequence of meetings will include:

- An introductory session, where participants are given some background knowledge about the modelling approach. The problem issue is introduced and described and some initial stock-flow and causal maps are generated. This is similar to process mapping but generally at a more aggregated, higher-level view of each process;

- A second session where perhaps an embryonic model based on the maps generated at the first meeting is presented. Moving quickly to simulation can accelerate learning, even if the actual model at this stage lacks detail. After this meeting the modeller will go away and aim to build something that begins to resemble the model;

- A third session where the model is presented; this will likely lead to a discussion that covers:
  - The need for some amendments to the model structure;
  - Evidence that the group is beginning to learn something new about the issue;
  - A discussion about data and data sources;
  - Identification of a range of scenarios that might be tested;

- At least two more sessions, in between which further work is carried out by the modeller and different group members, including a search for data/evidence.

Normally, a group will not manage to meet any more frequently than monthly (based on experience of how much time senior managers and frontline staff can afford to devote to, and meet about, any issue) – half-day sessions are probably sufficient. Any less, and not enough material is covered, but spending a whole day on model-building, requiring detailed work and plenty attention, can be draining. Achieving a consistent group attendance makes a difference because this avoids the need to engage in lengthy re-explanations of previous material in order to bring newcomers up-to-speed. Against that, if a group becomes too exclusive there can be a risk of “groupthink”, where the group adopts a particular interpretation that becomes the received wisdom.
In an iterative learning cycle, it is possible to reconsider and revisit anything from one meeting to the next. However, as the project progresses the focus of attention moves from:

- Model structure and gaining an understanding of the problem/issue;
- As the structure is consolidated, so a more detailed conversation about data and evidence is more possible;
- Estimation of data parameters tends to come towards the middle of the process (unless a project has been triggered by some problematic observations in the data); the kinds of data sought will include:
  - Population data about levels of need and, more difficult, the dynamics of need – so not just how many people need something but what is the rate at which people enter into this state and exit from it;
  - Service capacity and throughput data;
  - Data about interventions and their impact;
  - Financial data, mainly the costs of services and interventions;
- As the model moves from being a mere diagram to a functioning simulation, the focus moves on to an exploration of possible scenarios;
- At this point, to use Wolstenholme’s phrase, the group experiences a move from “speaking to” the model to have the model speaking to them;
- The issue of how the model is to be used beyond the immediate project will become more prominent; the shelf-life of some models may simply be for the duration of the project, other models might be intended for continued use by a client or even distributed.

This account reflects the author’s own experience of using SD mainly within consulting engagements, where a typical client would be a service commissioner or government agency/department. Given the low level of use of SD within the social care domain, and the lack of published research, it is difficult to outline precisely how SD might be used within longer-term research studies. The key requirement is always to secure client engagement in an iterative learning process. Lane et al. (2003) outline a slightly different approach in a fast-moving setting where it would have been difficult to negotiate time out to engage a clinical team in a more structured group model building process.

To the extent that social care researchers undoubtedly do already use models, normally qualitative approaches such as influence diagrams and logic models, and given the parallel preoccupations of fields such as health economics (concerned with understanding the impact of interventions and expressing that impact in terms of costs and improved quality of life), it is perfectly possible to envisage that SD could make a positive contribution to
any study with some of the following characteristics:

- There is an interest in population need and how it might be met;
- There is an intervention, the impact of which is to be evaluated and understood;
- There is complexity arising from interconnectedness, for example there would be a feedback loop involving the intervention responding to a demand and changing the nature of that demand;
- There is complexity arising from a service system that crosses agency boundaries, where a change in an intervention in one sector might have a bearing on another;
- There may simply be a need for a more dynamic understanding of a simple capacity issue;
- The issue of cost must be understood within that complex network; the policy-maker’s dream of “invest to save” might actually turn out to increase demand somewhere else, in the short-term at least;
- There are long time delays, such as exist in workforce chains, or before treatment effects are shown.
A WORD ABOUT DATA

It is often said that a model is only as good as the data going into it. Most system dynamicists would challenge that assertion having learned from experience that if any enterprise depended on the availability of perfect data little would ever get done.

While it goes without saying that we want to use the most accurate data available, the modelling process will typically highlight gaps and inaccuracies in official data while providing a specification of a set of data that would be really useful.

Data collected for performance-reporting purposes may not be so useful for planning purposes. A good example is length of stay in hospital. Typically, health organisations can provide a performance measure based on a trimmed median calculation. That is fair enough – it avoids judgements of performance being made based on extreme cases (statistical outliers) that do not reflect the performance of the organisation in question. However, as a basis for planning, we need the real average length of stay, a number that is captured by the mean, a summary measure that is influenced by outliers. Why? Because unless we are planning a new system in which there will be no outliers (begging the question of how that will be achieved) we need a realistic measure of turnover.

Sometimes the structure of a model provides much more information than any data could. Consider the workforce planning problem that arises from the implementation of the European Working Time Directive as applied to doctors. A simple workforce diagram shows how doctors move through three stages (stocks), junior, middle grade and consultant (Figure 40).

**Figure 40 Simple medical workforce chain**

With new restrictions on the number of hours that can be worked, the demand for middle-grade doctors rose because of their crucial contribution to on-call rotas. This meant, in effect, that there would be a new target requirement for middle-grade doctors...
but perhaps no change in the number of consultants required. It is immediately clear from this diagram that this thinking is flawed. The reason is that future consultants are a function of present-day middle grades. It is impossible to generate more middle-grade doctors without producing more consultants in a few years. The point is that it is possible to have this discussion based on the structure of the model diagram alone; we need no data to gain the insight that setting independent targets for two stocks in the same supply chain is impossible. An insight can be gained before any data are collected. The quality of the model does not depend entirely on the quality of its data inputs.

Nevertheless, a dynamic modelling methodology will inevitably be hungry for dynamic data. Typically, we want to know not just how many of x? but more challengingly, where were people in x this time last year? How have things changed? Of people entering this process, in which proportions will they exit to the various onward-movement options? And sometimes, a model will function as a lie detector, highlighting that reported combinations of official data items cannot possibly co-exist in reality (for example data about referrals, duration of service, service caseload and onward-movement).
CONCLUSION

This review builds on an earlier SSCR general review of mathematical modelling by presenting one method in more detail. SD provides a way of synthesising diverse types of evidence, enabling multi-disciplinary groups of stakeholders to explore a range of future scenarios in a risk-free environment. Its use of stock/flow and causal loop diagrams combined with user-friendly dashboards, and its suitability for use in a group learning environment, provide an accessible introduction to modelling for non-mathematicians. The review outlines when such an approach should be considered and points to some of the problems that are amenable to such a study. The literature contains many examples of applications in health; social care examples exist predominantly at the interface between health and social care. Nevertheless, social care and social work theorists have drawn on qualitative systems ideas, and grapple with complex non-linear problems, so might be receptive to a modelling approach based on these.

Some of the smaller models described in this review will shortly be made available in web-accessible format at www.symmetricpartnership.co.uk/scr.
REFERENCES


Lane DC (2001a) Rerum cognoscere causas: Part I – how do the ideas of system dynamics relate to traditional social theories and the voluntarism / determinism debate? System Dynamics Review, 17, 97–118.

Lane DC (2001b) Rerum cognoscere causas: Part II – opportunities generated by the agency/structure debate and suggestions for clarifying the social theoretic position of system dynamics, System Dynamics Review, 17, 293–309.


Modelling social care complexity: the potential of System Dynamics


Further reading