Simulating Dynamic Systems in Health Psychology

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Arts

in

Psychology

at Massey University, Palmerston North,

New Zealand

Susan Elizabeth Street

2001

Abstract

Despite their advocacy of the biopsychosocial model, health psychologists use a relatively narrow repertoire of techniques for developing and testing theory. These techniques have limited application to research questions concerning phenomena that are multidimensional, multilevel and change over time. This thesis demonstrates an alternative, dynamic systems approach to such questions in health psychology. It introduces some ideas in systems and dynamics and how we might model these. It uses an example to demonstrate the use of these ideas to develop a dynamic systems model in a health psychology context. The example is drawn from the epidemiological finding of a positive correlation between income inequality and mortality, and the proposal that this relationship may be mediated by processes that result in social disruption. The thesis explores the construction of a dynamic systems model to examine how a change in income inequality might affect the network of social relationships in a population. Social relationship processes in the model are based on some findings from social psychology, and these are incorporated into a network model, which is realised as a computer simulation.

Simulation runs suggested that an increase in income inequality can produce a ripple of relationship breakdowns. Contrary to intuition, the number of relationships lost was limited if the change was introduced suddenly, and if there was a high rate of making and breaking relationships. Further, reversing the change did not reverse the loss of relationships. The development process and the results obtained are discussed, and it is argued that dynamic systems simulation may be useful for developing and testing theory that applies to multilevel, multidimensional processes in health psychology.

Acknowledgements

The ideas presented here arrived very quickly. Their refinement and development into a thesis was a much more demanding task. John Spicer helped me greatly with this process. Between the rarity of methodological theses and my attempts at translating ideas foreign to psychology, I needed to draw heavily on John's wisdom to guide me. His advice was delivered with such a gentle hand that I often didn't notice it had happened, except that things suddenly seemed to work. Having given his advice, he let me find my own way. As a result, I cannot hope to identify everywhere that he has influenced the final product. But he did make some rather fundamental contributions, including pointing my writing in the direction that the rest of me was going. John has been generous with his time, extraordinarily patient, consistently good humoured and unfailingly kind. It has been a delight and a privilege to work with him.

Mike, Mum, and Ralph all helped by reading early (unreadable) drafts, providing comments and saying the right things. Angela read and commented on parts of a later revision. She was absolutely right, it does go better in this order. Rach helped with advice on layout. My family, Gamelan Padhang Moncar and friends all kept me going.

This thesis depended on my having access to a computer and software. This was provided by Mike, as was everything else for its material production, my upkeep and intermittent spells of sanity while I was writing it. He put up with strange meals, my strange hours, and my glooms and enthusiasms. He also provided computation advice, including pointing out that neither Minesweeper nor Freecell are word processing tools.

Contents

Figures

Tables

Chapter

1

Introduction

Health psychology, by its very nature, tackles complex problems in which psychological, social and biological factors interact. The discipline's dominant model, the biopsychosocial model, explicitly recognises this complexity, suggesting that we should consider interactions along biological, psychological and social dimensions (Sarafino, 1998). Other factors also add complexity in health psychology. For example, we are interested in phenomena over a range of degrees of aggregation. Topics in health psychology can range from how individuals perceive and cope with illness, to the influences on the health of social groups, to the design and assessment of health promotion activities targeted at whole populations. Where the topic touches these higher degrees of aggregation, we cannot simply add individual behaviours together to explain the behaviour of a population.

Another source of complexity lies in understanding the dynamics of phenomena. These determine how individuals or populations respond to change, and how long term processes unfold. While we usually investigate static relationships, what we are often really interested are the effects of change. For example, we may be interested in what effect an intervention might have or what effect a change in the environment might have. To understand responses to change we need to understand the underlying processes, and to recognise an element of dynamics in health psychology.

Complex interactions of a number of individuals are characteristic of systems problems, as are questions of the dynamic response to change. Psychologists have long recognised that systems can be important in determining behaviour. But despite acknowledging likely systems effects, in health psychology we rarely incorporate explicit systems ideas into theories. The biopsychosocial model is a good example of this. While many papers in health psychology begin by calling on the model in what seems a systems approach, in practice the subsequent analysis rarely addresses systems aspects. One reason for this may be that psychology uses few tools that allow us to conceptualise and analyse systems.

Although systems approaches have been relatively unusual in psychology, this is not to say that they have not been used. Recently there has been interest in incorporating dynamic systems approaches into social psychology (see for example, Vallacher & Nowak, 1994a). This literature often seems to jump straight into large non-linear and chaotic systems. Unfortunately, this gives the impression that systems approaches are necessarily arcane and difficult. Contrasted with these complex methods, simpler approaches may look too trivial to be of value. This is deceptive, as complex and interesting patterns can emerge from models that appear to be simple (Holland, 1998).

This thesis outlines and demonstrates a tool, computer simulation, that we might use to implement a dynamic systems approach to a phenomenon from health psychology. Here I identify and describe a methodology, computer simulation of a dynamic systems model, and make some claims for its strengths. I also demonstrate the methodology in action, to show that it can be practicable and fruitful. This is done through a concrete example: how a change in income inequality might disrupt social relationships in a population. This forms one link in a hypothesised causal chain from income inequality to mortality mediated through damage to social relationships (Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997). The example has some features that are common in problems in health psychology.

There is epidemiological evidence of a strong relationship between income inequality and mortality (Carroll, Davey Smith, & Bennett, 1996; Lynch & Kaplan, 1997; Wilkinson, 1996). This indicates that countries with a highly unequal distribution of income tend to have higher mortality than do more egalitarian countries. This is independent of their absolute wealth, and so is not readily explained by material mechanisms. Over the last twenty years income inequality has increased in many countries, including New Zealand. This might lead us to wonder how a change in income inequality might affect mortality.

One potential explanation suggests that the social fabric is disturbed in societies that have high income inequality (Wilkinson, 1996). Some evidence exists to support this. In England during the Second World War, a period of greater income equality and lower mortality also saw a population united by a strong identity and common goal (Wilkinson, 1996). People tend to trust each other less in societies with a high degree of income inequality than in more egalitarian societies (Kawachi et al., 1997). While we have evidence of trust being an important mediator, it is not obvious what mechanism might be involved. Other evidence suggests that strong social relationships are associated with reduced mortality (Berkman & Syme, 1979; House, Landis, & Umberson, 1988). We might ask whether the mechanism indicated by differences in social trust might involve an effect on social relationships.

This brings us to the first problem. While income inequality exists only as a characteristic of a population, social relationships can be conceptualised at many different levels. For instance, an individual might have a unilateral perception of being supported within a particular relationship. Between two people a relationship might be characterised by its degree of interaction, intimacy and function. Beyond that, each individual will have a wider social network of some size and composition. Finally, social networks are themselves crosslinked throughout the population to produce a connectedness between individuals in the population. For this example, we need to tie individual level data and knowledge about social relationships that we have from social psychology to a population level phenomenon like income inequality. We have few means to tie together these different levels.

The second problem is that if we want to know how a change in income inequality might affect social relationships, we have a dynamic element. Where we have changes, we have an initial response to the change, a final state and a period of transition between these. It is difficult for us to conceptualise the dynamic behaviour of psychological processes, as the methods that we use to develop theory in health psychology do not provide for or stimulate thinking in terms of dynamics.

The third problem is that the question spans economic and social dimensions. This multidimensionality is a common feature of phenomena in health psychology and the discipline has responded by favouring the biopsychosocial model as a theoretical base. The problem is that the biopsychosocial model provides no hints as to how to link constructs in different dimensions.

Studying the effects of a change in income inequality on social relationships highlights these three features; multiple levels of aggregation, the dynamic nature of many processes, and the interaction of different dimensions. These features are not exclusive to this phenomenon, and are commonly found in health psychology. Although common, we do not have methods that give us the means to explore phenomena with these features. Dynamic systems approaches offer ways to conceptualise and model processes in such a way that they incorporate multiple levels and different dimensions. Dynamics behaviours can be explored by building models and setting them into action. The example of the effects of a change in income inequality on social relationships provides both the challenge and the opportunity to demonstrate a dynamic systems methodology.

In this case, I have used a computer simulation to model how a change in income inequality might affect social relationships. The model incorporates a small population of individuals and relationships. A network of individuals and their relationships will be governed by characteristics of both, but the dynamics of such a network are usually dominated by the characteristics of the relationships. This is partly because relationships between people are more changeable than are individual traits. It also comes from the role of relationships in making the network linkages. We might draw an analogue with a human pyramid. Individually the people participating will be steady on their feet, and able to shift their balance significantly, shifting their weight onto one foot if necessary. When they build a human pyramid they make a mesh of connections between the individuals. The dynamic behaviour of the pyramid depends on those connections. If one connection is lost through someone losing their footing the consequences will ripple though the pyramid, and it will collapse.

The model is governed by a simplified set of characteristics, derived from some observations from social psychology about the making and breaking of social relationships. The first group of these relates to who we form and maintain relationships with; people who are similar and with people who are geographically nearby. The second group relates to decisions about maintaining or terminating relationships where the costs of maintaining the relationship have increased. In the third group, people with broken relationships will tend to change their level of participation in and demands of their remaining relationships.

Having defined the model in terms of individuals and relationships with a set of characteristics, it can be constructed. For this example, the model is constructed in the form of a computer simulation. Running the simulation programme sets the model into motion, and allows us to explore the effects of income inequality on social relationships in a population.

This example provides a demonstration that simulation can allow us to explore systems in social and health psychology. Some benefits of the methodology are demonstrated through the process of developing and exploring the model. These relate particularly to the way that we think about problems in health psychology. The initial development of the model forces us to think explicitly about the processes that might lie behind an observed phenomenon. Once the simulation has been developed, we are free to manipulate the model to try out different ideas, exploring its behaviour under different conditions. Some surprising outcomes emerge even in the development stages of the computer simulation. The development process is itself dynamic, with simulation runs feeding back some information about how a process might unfold. Running the simulation provides us both with the experience of patterns emerging from a dynamic process and with results that trigger further questions. The example demonstrates that this methodology can offer us a new way to think about the complex problems characteristic of health psychology. It forces us into a systems mode of thinking, in which we consider some processes that govern how individuals might interact.

This thesis brings together ideas from a number of different areas relating to the methodology itself and to the example that I use to illustrate the use of the methodology. I conclude this chapter with a summary of the order of presentation of these ideas.

The second chapter gives the background to this example. Most of this material is drawn from epidemiological research, where a relationship between income inequality and mortality was first reported. Some material is also drawn from the literature concerning the effects of social networks and social support on health. The chapter goes on to explore some features of the example that make it awkward to investigate using conventional analyses. These include problems of population level analysis and of the response of a large system to changes.

The third chapter notes some features of dynamic systems, particularly their characteristics and ways that we might approach them. The fourth chapter looks more closely at the methodology surrounding modelling. These chapters prepare the ground for the use of different types of models for dynamic systems, and lay out some ideas as to how we might use computer simulations in building knowledge.

The next three chapters describe how the example might be modelled using a computer simulation. The fifth chapter gives a detailed description of the development of the model and computer simulation. The sixth chapter describes the further refinement of the model through the early parts of the simulation and goes on to present some results obtained from running the simulation. A discussion of these simulation results follows in the seventh chapter.

Finally, in the conclusion I make some comments on the potential for the use of computer simulations of dynamic systems in health psychology, and how they might shape our thinking about health psychology.

Chapter

2

Income inequality and mortality

In this chapter I introduce the example, the effect of a change of income inequality on social relationships. The chapter draws on findings from the epidemiology literature that income inequality and mortality are correlated. A number of possible mechanisms for this correlation have been proposed (Lynch & Kaplan, 1997), including a hypothesis that there is a causal chain in which increasing income inequality damages social relationships resulting in increasing mortality. The chapter also touches on a different set of literature, that has explored the effects of social networks and social support on health. From this we have a finding that there is a link between the combined size and function of an individual's network of social relationships and mortality. The chapter goes on to consider some of the methodological difficulties that appear when we attempt to investigate a link between income inequality and an individual's network of social relationships.

There is a correlation between the equality of income distribution and health, such that societies with an egalitarian distribution of income have higher life expectancies. This has been found at many levels of aggregation; between countries (Rodgers, 1979; Wilkinson, 1990), across states in the United States (Kaplan, Pamuk, Lynch, Cohen, & Balfour, 1996; Kennedy, Kawachi, & Prothrow-Stith, 1996) and at the level of local electoral wards within England (Ben-Shlomo, White, & Marmot, 1996). While the correlation is found across all levels of aggregation, the pattern differs at different levels of aggregation (Soobader & LeClere, 1999), suggesting that aggregation is itself important in producing the effect.

Income distribution is usually conceptualised as some form of income inequality. This is the difference between the actual income distribution and the distribution if income was equally shared. Measures of income inequality can be a simple proportion of wealth held by the some part of the population, or can be expressed as a calculated index. Of the calculated indices, the Gini coefficient and the Robin Hood index have been most widely used in epidemiological studies. While a variety of measures have been used by researchers, the correlation between income inequality and mortality seems to hold whichever measure is used (Kawachi & Kennedy, 1997b).

The earliest studies of income inequality and mortality were based on figures drawn from international income studies (Rodgers, 1979; Wilkinson, 1990). Measures of income inequality are very sensitive to the income data used, so international income data are potentially not comparable because different countries have different economic cultures. For example, reported household incomes depend on how raw figures are adjusted for the numbers of people in a household and on taxation structures. Later studies have used data drawn from single countries, producing income data can be compared directly. These studies have continued to find a strong correlation between income inequality and mortality (Ben-Shlomo et al., 1996, Kaplan et al., 1996; Kennedy et al., 1996).

As with income data, health data from different countries are not necessarily comparable, so most studies substitute mortality data as a proxy for health data. While this is not an entirely satisfactory substitution, it does have some benefits. Death is reasonably consistently identified and reported, and international mortality statistics are readily accessible. Mortality data are adjusted for the age distribution of the population, with either age adjusted mortality or life expectancy at birth reported. Some researchers (Ben-Shlomo et al., 1996; Geronimus & Bound, 1998; Soobader & LeClere, 1999) have suggested that income inequality has a particular effect on premature deaths and have limited their data to reflect this.

Higher income inequality is associated with higher mortality and with lower life expectancy. The strongest effects have been found at the highest levels of aggregation, in international studies. Wilkinson (1990) found a strong correlation ($r=0.83$, $p<0.001$) between life expectancy at birth and the proportion of income going to the bottom 70% of the population. This study used income data from nine developed countries, but the comparability of this international data is questionable. This problem was avoided in a later study that compared census data from the individual states in the United States (Kaplan et al., 1996). This found a correlation ($r=0.62$, $p<0.001$) between mortality and the proportion of income going to the bottom 50% of the population. At the electoral ward level (Ben-Shlomo et al., 1996) the effect of the variation in deprivation in the ward on mortality was 7 per 100,000 per quartile of variation (95% confidence interval 4 to 9, $p<0.001$). This study also looked at the distribution of the mortality effect, and found that the effect appeared across all wealth quartiles, suggesting that the income inequality may affect the whole population.

The increasing significance of the effect at higher degrees of aggregation in the above studies suggests that the level of aggregation may be a factor in the strength of the correlation found. Some studies have directly investigated the effects of different levels of aggregation. These have found that median income accounts for most of the income inequality effect at lower levels of aggregation, but not at higher degrees of aggregation (Soobader & LeClere, 1999). At higher levels of aggregation, income inequality is an important factor in its own right.

The above studies strongly indicate that there is a cross-sectional association between income inequality and mortality, and that an aggregation process is important in this. Over the last twenty years many countries have seen increasing income inequality (Statistics New Zealand 1999, Wilkinson, 1996). While we might suspect that that changing income inequality would produce a corresponding change in mortality, cross-sectional studies are not sufficient to demonstrate this. We would need longitudinal data to demonstrate that changes in income inequality are accompanied by changes in mortality.

As with the above, most studies of income inequality and health have been cross-sectional, but a few studies have been longitudinal. In most of these longitudinal studies the sets of data available were too small for formal analyses, and so they have relied on graphical analyses (Pamuk, 1985; Wilkinson, 1992, 1996). These suggest that longitudinal changes in income inequality might be accompanied by changes in mortality. Wilkinson (1992) also investigated the effect of changes in income distribution on life expectancy. In two studies he found a correlation between the annual change in the proportion of income going to the bottom 60% and the annual change in life expectancy ($r=0.8$, $p<0.005$ and $r=0.47$, $p<0.05$).

Other researchers have investigated changes in mortality across periods of great economic change, both in Western countries (Ferrie, Shipley, Marmot, Stansfeld, & Davey Smith, 1995; Morris, Cook, & Shaper, 1994; Pappas, Queen, Hadden, & Fisher, 1993; Phillimore, Beattie, & Townsend, 1994) and in Eastern Europe (Walberg, McKee, Shkolnikov, Chenet, & Leon, 1998). These researchers have all taken data sets across a few intervals and have used conventional statistics to assess changes between the two time points, implicitly ignoring the time series underlying the process. This research design is based on the assumption that the process linking income inequality and mortality is reversible, and that responses to increases and decreases are symmetrical.

Some writers (Berkman & Syme, 1979; Durkheim, 1951) have suggested that the rate of economic change may also be important. A study investigating changes in mortality across regions in Russia following the collapse of the Soviet Union did find that the rate of economic change has been important (Walberg et al., 1998). If mortality depends on both economic conditions and their rate of change, this raises the stakes for understanding the dynamic nature of the process, as processes that depend on rates of change tend to exhibit dynamic behaviours.

Income inequality is strongly correlated with a population's mortality. The importance of aggregation in the process, and the importance of understanding the effect of changes indicates that there may be a significant systems component at work in this phenomenon. This suggests that the phenomenon may be the source of a suitable example for demonstrating a dynamic systems methodology. We now consider the potential mechanisms that have been offered as explanations, looking for a process that might form the basis for a systems formulation.

Explanations for the income inequality - mortality connection

Most proposed psychosocial mechanisms linking income inequality and mortality have been framed at either a societal level or at an individual level. At the societal level, explanations tend to be in terms of constructs that exist at high levels of aggregation, like social cohesiveness. Individual level explanations are based on individual level constructs, like the psychological and social stresses affecting individuals. Both societal and individual level explanations tend to be fixed within their frames of reference, with no sense of continuity between explanations at different levels.

Berkman and colleagues (Berkman, Glass, Brissette, & Seeman, 2000) have attempted to incorporate multiple levels into an overall model. They have proposed that there is a causal chain from the most aggregated level to individual biological pathways (figure 1).

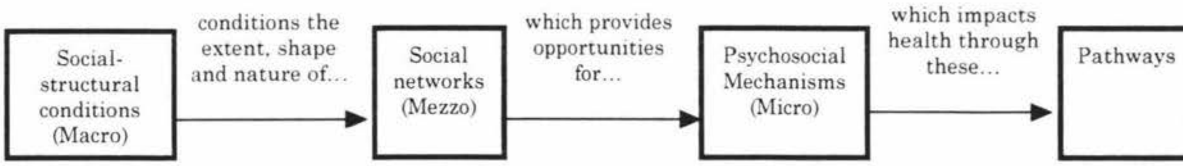

Figure 1. Chain of causality flowing from the most to the least aggregated level. (Adapted from Berkman et al., 2000).

This model is unusual in that it identifies processes associated with the arrows between the boxes. Further, the authors note that these processes are dynamic. Starting at the left hand side and fitting two pieces from the causal chain proposed for income inequality and mortality, the model might suggest that income inequality conditions the extent, shape and nature of social networks. This multilevel model is explicit in suggesting a causal chain working down the levels of aggregation from the highest to the lowest. Effectively this parallels a process of disaggregation; in which a high level social or structural condition works its way into lower level constructs. But the model could as well be operating bidirectionally. Between the two most aggregated levels social and structural conditions might shape social networks. And, in turn, social networks might aggregate to shape social and structural conditions. This would be consistent with the feedback effects that we would expect to see in a stable dynamic process.

The question of level of analysis is fundamental in relation to income inequality, as this is an inherently macro level construct. Any effect of income inequality on social networks and the mortality of individuals span the levels of figure 1. Income inequality is located at the highest level of aggregation and the biological pathway to individual mortality is at the lowest level of aggregation. Neither exclusively individual nor exclusively societal level explanations are entirely satisfactory, as neither satisfactorily addresses more than one level. Despite this, both explanations have constructs that overlay common ground in social relationships.

Social relations at aggregate levels

Some researchers have focused their attentions on the social environment of societies where income inequality exists (Kawachi & Kennedy, 1997a; Wilkinson, 1990), and have suggested that income inequality might be damaging to the social relationships in a society. Constructs associated with this have been called social capital, social connectedness and social cohesiveness (Kawachi & Kennedy, 1997a; Kawachi et al., 1997; Wilkinson, 1996). Features identified as indicative of these constructs overlap with many other equally vaguely defined concepts, such as social integration, social networks, social support and the density of relationships (Berkman et al., 2000; Sampson, 1991).

Of these, social cohesiveness is most commonly used in this context, but this is not a particularly well defined construct (Coburn, 2000). A broad range of features have been identified with social cohesiveness. An early hint of a similar construct is found in Durkheim's (1951) work on suicide and social integration. This extended a causal net beyond the individual to consider the integration of individuals into society. In Durkheim's work, the social integration of individuals is associated with a society's cohesion, and with a sense of identity, unity and shared actions and beliefs among its members.

This definition of social cohesiveness shares much common ground with concepts of social capital from political science. Putnam (1993, cited in Putnam, 1995) initially associated a society's social capital with its citizens active participation in their communities, and particularly with their participation in democratic processes. He later extended this to encompass "features of social life—networks, norms and trust—that enable participants to act together more effectively to pursue shared objectives." (Putnam, 1995) This continues the sense of shared actions, beliefs and purpose of Durkheim's social cohesion, while adding a component of trust between members of the community. A shared sense of purpose and action can also translate in part to a willingness to help others, a factor that has also been included in measures of social cohesiveness (Sampson, Raudenbush, & Earls, 1997).

Merging these social capital and social cohesion concepts we are left with a somewhat broad picture of social cohesiveness. It has a sense of unity and shared values among individuals, with individuals interacting with and supporting each other and having a sense of belonging and identity. A society with high social cohesiveness might be characterised by high levels of participation in community activities, by shared values and purpose and by a high degree of trust. In a broad sense social cohesiveness might be associated with number, function and strength of relationships in a society, including the more distant relationships with people that we interact with but do not know.

Having some sense of what social cohesiveness is, we might wonder how income inequality might be associated with changes in social cohesiveness. It has been suggested that income inequality may affect perception of the social environment (Lynch & Kaplan, 1997). It may be associated with diminished levels of trust and isolation of some groups (Soobader & LeClere, 1999), or with changes in social norms (Kaplan et al., 1996).

Kawachi and colleagues (1997; Kennedy, Kawachi, Prothrow-Stith, Lochner, & Gupta, 1998) have used United States data drawn from the census and from the General Social Surveys to directly investigate the correlation between income inequality, mortality and some aspects of social capital. Their indicators of social capital broadly fit some the features of social cohesiveness. Three of these were drawn from questions in the General Social Surveys relating to trust in other people, belief that others would behave with fairness, and the perceived helpfulness of others. Their fourth indicator was the level of civic participation, measured by the per capita number of groups and associations in a state. They used the Robin Hood Index, calculated from census income data for each state, as their measure of income inequality.

In path analyses they found a strong correlation (zero order correlation 0.77) for income inequality and mortality mediated by perceived fairness (Kawachi et al., 1997). Income inequality was very strongly correlated with social mistrust ($r=0.71$, $p<0.05$), with a perceived lack of fairness ($r=0.73$, $p<0.05$) and with a perceived lack of helpfulness ($r=0.71$, $p<0.05$). It was also correlated, but less strongly, with civic participation ($r=-0.4$, $p<0.05$).

Some researchers have suggested that wealthier people isolate themselves in enclaves of predominantly wealthy people (Kawachi & Kennedy, 1997a; Soobader & LeClere, 1999). It may be that, in addition to withdrawing geographically, wealthy people are withdrawing socially from relationships with poorer people.

Isolation into enclaves may have effects on the resources allocated to social spending in wealthy areas (Kawachi & Kennedy, 1997a; Kawachi et al., 1997; Soobader & LeClere, 1999), in addition to effects on social cohesiveness. This effect on resource allocation forms the basis of an alternative mechanism, the neo-materialist mechanism (Lynch, Davey Smith, Kaplan, & House, 2000), in which income inequalities are allowed to exist in societies that are also prepared to reduce investment in public goods such as education, welfare and health services. Lynch and colleagues (2000) do allow that income inequality is likely to have psychosocial effects but, as presented, the neo-materialist mechanism does not incorporate these.

While aspects of a neo-materialist mechanism are likely to be important, they too neatly separate income inequality and its material consequences from its social consequences. For example, structural disinvestment in the United States has included policies of targetted neglect of some neighbourhoods. In addition to their material effects, these policies have resulted in the isolation of some groups through the breakdown of weak links in social networks (Granovetter, 1973), and in the diffusion of both social disorder and health problems into neighbouring areas (Wallace & Wallace, 1997). Coburn (2000) notes that income inequality accompanies the adoption of neo-classical economics and that reductions in social cohesiveness go hand-in-hand with the individualism central to free-market economic policies.

Social networks, social support and health

Some writers have suggested that social cohesiveness has an micro-level counterpart in the social network surrounding an individual (Kawachi et al., 1997; Wilkinson, 1996). The effects of social networks on health have been investigated both by epidemiologists and by health psychologists. This research has taken two approaches, considering social networks either as sources of social support or as indicative of social integration. The social support approach suggests that social support might influence health by buffering stress, and that social support is only effective in the presence of a stressor. The social integration approach suggests that extent and functionality of a social network can itself produce a main effect on health (Cohen & Wills, 1985), whether or not there are stressors. The social support and social integration approaches have largely acted as competing alternatives, albeit with occasional acknowledgement that the relative value of the two modalities may depend on the situation (Cohen & Wills, 1985; Maton, 1989, Morgan, 1990).

As with the social cohesiveness literature, the terminology used in the social networks and social support literature is inconsistent. One definition of social support requires that social support consists the specific actions, intentions and interpretations that either giver or receiver interpret as supportive (Pierce, Sarason, & Sarason, 1990). In contrast, a social network is defined by its structure; a set of individuals and the relationships between them (Wasserman & Faust, 1994).

Unfortunately, the boundaries between social networks and social support have blurred, moved and, in some literature, disappeared entirely. Within the social support literature, the terms social networks, social cohesiveness, social integration (Cohen & Wills, 1985) and the density of relationships (Sampson, 1991) are used almost interchangeably with social support. The effect has sometimes been that social network effects have been subsumed under a more limited, but less defined, social support construct (for example, see Felton & Shinn, 1992). In this form the construct lacks the features of social support and reflects a general confusion of approaches. If we limit social support constructs to those meeting the Pierce, Sarason and Sarason (1990) definition above, it is not at all clear that the health effects of social networks are entirely due to social support. Other factors associated with social networks, such as social integration (Felton & Shinn, 1992; Morgan 1990, Thoits, 1995) stability and regulation (House et al., 1988) and collective efficacy (Sampson et al., 1997) may be as important as specific social support actions.

Cohen and Wills (1985) noted that the appropriate choice of a social support or social integration construct varies with the mechanism targeted. If a construct involves social integration, it should encompass both the number of relationships and the roles and functions that those relationships have for the individual. This provides for a construct that better captures the richness and variety of strengths in an individual's social networks.

This is supported by the findings of epidemiological studies. Using social integration measures that include both structural and broad functional aspects of social networks, these have found that social integration correlates with physical health and mortality (Berkman & Syme, 1979; Cohen & Wills, 1985; Felton & Shinn, 1992; House et al., 1988). When social network structural factors alone are measured, the evidence for a relationship between social network size and physical and mental health is weak (Cohen & Wills, 1985).

Social support has also been a focus of some interest in health psychology. It has been linked fairly consistently, although not necessarily strongly, with health. This effect is possibly mediated by mental health factors (Thoits, 1995). Research into social support and health outcomes has tended to focus on the types of social support provided and its effectiveness (Morgan 1990). Social support researchers have shown particular interest in the social support provided by close relationships and through helping relationships (Felton & Shinn, 1992), and have been less interested in the effects of an individual's wider social network.

Evidence from epidemiological studies indicates that social integration measures correlate more strongly with measures of physical health and mortality than do social support measures (Felton & Shinn, 1992; House et al., 1988). In addition, social integration measures produce a more consistent effect on health in the absence of a specific stress event than do social support measures (Cohen & Wills, 1985; House et al., 1988; Maton, 1989). This tends to support the existence of a main effect of social integration on health.

The social integration construct, which is based on the number, role and function of an individual's relationships, would appear to have some commonality with the social cohesiveness construct. The sense of a society's unity and shared values, interaction and support of social cohesiveness are located within the network of social relationships in that society. The effects of income inequality on social cohesiveness points to a possible damaging effect of income inequality on the network of social relationships. If social networks are damaged, the research from the social networks and social support literature would lead us to expect that this would damage be reflected in the physical health of individuals in a population.

We now have a phenomenon of interest, the effect of a change in income inequality on mortality. We also have some idea of a mediating mechanism, damage to the network of social relationships in the population. With a mediating mechanism, there are two steps to the process. In this case, these link income inequality to social relationships, and social relationships to mortality. For our example, we only need choose one of these to demonstrate the use of a dynamic systems approach. The first stage, a mechanism linking income inequality and social relationships illustrates features characteristic of a dynamic system, aggregation, dynamics and multiple dimensions, and so will form the basis for our example.

Methodological difficulties

If we try exploring how changes in income inequality influence social relationships, we quickly meet two methodological difficulties. At best, these make it difficult to apply conventional methodologies. At worst they leave us unable to theorise, and consequently reluctant to research the phenomenon.

Aggregation

The first of these difficulties is aggregation, or rather two aspects of aggregation. The first of these is that income inequality is located only in the aggregate, while social relationships are located at lower levels of aggregation. This means that our analysis has to span between the highest and lowest levels of aggregation. One approach might be to either aggregate the social relationships, or to disaggregate income inequality. Both aggregation and disaggregation (Susser, 1994) are risky processes that can distort the data.

The second is that there is evidence that the role of income inequality becomes stronger as the level of aggregation is increased (Soobader & LeClere, 1999). This is direct evidence that the process of aggregation may be an important part of the mechanism linking income inequality and mortality. With aggregation itself important in shaping the phenomenon, we are forced to consider including an aggregation process in our theories. But how might we theorise about a process that involves aggregation?

An approach taken in epidemiological research has been to theorise only at the macro level. This avoids the dangers of ecological inference (Kawachi et al., 1997; Susser, 1994), as macro level theories say nothing about what happens at other levels. One problem with this approach is that it is not always entirely clear what meaning we might associate with macro level constructs. At this level, simply defining the constructs can be difficult, and we see a hint of this in the confusion of terminology surrounding social cohesiveness. We might expect to have even more problems theorising about constructs that we find difficult to define. Approaching processes involving aggregation at the macro level evades the process of aggregation, rather than illuminating it.

Techniques that address aggregation are rarely engaged in psychology. Having focused on individual subjects, our methods are not particularly well suited to theorising about processes involving aggregation. In the case of the example, we have an aggregate process that may involve social relationships. Social psychology has a body of knowledge about social relationships, but health psychologists have not invoked this when investigating an association between income inequality and social cohesiveness. This may be because it is difficult to see how we might do so. We need to find a way into aggregation processes so that we can apply micro and pathway level knowledge to problems involving aggregation.

Dynamics

Many of the studies investigating effects of income inequality on mortality have been cross-sectional. Few of these have questioned whether cross-sectional studies reflect the contemporaneous income inequality or income inequality of an earlier period. This is worrying, as most studies have been carried out through the 1980s and 1990s, during a period that has seen rapid and large increases in income inequality (Statistics New Zealand, 1999; Wilkinson, 1996).

The approach taken in most studies has been to correlate cross-sectional data. This makes the assumption that both the independent and dependent variables have settled to some stable state. But neither mortality nor the economic policies that bring changes in income inequality are static. Economic conditions and the political ideologies that drive them are cyclical. Mortality has tended to drop steadily throughout history (Evans, 1994), and periods of static or increasing mortality have been rare, so that the steady state condition in mortality has been an annual reduction in mortality. As a result, an effect may appear as a failure of the trend of reducing mortality, rather than as an increase in mortality, with increases in mortality indicating an extreme effect. Cross-sectional designs are of limited value in understanding the effect of economic changes, as to understand them we need to understand the dynamics of the process.

Again, we are not particularly accustomed to building theories about how a dynamic process will unfold. Our theory building has become bound both by the culture of previous research and by the familiar statistical techniques that we prefer to use. In general, there are only very limited statistical techniques available to apply to dynamic processes. If we restrict our theorising so that we only build theories that can be tested by statistical techniques, we are limited to a theoretical construction set whose only dynamic components are trends, noise and regular cyclical processes.

In the first chapter I noted three features, aggregation, dynamics and multiple dimensions, that are indicative of dynamic systems. We have seen that the example exhibits both aggregation and dynamics. These make it difficult for us to theorise, as our primary mode of theory building relates to static states at a single level of aggregation. The example also features multiple dimensions, as one construct is located in the economic dimension and one in the social dimension. The example exhibits all of the features that might lead us to consider a dynamic systems approach, and so is a good candidate to demonstrate the application of a dynamic systems approach. Before I do so, I will introduce some concepts of dynamic systems, and to lay the ground for the methodology that I will use in this example.

Chapter

3

Dynamic systems

In the previous chapter I introduced the example, and noted that it features aggregation, dynamics and multiple dimensions. I have indicated that these are characteristic of dynamic systems, so now I will step back from the example for a brief introduction to the nature of dynamic systems.

A dictionary definition of a system is that it is "a complex whole; a set of connected things or parts; an organized body of material or immaterial things." (Allen, 1990). The key ideas in this definition are that the system is a whole, and that it consists of a number of connected parts. As we shall see, these ideas are more powerful than such a simple definition might suggest. It is not difficult to think of examples that meet this definition of a system; every discipline has its systems, and consequently its systems problems, ranging from computer networks to stock markets to our biology, families and organisations.

The first disciplines to take a formal approach to systems were the physical sciences. These developed a mathematical science of the theory and analysis of systems (Elgerd, 1967; Power & Simpson, 1978). The first systems studied were linear physical systems, in which the equations governing the behaviour of the components describe straight lines. Although the equations are linear, the analysis of linear systems is not straightforward, because the equations also include a time component. The interconnection of components means that we have to know the state of each component at any time to understand how it is affecting its neighbours. This means that to understand how the system behaves; we need to understand its dynamics. The central role of dynamics in understanding system behaviour is almost as defining of a system as is its consisting components and their interconnections.

Some features of linear systems emerge as a consequence of their linearity. If a system behaves linearly, its steady state output depends only on its inputs. If it is stable, it will always follow a trajectory to the same stable output for the same inputs, and will settle when it reaches that output. As a stable system will always move to the stable output, we can investigate some features of a linear system using conventional statistical techniques, as long as we leave it long enough to settle after a change.

Much of the effort in linear systems science has been to identify whether a system is stable, or how it can be made stable. To be stable, a system needs a self-correcting mechanism. This works through a reverse path, the feedback path, from output back to input. Systems without a feedback path cannot be stable, and a change of input to such a system results in the system collapsing.

An important feature of linear systems is that we can add its responses to different inputs together. This is a direct consequence of the simplicity of the linear equation. For example, if the system produces an output Y for an input X according to the equation

$$Y = 4X$$

$$\text{If } X = 12,\ Y = 12 \times 4 = 48$$

This output is the same if we split X into components so that

$$X = 1 + 2 + 3 + 6 = 12$$

we get the same result for Y when we separate the components of X

$$Y = (1\times4) + (2\times4) + (3\times4) + (6\times4) = 48$$

This is not the case for a non-linear equation, so if

$$Y = X^2$$

$$\text{If } X = 12,\ Y = (12)^2 = 144$$

but this time if we separate the components of X, we do not get the same result

$$Y = (1)^2 + (2)^2 + (3)^2 + (6)^2 = 50$$

The ability to separate inputs and add the system's responses to each is an important and universal feature of linear systems, called the principle of superposition. It is a very useful device for simplifying problems, as complicated inputs can be divided into simpler components, and the system response to each calculated. It also means that we can investigate the system response to two separate inputs one at a time, and derive the system response to the combination by adding the two results.

To give an example, we might apply some system ideas to a model of trying to lose weight. The input is the target weight, and the output is the actual weight. First we might start with a non-systems approach, as shown in figure 2. This has the output weight as simply a function of food intake and their activity level. So all we need for weight loss is to know the target weight and a function for setting diet and activity to produce the target weight. Unless the initial food and exercise combination is correct for the target weight, a one off intervention setting food and exercise levels might never reach the target or it might overshoot.

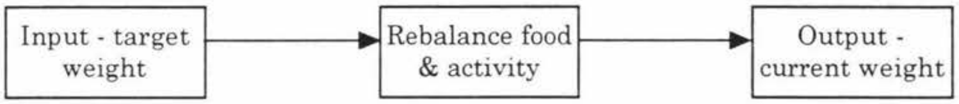

Figure 2. Non-systems approach to setting weight.

In reality, the system has a number of feedback mechanisms that function to make a system of weight maintenance. An essential feature of feedback is that the output, the current weight, is compared with an input, the target weight. This tells the weight control system whether there is a discrepancy between the target and current weights. Figure 3 shows the block diagram modified to show this feedback comparison. Making this comparison repeatedly allows the rebalancing of food and activity to become a continuing process, and for it to adapt if the target changes.

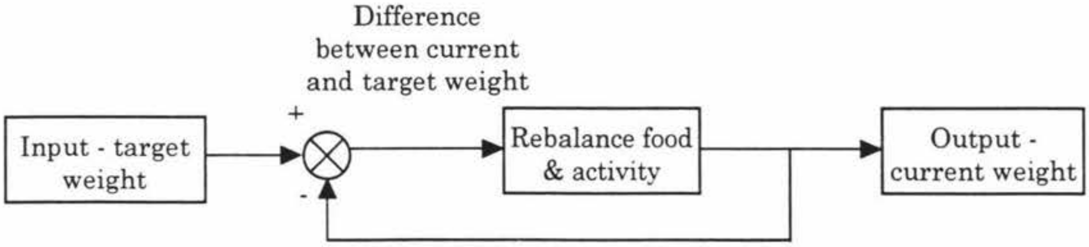

Figure 3. Weight control system with feedback.

The feedback path for weight control is much more complex than a simple comparison. It includes negative feedback effects, like feeling hungry or social pressures to eat, that tend to return weight to its original level. These negative feedback effects tend to have a stabilising effect. The feedback path also includes positive feedback effects like rewards for weight loss that may tend to amplify the weight loss process. Positive feedback tends to have a destabilising effect on a system.

If a system contains any components that behave in a non-linear manner, including any which have a discontinuity, it is a non-linear system. Non-linear systems do not behave as neatly as do linear systems. They may have a number of different stable outputs. Similar inputs may result in quite different outputs or even an oscillation around a point or between two points. These oscillations can be so complex that, at a glance, they may look like noise although they are not random. And, counter-intuitively, from these complex processes simple patterns can emerge. Superposition does not hold for non-linear systems, so we cannot simplify the analysis of a non-linear system by dividing inputs into easier to handle components. Except in a very few, very small, very simple examples, we cannot solve the equations of non-linear systems.

This is not to say that non-linear systems are intractable to investigation. Firstly we can investigate whether what looks like noise is really noise, or whether it is the result of a complex, non-linear, but deterministic, system. As with linear systems, we can look at the system's stability. We can look for the stable points or attractors, and investigate the trajectories that system follows to reach these attractors. We can investigate the linking mechanisms in the system and how its feedback mechanisms work.

Returning to the weight control example, a person's weight control system may, or may not, be stable and it may, or may not, be linear. If linear it will have a single attractor, or stable point. This is likely to be the person's original weight. Conversely, it may be non-linear. If so it may have other stable points, or attractors, and there may be other stable weight outcomes possible. A person's weight may be stable within a certain range, but if nudged outside that range it might become unstable, so that it falls or rises catastrophically. An attempt to modify weight might find a trajectory that oscillates between a high and low weight without ever settling.

This illustrates some of features that we might expect if weight maintenance behaves as a dynamic system, rather than as a simple single step process. Treating weight maintenance as a dynamic system, rather than a simple process, would lead us to expect some of the otherwise inexplicable outcomes to emerge from people's attempts to modify their weight. It provides a theoretical base to understand these effects that simple steady state theories cannot.

In addition to its uses in investigating feedback systems, a dynamic systems approach can allow us to investigate an aggregation process. This follows directly from our definition of systems being a number of components and their interconnections. In such systems in which a large number of components are interconnected, the simple responses of a large number of components can produce interesting and informative patterns in the system.

Conventional statistical methods are not well-suited to assist in the generation and testing of hypotheses about dynamic non-linear systems. Statistical analyses are based on an assumption that there are two components to a set of data. One component is determined by a relationship between the dependent and independent variables. The other component is purely random noise. These components are simply summed, resulting in data sets that behave well statistically. This approach makes two assumptions. Firstly, we have seen earlier that we can only add components if a process is linear. Conventional statistical approaches assume that the processes under investigation are linear.

The second assumption is that there is some source of entirely random noise in a social system. The strongest statistical results in psychology might produce correlation coefficient of 0.7, and so explain about 50% of the variation in outcome. Using statistics we implicitly assume that the remainder is due to people behaving in an entirely random way.

An alternative is that what looks like noise to our statistics is the result of a trajectory in a dynamic and complex system (Nowak & Lewenstein, 1994). These trajectories may arise when a large number of people act in non-linear but predictable ways within a large system. The behaviour of such a system is deterministic, not random (Alligood, Sauer, & York, 1996).

The social sciences have long recognised the importance of systems can be important in understanding behaviour. But while a need for understanding problems in a systems framework has long been identified, the development of the means to do this has been slower. This slowness is not confined to psychology or even to the social sciences (Barabási & Albert, 1999), but it is more pronounced in the social sciences because the systems involved are complex and non-linear. Although the significance of non-linearity and complexity to psychology has been questioned (Puddifoot, 2000), we inherently behave non-linearly. Our lives are full of discontinuities; we change workplaces, join new activities, have children, break up with other people, suffer disasters and win lotteries. Each of these events represents a discontinuity and consequent nonlinearity. Until recently the mathematics of non-linearity were not well understood, and the patterns that arose from complexity were not recognised. Recently, understandings of these systems has advanced mathematically, and the results have been brought to a wider audience (see for example Gleick, 1987). This has made them more accessible to social scientists.

In health psychology, we have a further factor driving the need for a systems approach. The biopsychosocial model encourages us to contemplate complex mixes of biological, psychological and social mechanisms. Beyond this the model does not suggest how we might achieve this (McLaren, 1998), as the model is prescriptive rather than descriptive. A problem with conventional analyses has been that, at best, the social component sits unevenly among more readily individualised components. At worst, the social dimension tends to disappear entirely.

One method of accessing social factors has been through qualitative techniques. These allow us to develop a very detailed picture of an individual and their interactions with structural mechanisms. While qualitative techniques are useful, the information that they can get about broader social patterns in a large system depends their direct effect being reported by participants.

A dynamic systems approach offers a different way to integrate social factors with factors in other dimensions. This is because conceptualising a systems formulation forces us to think explicitly about the processes behind the interaction and interconnection of constructs. When we do so, the different dimensions merge into a single picture at a systems level. Systems approaches can allow social structures to be represented explicitly in our theories, integrated with psychological biological and economic factors.

I have already indicated that a dynamic systems approach offers a way into a problem that involves aggregation. Where we build theories of a system that aggregates individuals and their relationships, we are constructing a theoretical representation of a social network. The behaviour of large networks of individuals tends to be dominated by the relationships between them, rather than by the characteristics of the individuals in the network. To illustrate how this might happen, I will begin with two simple components, balls and springs. These components can be set into motion by an external environmental agent, a simple force pushing on one ball. The movement of an individual ball in response to a force acting on it from the environment (figure 4) is determined by the individual characteristics of the ball. Its motion in response is simple and easily understood.

Figure 4. Individual ball subjected to an external force.

Next, considering a dyad of balls connected by a spring (figure 5), the movement of these in response to an external force on a ball depends not only on the characteristics of the balls but also on the characteristics of the spring. The movement of the dyad will be more complex than for the individual ball, but still relatively simple.

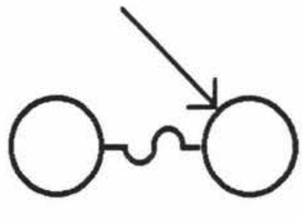

Figure 5. Dyad of balls connected by a spring with one ball subjected to an external force.

We move finally to a network of balls connected to each other by springs (figure 6). The movement of the balls in response to an external force now depends to a great extent on the network of springs linking the balls to their neighbours, and linking those neighbours to their neighbours.

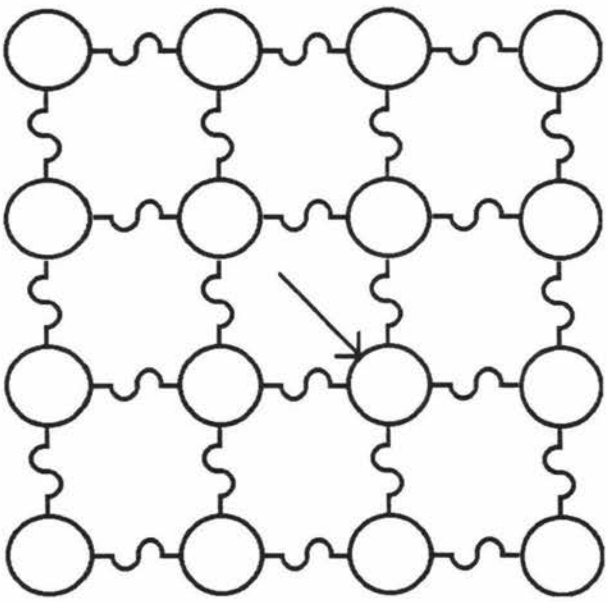

Figure 6. Network of balls connected by a spring with one ball subjected to an external force.

These ball and spring models are simple, but that simplicity allows us to take a first step into understanding the importance of relationships at higher levels of aggregation. The simplicity of the physical model may lead us to believe that relationships may not be so important in systems where there is a great deal of variability between individuals.

For an example that involves people we might consider a sports team. Children who are beginning to play a team sport act as individuals with no relationship linkages. They all pursue the ball, and seem to regard team mates and opponents equally as obstacles. At this level, a soccer match looks like a group of twenty children chasing a ball around the field, and could be thought of as twenty unrelated individuals all trying to get the ball.

At senior levels every player has a role in relation to other players, both team mates and opponents. This dominates both their position and movement, and the ball is passed along these relationship linkages. Patterns emerge in whole teams, and form the basis for tactical moves. The effect is that the team operates as a system. Although some players may be more skilled than others the performance of the team depends on the interaction of all players.

When we investigate systems, the most interesting patterns can emerge independently of the variability in the characteristics of individuals making up the system. This is not to say that individuals are not important and interesting. Their own characteristics and environment are, at the very least, important to themselves. But drawing on a systems conception incorporates individuals into their environment. It allows us to recognise that our web of relationships is itself important in determining our environment, and affects us in ways that we may not necessarily be able to observe from our vantage point within the web.

This brief summary alerts us to some things that we might be alert to in beginning to assemble a dynamic systems approach to our example. To this point, all we have decided is that we are interested in the effects of a change in income inequality on social relationships.

Having social relationships as one of the key constructs alerts us to the possibility that system might be based around a network of relationships. These can be powerful determinants of the behaviour of a system. A system that is based on social relationships is likely to be non-linear, as the paths to formation and dissolution of relationships differ, and relationships tend to breach suddenly. Social systems are, overall, relatively stable; despite constant change they rarely collapse totally. Our interest in how they change might lead us to be interested in what the stable outcomes are, what is important in determining this, and what trajectory the system takes in reaching stable outcomes.

This gives us a beginning idea of the form that the example might take, but to this point we have no indication as to what tools we might use to think about this, develop our ideas and explore such a system. This depends on how we model the system. In the next chapter I outline some aspects of methodologies for modelling a system.

Chapter

4

Methodology

In the previous chapter I introduced some dynamic systems concepts. So far, I have made no comment about how we might investigate them in practice. As we have seen, the example exhibits many features that are characteristic of dynamic systems. These features are widespread in phenomena in health psychology. To understand these phenomena and the dynamic processes involved we need to draw on different tools from those used in static analysis.

Many techniques exist for the investigation of systems. Some are specialised to the systems that they were developed for, so programmes used to calculate power system stability are unlikely to be useful in social systems research. Other techniques have more general application, including analytical and simulation techniques. No matter the technique, the model is the foundation on which we base our theory and understanding of a dynamic system.

Models and modelling

A model is an aid to understanding. It may take many forms, but common to all is that they are simplified representations of more complex entities.

We are already familiar with making models in health psychology, as model-making at many levels is an integral part of our standard research practice. Our psychological model-making is at its most explicit when we draw a flowchart diagram, and in this activity we work with multiple levels of model. We work at a theoretical level with construct models. We map these to another type of model, measurement models, and we realise them as statistical models.

In drawing a flowchart model, other, less obvious, models are developed along the way. Although it is perhaps not obvious that constructs are also models they are simplifications, representing far more complicated concepts. For example, a construct called social cohesiveness may be a greatly simplified representation of selected aspects of a multifaceted phenomenon that includes aspects of co-operation, trust, participation and social activity. Defining a model, a construct called social cohesiveness, allows us to identify some common features and to define some boundaries to a concept. It allows us to get a handle on an idea without the definition sliding around, and it allows us to contain the complexity to something that we can understand. The danger is that we risk mistaking a construct for something real. This illustrates a general danger in building models - we can become confused about the status of our model (Brogan, 1974).

Despite the routine use of modelling in health psychology, we are relatively unfamiliar with building models from scratch. As a first step in creating a model, we should have a clear idea of the purpose of a model. It is an aid to understanding things that happen in the real world. The phenomenon of interest (the target) may be too large for us to observe, too complex for us to understand, too inaccessible for us to study or too dangerous for us to manipulate safely. In lieu of studying the target, we can study something (the model) that is similar to the target in some important ways. A well-designed model should behave similarly to the target in ways that we are interested in, while being easily manipulated. If we design a model well, we can find out things about the target by studying the model.

I have indicated that the model is intended to aid our understanding of some target phenomenon. Looking more closely, we might wonder how we would use a model to do this. One advantage of using standard models and standard methods is that they come pre-packaged with a methodology. In the case of conventional quantitative methods, this methodology depends on testing a falsifiable hypothesis (Hilborn & Mangel, 1997; Scruton, 1996), through statistical hypothesis testing techniques. As previously noted, these techniques are not particularly well suited to dynamic systems problems in the social sciences. Adopting different methods and different models brings with it a question about how we learn anything from these models. Consequently, not only do we need to design new methods, but we also need to think about how these can be good science.

There are two major alternative ways that a model can inform us; we can take a rationalist or an empiricist approach to modelling. A rationalist model is constructed logically from axiomatic rules and derivable relationships between rules. This approach has worked well in the physical sciences, but the social sciences are short of fixed rules that might even give us a start into making rationalist models.

An alternative is to take an empiricist view, and to assume that there is a real world and that we can find out about it through observation. We do this in conventional research when we study individual subjects, as we can observe them directly. It is much more difficult to observe large social systems directly. Instead we might substitute for the target with a model, and observe the model. Of course, this may tell us about the model, but how can we be confident that we can learn something of the target from this?

That confidence can be built through hypothesis testing. The understanding that we gain from studying the model should allow us to make hypotheses about the target system. Where possible, these can then be compared with observations of the target. To the extent that the model can assist us with understanding the target, it is a good and useful model. This remains an empiricist approach, as it is still based on observation of a real world system. But the real-world system that we observe directly is the model and the target system is only partially observed. Where we can observe the target system, we can compare the model and target and can judge how well the model matches the behaviour of the target system.

Some design considerations

Models may be constructed in many ways, limited only by our imagination and the evidence to support its design. We should not aim for the model to match the target in every respect. If we did so, it would be identical to the target (Holland, 1998), and would be as difficult to investigate as the target. Rather, we should attempt to simplify and capture some of the behaviour of the target in the model (Hilborn & Mangel, 1997). We should include only relevant detail, as identified by our theoretical account, while excluding irrelevant detail. This means that creating a model needs thoughtful design.

We are familiar with models based on simple mathematical equations, because psychologists use these mathematical models routinely. Although the statistical analysis that surrounds them can be very complex, the mathematical models themselves are usually extremely simple. For example, in univariate linear regression, we model a relationship between constructs as a straight line. The first stage of the process is to fit a line of the form $Y = A + BX$ as closely to the data as possible. In the second stage we test how well the line fits the data. In practice, we find these mathematical models extremely uninteresting. So much so that we sometimes do not report the detail of the model that we are testing, and only report how good the fit was. For example, very few of the articles reporting findings in the income inequality and mortality literature report the parameters of their regression models.

More complex mathematics can also be used for modelling. For example, the vibration of a drum head can be modelled by Bessel's equation.

$$\frac{d^2W}{ds^2} + \frac{1}{s}\frac{dW}{ds} + W = 0 .$$

Mathematical modelling has been very successful in the physical sciences, and has occasionally been assumed to be the ideal form of model for the social sciences (Chattoe, 1996; Doran & Gilbert, 1994). But the above equation would not be especially helpful if we were trying to understand how to play a drum. The ideal model depends on what we are interested in, so if we wanted to know how to play a piece we would model the vibration of a drumhead in an entirely different way.

Figure 7. A model of the vibration of a drum head.

Most models used in psychological research are static, and assume a steady state relationship between their variables. This static nature is not always apparent. For example, flowcharts are a familiar and common modelling device in health psychology. Although the name is suggestive of motion, our flowcharts are somewhat ambiguous representations of static relationships with arrows used to indicate causal direction. These arrows suggest some process underlying them, but the boxed form of the constructs can distract interest from those underlying processes (Spicer & Chamberlain, 1996). Causal direction is an uncomfortably dynamic interloper in what is, otherwise, a comfortably static model. And while we may have a causal direction, a static model does not address how things change. Our conventional flowcharts, with boxes representing constructs and lines representing causality and correlation, cannot represent dynamic systems particularly well.

Other flowchart styles, such as the block diagram of the weight maintaining feedback system seen in the previous chapter, can give us better models of dynamic systems. Again, the design of a flowchart model should depend on what we are interested in.

Computers can allow us great flexibility in constructing models. Here the model is not the computer programme itself, but the collection of algorithms and rules that are realised in the programme. These may be derived from mathematical models, or they may be other forms of algorithm such as logical statements or fuzzy statements. Computer models can be used to investigate the dynamics of a system when the algorithms are set into action by running the simulation.

Scale models are a common type of model. They are useful for visualisation, for example for showing what a new building will look like. A scale model contains quantitative information about the target. In a 1/100th scale model we would expect to be able to determine the distance between buildings by measuring the corresponding distance on a model and multiplying by one hundred. While we are unlikely to use physical scale models in health psychology, our models do sometimes incorporate scaling. In health psychology, scaling could be applied to quantitative measures, like time, and incorporated into the model. More often this will not be the case, and if we expect the output of a simulation model to correspond with a quantifiable time scale we need to design it for this. Whether or not we can do so will, however, be dependent on the quantitative information that we have access to. So to scale for time, we would need to have a good understanding of the time constants and time delays in the processes that we are modelling. This information is rarely available in health psychology.

While a scaled model might seem the ideal, in fact the need for scaled models is rare. Again we need look no further than conventional psychological research for an illustration. Although we use quantitative techniques, we do not take quantities from our results. Rather we state our findings using qualitative statements like "there was a significant difference in the scores". Equally, we do not need to set ourselves the task of producing quantitative results from scaled models. Qualitative models will be every bit as helpful in allowing us to develop an understanding of the complex patterns emerging from a dynamic system (Chattoe, 1996; Doran & Gilbert, 1994; Holland, 1998).

Relationship with theory

The process of building a dynamic model is similar to the process of building the static models that we are already familiar with. We start with a question, and some idea of the factors that might have some role. With a static model we might be happy to find that the constructs are correlated. Where we are concerned about how this correlation comes to be, we are likely to seek other constructs that mediate, moderate or confound the correlation. Even if we were to find all of these, they would tell us nothing of the process by which one construct affects another. This is not the case with dynamic models. If we are building a dynamic model, we need to understand the processes themselves. We cannot understand the dynamics of a system without understanding the processes that govern it, and so identifying and modelling processes is an important part of building a dynamic model.

I should note here that the construction of a model is closely related to building theory. We have some experience of this relationship between theory and model from our flowcharts. These have allowed us to build theories of a particular form; a selection of constructs linked by correlations and causations. This is a very rudimentary form of theory building. Its dominance in psychology has resulted in a set of fragments of knowledge, with very little to link those fragments into an overall understanding. This has been a particular problem in social psychology (Vallacher & Nowak, 1994b).

There is a different quality to theories in the physical sciences (Phillips, 1992). Physical scientists look for their theories to explain why an observation holds. This needs an explanation that is distinct from the observation; we cannot simply claim a correlation as a theory for itself. While theory in the social sciences is unlikely to ever reach the unity of physical theory, we can use the style of theorising in the physical sciences as a guide to raise the bar a little in our theorising. We can require that our theories explain a phenomenon through the explaining the processes that generate it. With this requirement for our theory, we cannot claim a relationship between changes in income inequality and social relationships as theory. It is simply an observation. We need to look for some underlying process that might lead to this phenomenon. Here we might recall our requirement for building a model of a dynamic system, with its focus on the processes governing the system. In effect, setting ourselves the task of building dynamic models forces us to develop theories of social processes.

Building a dynamic systems model

When we have a reasonably clear sense of what aspects of the target we are interested in, of what questions we might want to ask and of what features we might want to include in the model, the next step is to design the model. The first design step has a creative component that is dependent on having an idea. There are no magic formulae for generating ideas. But this should not be a problem, people have ideas all of the time. However, having an idea is only the start. Most ideas are not especially interesting in themselves, and to make something interesting of we need to develop the idea. While we cannot learn to have ideas, we can learn techniques for developing those ideas into something interesting (J.S. Body, personal communication, January 13, 2001).

Here I have taken a technique for developing a dynamic systems model from an engineering text (Brogan, 1974). Engineering techniques are obviously not directly applicable to psychology problems and so I have modified some of the steps slightly to reflect this. This development technique is followed in control systems engineering, a discipline with a long established history of working with dynamic systems. This is not to suggest that this is the only technique that we might use for developing a dynamic systems model, but it does give us a convenient starting base.

•Specify purpose of the model.

•Define the boundary of the target system, and its inputs and outputs.

•Break the model down into simple idealised components.

•Identify the variables of interest.

•Specify the characteristics of each component.

•Specify rules governing the interconnection of the components.

•Manipulate the rules and characteristics into a form that allows solution.

•Analyse the behaviour of the model.

•Do not confuse the model with the real world.

The first two steps relate to the first parts of the process described earlier. We need to define the target that we are interested in modelling, what aspects we are interested in and what can be left out for simplification. We also need to have some idea of what we would like to know from the model.

Going back to the definition of systems as a connected set of things or parts, we need to decide what things and parts, or components, will be connected to construct our system. And we need to decide on their characteristics and the rules for how they can be assembled. In the next group of four steps we define the components of the system, the ways of characterising our components, and the ways in which they interact. In psychology obvious components might be individuals, groups and organisations. Less obviously, our system components might include relationships between people.

In the next two steps we build the model, then we analyse its behaviour. Where we use a computer simulation to build the model, we analyse its dynamic behaviour by running the simulation and monitoring its output.

The final step is a reminder of the status of the model. Throughout the process the model remains simply a model, an aid to understanding what might happen in an otherwise incomprehensibly complicated system. It is not intended to incorporate the full complexity of reality, it is merely a limited tool to allow us to understand some aspects of a complex whole. The most that we should ask of it is to provide hints as to what we might find if the processes modelled are a significant determinant of system behaviour.

We now have some ideas about how we what our purpose is in modelling a system, and how a model might inform us. Armed with this and a step-by-step plan for how we might go about developing a model, we now return to the example to illustrate how we might begin to develop a dynamic systems model for the example.

Chapter

5

A model of income inequality and social relationships

At this point we return to the example, to show how we might work through the development of a model and its implementation and application through simulation. To recap, for the example we will develop a model of the effect of a change in income inequality on the social relationships in a population.

We must begin with an idea of a process by which a change in income inequality might result in disruption to social relationships. I have based the model on the possibility that an increase in income inequality disrupts the normal development, maintenance and dissolution of relationships. This is not the only possible mechanism. For example, another might be that an increase in income inequality disrupts the quality of relationships, so that, while relationships are maintained, they do not have the same functionality. There was nothing in existing theory that led me to prefer one or other of these, as we are only at an exploratory stage. There is some sense of an ongoing process in the cycles of relationships that suggest this as a good potential starting point to illustrate the development of a dynamic network model. The nature of the structural and dynamic content of relationship making and breaking may make it easier to visualise the dynamics of the process in action.

With this starting idea, we should look more closely at what is known of the formation, maintenance and dissolution of relationships. This will provide some theoretical base for modelling the process.

Relationship formation and dissolution

We are looking for a process by which income inequality might be linked with the making and breaking of relationships. While changing income inequality has no equivalent construct at the individual level, it does have a consequence for at least some relationships. Changing income inequality means that some people have gained or lost income in comparison with others. Where this gain or loss affects only one individual in a dyad this may have the effect of reducing the similarity of the people in the relationship. Similarity is an important factor in the making and maintenance of relationships.

In New Zealand the change in income inequality in the 1980s and the 1990s resulted from a large increase in the income of the highest earning decile (Statistics New Zealand, 1999), and a smaller decrease in the incomes of the lowest earning 80% of the population (Podder & Chatterjee, 1998). Our income determines many of the things that we do, and where we do them. An increase in income sees the beneficiary able to access different places, activities and interests and facing different types of problem to those faced by people with lower incomes. Increasing the income of the highest earning decile may reduce the similarity of the people in that decile with some of their friends.

We are more likely to make and maintain friendships with people with whom we share some similarity. The inverse of this is that one factor cited in relationship dissolution is that the individuals in a dyad find that they have a reduced degree of similarity (Fehr, 1996; Sprecher & Fehr, 1998). An increase in dissimilarity resulting from an increased income difference between individuals may trigger the termination of the relationship. This process forms the first step in the model.

It is also possible that a change in an individual's wealth may trigger a relocation. Physical proximity is also known to be important both in the formation and in the dissolution of relationships (Fehr, 1996; Sprecher & Fehr, 1998) and this in turn is also likely to influence both availability and frequency of interaction. It has also been cited by epidemiologists as being a factor in changes in social cohesiveness (Kawachi & Kennedy, 1997a; Soobader & LeClere, 1999). Relocation is not included in this model, but is likely to be an important factor that would be included as the model was developed. I have not included it in the example so that the demonstration is easier to follow.

This first step, in which relationships are broken when there is increasing dissimilarity, does not signal the end of the process. We go on to consider the consequences of these relationship breakdowns for the remaining relationships.

Where an individual has lost a relationship, that person may look to other existing relationships for support, and to pick up the functions lost with the broken relationship. There is some indication that this does happen. During and after the termination of a relationship individuals are likely to seek social support (Sprecher & Fehr, 1998) and to increase their participation in remaining and new relationships (Fehr, 1996; Sprecher & Fehr, 1998).

This will tolerated in many relationships, but in a few cases this loading onto a remaining relationship will push that relationship beyond breaking point. Although not directly cited as a reason for relationship dissolution in the literature, there is some suggestion of this as a reason for relationship breakdown. In the social exchange formulation, a balancing of costs and benefits goes toward a decision to terminate or maintain a relationship (Sprecher & Fehr, 1998). A relationship may be terminated when the costs of the relationship exceed its benefits. Dependency and smothering behaviours are also cited for relationship deterioration (Fehr, 1996) and these are also suggestive of excessive demands being put on a relationship.

A construct incorporating the time spent on a relationship, the costs of the relationship and the functions that the relationship fulfils will be incorporated into the model. This construct is similar to the cost side of the social exchange formulation. It could also be visualised as stress on a relationship, or as seeking or giving social support. I haven't identified it with any particular one of these elements. This is because each of these comes prepackaged with a previous history and definition in psychological research. As psychologists, we have some formal knowledge of these constructs, but any knowledge that we have is based on static research and theorising. These constructs may not translate directly to dynamic equivalents. This construct is also similar to the relationship function constructs in the epidemiological literature relating to social networks. It was these constructs, for example the social network scoring used in the Alameda county studies

(Berkman & Syme, 1979), that have been found to have a strong effect on health. This construct is called "loading" in the model.

We now have a collection of components relating to relationship processes. To summarise, these are the importance of similarity in maintaining relationships, a redistribution of load following relationship breakdowns having the effect of increasing loading on remaining relationships and a limited tolerance of dissimilarity and loading costs in a relationship. We can now assemble these components and produce a statement of the proposed process. An increase in income inequality is reflected in some relationships as an increase in the difference in income between two individuals. In some cases this will increase the dissimilarity sufficiently for the relationship to dissolve. Having lost a relationship, a person may redistribute the loading from that relationship onto his or her remaining relationships. Again, in some cases this redistribution of loading will exceed the ability of the relationship to support the additional loading, and the relationship will be terminated. The result might be a cascading series of relationship breakdowns throughout the population, initiated by a change in income inequality that initially directly affected only a few relationships. This proposed relationship redistribution and the resulting relationship breakdown cascade forms the basic structure for the model.

Building the model

We now have a statement of the idea for the process involved, and we can now begin to design the model. The design will follow the steps outlined at the end of the previous chapter. We begin by considering the purpose of the model.

The relationship formation and dissolution process described above is located at the level of individuals and dyads. Although we can see that the process might aggregate to produce a ripple of relationship breakdowns through the population, it is not obvious what the pattern of the spread of breakdowns might be. For example, it is not obvious whether the proposed process described is self-limiting, whether it would result in the loss of all relationships in particular groups, whether it is only a passing effect cancelled out by new relationships beginning, or whether it might result in a total collapse of all relationships. The model will give us an opportunity to investigate whether the process can result in a series of relationship losses in a population. If it does, we can use it to investigate some of the consequences of a change in income inequality. For example, we might use it to identify who might be affected by these changes, and to see whether anything might mitigate the effects.

As we are interested in income inequality and relationships in a whole population, our target system is a whole population. This population should be largely self-contained, so that very few relationships extend outside the target population. Although the target population will be large, the model does not need to be equally large to effectively model the population. All we are looking for is for the model to be large enough for patterns to emerge.

There is only one external input of interest to us, income inequality. This arises as a consequence of the economic environment, and so is generated outside the target system.

We also have to select an outcome of interest. A simulation allows us to track the status of each relationship at each iteration and so we have a complete history of each relationship. We can summarise relationship data across the whole population in a number of different ways. For example, we could select a structural summary, such as the number of relationships lost, or a summary that reflects the loss of functionality, such as the total loading summed across all relationships. The choice of summary variable should reflect our hypothesis. In this case, we are concerned with a combination of structural and functional elements as this is the form implicated in both social cohesiveness and mortality. Summing the loading across the population obviously includes the functional element, but also implicitly reflects structural changes, as large losses of loading result from the loss of relationships.

Two outcomes are used in the results. One is the sum of the loading across the whole population. This is called the Population Relationship Score in the results. A second outcome is a count of the number of individuals who lose a significant proportion of their relationship loading.

We need to choose the components for our model. The most obvious choice of component for psychologists is the individual, but the unique qualities of individuals are such that they are not always the best choice of component in a system that we hope to simplify. The proposed process gives us some hints as to what components we might select. It suggests that a change in income inequality might reduce the similarity between individuals and so affect their relationships. Therefore we want to include sufficient information about the individuals so that we can determine the relevant similarities (or differences) between them. And we want to include information about their relationships. One way to achieve this is to represent both individuals and their relationships as components in the model.

As the relationships interconnect the individuals, the model has a network form. The relationships are likely to dominate the dynamics of the network, with the characteristics of individuals relatively unimportant to the dynamics. The relative importance of the characteristics of relationships over the characteristics of the individuals allows us to use a very simple model of individuals. This is not to imply that the individuals are all the same, but that we only need include sufficient information about them to be able to determine differences that are significant to their relationships.

The model includes only relationships that are positive in tone. This is partly because we begin with as simple a model as possible. It is also because previous research into social relationships and health has concentrated on positive relationships. The model is based on existing knowledge, and this mostly relates to positive relationships. The relationship is also assumed to be symmetric within a dyad. While this is not always the case, unbalanced relationships tend to be unpleasant for at least one individual (Franzoi, 1996) and might not be expected to be either positive or stable.

In a network simulation, we need two distinct types of variable. One variable should capture the difference between individuals. Another should capture the traffic in the links in the relationships between them. The individuals in the model are characterised only by their location and their income. Although this is a minimal set of traits, it is sufficient to capture some major aspects of similarity and difference between individuals. Activity in the relationships is captured by the loading construct. This represents the costs invested in maintaining the relationship.

In addition to loading, each relationship also has an associated degree of intimacy. In this model this is fixed at its initial value, relationships begin and end with the same degree of intimacy. Individuals have a larger number of less intimate relationships and a small number of more intimate relationships. The degree of intimacy is used to determine the degree of tolerance in the relationship. In the model intimacy characterises the relationship, rather than representing any activity in it.

According to our proposed process, when relationships become overloaded they break. While the loading remains at a tolerable level the relationship will survive, but when this loading is exceeded the relationship will be dissolved. The maximum load that the relationship can take before breaking is a characteristic of the relationship. It depends on the intimacy of the relationship, as we are prepared to put more effort into our more intimate relationships. It also depends on the stress on the relationship, represented in the model by the degree of dissimilarity and distance between the individuals.

When a relationship dissolves, the function that it fulfils for the individual is lost. But people tend to fill up the space that is left by broken relationships. For example, they may spend more time with other people and seek more support from their other relationships. This process is incorporated into the model as a redistribution phase. This will affect people who have lost relationships, who will attempt to redistribute lost loading by increasing the loading on their other relationships. It will also affect individuals who have not lost relationships, but who are connected to someone who has. These will experience an increase in loading on that relationship, and will attempt to reduce their loading on other relationships. To the extent that it is possible, individuals redistribute loadings to keep the sum of their relationship loadings constant. This is not always possible, either because both individuals in a relationship may be attempting different modifications to the loading in the relationship, or because the relationship may itself become overloaded and fail.

The foregoing describes the algorithms of the model. In this form, the model cannot be manipulated or examined, so we need to realise the model in a form that we can use. We should note that the algorithms include a non-linear component; the decision to maintain or to dissolve a relationship is discontinuous at the point where the tolerable loading is exceeded. Computer simulation offers a method to construct and manipulate a non-linear network model. The algorithms of the model provide the basis for a computer programme that will allow us to manipulate and explore the model system.

Once completed, we can set the computer programme running, allowing us to explore the behaviour of the model. Setting the programme running involves starting with a stable network of relationships. When we change the income of a few individuals, this changes the income inequality in the network, and sets the simulation running. We can then observe the dynamic behaviour of the model. The first simulation runs tested the model, and its sensitivity to the parameters chosen for its realisation. I also made some simulation runs to test whether the model behaves sufficiently like a linear system to be considered linear.

Following completion of the simulation runs testing the model, I changed the form of the input to simulate the response to different patterns of change in income inequality. The initial simulations all started with a single large step change in income for the wealthiest individuals. Once the model was tested on this step change, I applied two questions. The first question was whether the effect of a step increase in income inequality was reversed if we restore the original income distribution: will the network of social relationships repair itself if we close the gaps? The second question was whether the rate of change of income distribution would affect outcomes, so are the effects of rising income inequality mitigated by introducing changes slowly.

All of the above has produced a model of a proposed process that might link changes in income inequality with effects on social relationships. But it is only ever a model, and it is only useful within its own design boundaries. For example, the model was designed to represent the dynamic response to a change in income inequality. It does not tell us anything in its static condition, and only becomes informative when it is set into motion.

The set of algorithms above are stated qualitatively. The model makes no claims as to scaling, including scaling of time. It is a qualitative model, even though the mechanics of its realisation depend on calculations in the simulation programme.

Details of the simulation programme

The model has been realised as a MATLAB programme. MATLAB is a high level programming language, specialised for the analysis of arrays, a feature that made the task of programming a network model easier. The programme listings are included as an appendix.

The variables used in the programme overlay their corresponding constructs, but do not exactly equal these constructs. An obvious difference between the constructs in the algorithms and the variables in the simulation is that the constructs are qualitatively stated, but the variables are represented numerically. Their values are not representative of any scaling. Rather they are selected for convenience, so that the simulation matches the qualitative behaviour of the algorithms.

The simulation model population consists one hundred individuals. This population is sufficiently large for aggregation effects to appear, while being small enough for simulation runs to run in a reasonable time.

Characteristics of individuals

The programme has a set of variables describing the individuals, from which we can derive the differences between them. Two variables are used to represent individual characteristics, their income and their geographical location.

Income

Income represents the income of an individual. It has a range of 0-20. Income is allocated to individuals income randomly, so that any patterns are a result of patterns arising in the system, rather than being a consequence of the initial data. Incomes are generated so that the income distribution is approximately normally distributed. This gives a distribution that is simple to generate and that approximates a realistic shape at the high income end, where the income changes are to be made. The population used for the simulation runs had a mean income of 10.2 with a standard deviation of 2.7. The top decile had incomes between 14 and 20.

Position

Position is a vector that represents the location of the individual on a 10×10 map. The geographical position of each individual is randomly generated to obtain an even spread of individuals around the map. Again this is so that any patterns generated are the result of patterns arising in the system, rather than being a consequence of the positioning of individuals.

Calculated from the characteristics of individuals

From the income and position variables we calculate geographical and income distances between the individuals. These distances correspond to the differences between the individuals.

Income distance

Income distance is the calculated difference in income of the two individuals in the relationship.

Physical distance

Physical distance is the calculated physical distance between the individuals.

Characteristics of relationships

Another set of variables describes the relationships. Three direct variables are used to characterise a relationship, the identity of the two individuals linked by the relationship, the degree of intimacy in the relationship and the initial loading on the relationship.

Intimacy

Each relationship has degree of intimacy. This is a whole number in the range 1-5, from the least to the most intimate. For simplicity, the model holds the degree of intimacy constant unless the relationship is broken, in which case it is reduced to zero. The model assumes that relationships do not become more or less intimate in time.

Intimacy is not generated randomly, but is allocated to relationships that are generated randomly. In the initialisation process each individual attempts to find a minimum set of five relationships of intimacy 1, four relationships of intimacy 2, and so on until one relationship of intimacy 5. This corresponds to a small number of close relationships, and progressively larger numbers of more distant relationships. Some individuals finish with more than the target number of relationships, as they may be allocated further relationships to make up a full set for others. Others may fail to find a full set of stable relationships, finishing with less than the target number of relationships.

Loading

The loading on a relationship is initially randomly generated. As the simulation progresses the loading on a relationship may change either as loadings are redistributed or as relationships break and their loading is reduced to zero.

Calculated from the characteristics of relationships

From the degree of intimacy and the geographical and income distances a third variable is derived for each relationship, the maximum loading that can be tolerated before the relationship is breached.

Maximum tolerable loading

This sets the breaking point for each relationship. The maximum tolerable loading is higher in high intimacy relationships, and increases as the distance between individuals reduces. It is calculated for each relationship according to the formula:

$$\text{maximum tolerable loading} = \frac{\text{tolerable.multiplier} \times \text{intimacy}}{(\text{income.distance} + 1)(\text{physical.distance} + 1)}$$

The tolerable multiplier is chosen so that some relationships do break when the simulation is run. It is not intended to have any meaning, but might be thought of as setting the degree of robustness of relationships.

Population relationship score

The loadings of all of the relationships are summed over the whole population to produce the Population Relationship Score. This is a summary value indicating the total level of functionality in the relationships in the population. It is also affected by the breach of relationships, as the loading is reduced to zero in these, and so this produces a combined measure of the total number and function of relationships. The Population Relationship Score is used to report results from the simulation. It could be thought of as indicating the level of social activity in the population.

This is a summary result for the whole population, and it does not give us any indication of the distribution of the losses of loading. The same score could represent an evenly spread loss of loading among all individuals, or a total loss of loading for some individuals.

Time

Time appears in the simulation in the form of iterations. As the model is not scaled, an iteration cannot be transformed into a measure of time. This is common in the simulation of social phenomena where discrete iteration steps do not translate to real time (Chattoe, 1996). We might note that there is no direct means to measure time in the physical world; we measure time by counting the number of occurrences of a regular physical phenomenon. Unlike many physical processes, most social processes are highly variable in duration. Where we measure time through social phenomena, such as through a count of generations, we have to have a high tolerance for variability in that measure.

Initialisation of the network and individuals

The initialisation programme first generates a population with randomly selected locations and incomes. It then sets each individual to find a set of relationships, and tests that all of the relationships are tenable. Following completion of this test, the individuals have a set of randomly selected relationships, all of which are initially tenable. This initial network is static and at equilibrium.

Random relationship making and breaking

This version of stable equilibrium in relationships is not very realistic, as it is static. In reality we live in a social environment where there is a continuous activity of people entering and leaving relationships. This effectively forms a background noise process that will overlay any effect perturbing the system.

If the system behaved linearly, we would expect to see this normal process of making and breaking relationships simply adding to any effect of a change in income inequality. This would look identical to the result without a background process of making and breaking relationships, but with noise superimposed. If this were the case, we would not have to consider this background process further as the system would be behaving linearly. Using the principle of superposition, we could derive the results from noise and changes in income inequality separately, and simply add them together. On the other hand, we might intuitively expect the normal process of relationship formation to replace lost relationships and prevent a wave of relationship breakdowns. If so, our intuition is leading us to expect a non-linear process.

The simulation model includes a subroutine for producing this continuous background noise of making and breaking relationships. This noise routine can be incorporated into the simulation if we find that the process is not linear. The initialised network with this noise process running remains stable, giving us a choice of starting points.

Running the simulation

The simulation begins at an initialised stable equilibrium, with or without the background noise. To increase income inequality, we increase the income of individuals in the highest income decile. The size of this increase can be varied, again according to convenience. As with other quantities, the size of this increase is not to scale. Increasing the income of the top decile results in an increase in income inequality with a similar pattern to that found in New Zealand. This sets the relationship breaking and redistribution process into action.

Following the increase in income, some relationships will no longer be tenable, as the similarity between people will be reduced in relationships with the highest income decile. To detect this, the programme recalculates the tolerable loading for each relationship. The tolerable loading will be reduced for relationships where the income distance has increased. The programme then compares the actual loadings with the recalculated tolerable loadings and terminates any relationships that are no longer tolerable. This ends the simulation of the system's initial response to the change in income inequality.

The programme then moves to the redistribution process. It starts a series of loops through a cycle in which individuals attempt to redistribute the loadings from lost relationships onto remaining relationships. If an individual has lost a relationship, they will distribute the load that was on that relationship to their remaining relationships. This will result in the remaining friends of someone who has lost a relationship all having load added to one of their relationships. In turn they will redistribute their loads, by trimming a little loading off each of their other relationships.

For example

if A has relationships with B, C, and D

and B has relationships with A, E, and F

if A and C break up their relationship A will spend more time with B and D

and B will balance this by spending less time with E and F.

The programme then tests whether A's relationship with B and D will tolerate this extra loading. If either cannot, that relationship will also break up.

Again the programme compares relationship loadings with the tolerable loadings and removes any which are no longer tolerable. A series of fifty loops through the cycle is sufficient for the system to settle following an increase in income inequality.

Throughout this process the programme logs the loadings on each relationship. We can extract the total loading across the whole population at the end of each cycle from this log. This gives us the Population Relationship Score reported in the simulation results in the next chapter.

Chapter

6

Simulations

This section contains two stages of output results from running the simulation. The first continues the development of the model from the previous section by testing its linearity. The results of this first stage illustrate how we can further develop a model as we gain some understanding of its behaviour. The second group of runs apply the simulation model to ask whether the process is reversible, whether making changes gradually reduces the impact on relationships and who is affected by relationship breakdowns. These runs are intended to demonstrate how we might use such a model to answer questions.

Stability of the model

I first tested the model for its ability to produce a stable outcome. We would expect that the overall activity in social relationships within a real population would be stable around an attractor, that is that it would remain within a certain range. Following a perturbation we would expect it to settle eventually to a new stable position, as we would not expect to see a total social collapse. The model should reflect this. An unstable system model might show an inexorable trend toward elimination of all relationships, or an inexorable trend toward everybody having a relationship with everyone else, or a trend of ever more extreme alternating gains and losses. An alternative might be that a system, although neither stable nor unstable, generates an entirely random outcome.

At this stage, it was possible that triggering the process of reloading and breaking relationships might result in a total collapse of the network of relationships in a population. Such a collapse would be extreme and inconsistent with the existence of any social activity, and would suggest either that the model is missing some stabilising factor, or that it is simply wrong.

The first two output results in figure 8 and figure 9 show the total loading on relationships in the population through fifty iterations of the simulation.

Figure 8 shows the response of the system following a 100% step change in the income of the top 10% of the population. The Population Relationship Score, representing the total loading on all relationships in the population is plotted on the y-axis. The number of iterations through the cycle of redistribution of loadings is shown on the x-axis. This axis represents the passage of time, but is not scaled.

This simulation run includes only the process of making and breaking relationships resulting from the increase in social distance. It excludes a normal background noise of beginning and ending relationships.

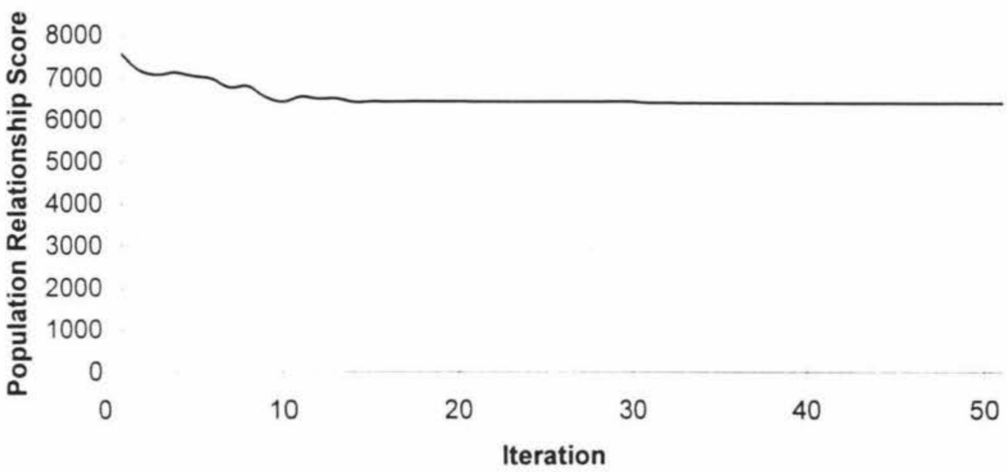

Figure 8. History of population relationship score following an increase in the income of the top decile (income increment=100%, tolerable multiplier=60, probability of relationship gain or loss during the iteration=0)

At iteration 0, the incomes of the top decile are increased suddenly. This equates to a sudden change in income inequality. Figure 8 shows that, following a process of breaking and redistribution, the social network eventually settles to a stable level after about 15 iterations. The final value of the Population Relationship Score resulting from the increase in income is approximately 15% lower than the starting value. With the caveat that this model is not scaled, we might think of this result as showing that the effect of increasing the income of the top decile is that everyone loses one day a week of their social activity, or that 15% of people lose all of their social activity.

Figure 9 shows the initialised population with a simulated background noise of beginning and ending relationships, but with no changes in incomes. This shows a process that is noisy but, again, stable.

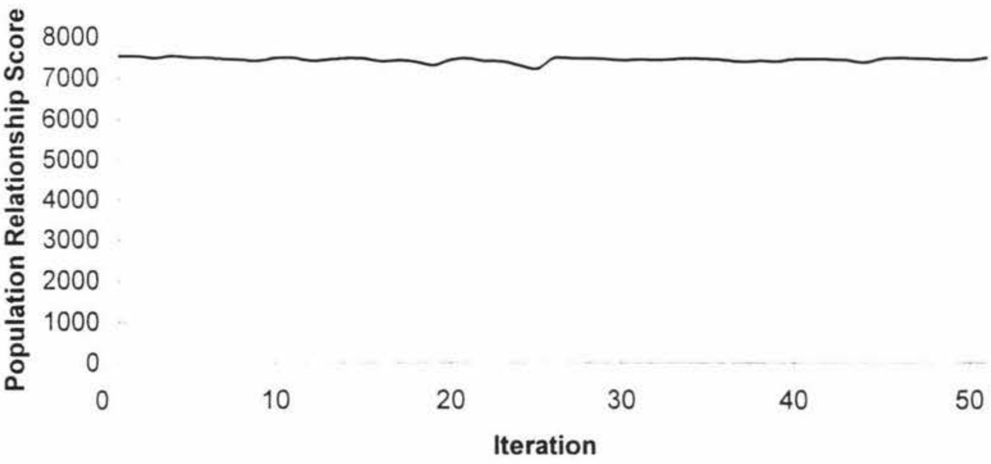

Figure 9. History of population relationship score with a background noise of beginning and ending relationships (income increment=0%, tolerable multiplier=60, probability of relationship gain or loss during the iteration=0.01)

These simulation runs give us reassurance that both the change in income inequality and background noise processes can produce stable outcomes. The system is realistic to the extent that it does not collapse entirely under either process.

Linearity

One of the aims in building the model is that it should be as simple as possible, and only as complex as necessary. One way of reducing the complexity of a linear model is to split the inputs into simpler input signals, and to add the results together to derive a complete response. This makes use of superposition in linear systems, and so it cannot be used where a system is non-linear within the range of interest.

Although we know that the model includes a non-linear process, the decision to terminate a relationship, it is still possible that the system behaves linearly in the range of interest to us. We need to test the model to find out whether this is the case.

With this in mind, I have investigated the linearity of the model across a range of values for the tolerable loading multiplier factor, and across a range of increases in income. The results of these tests is shown in figure 10 and figure 11.

For both figure 10 and figure 11, the y-axis now represents the difference between the initial population relationship score at iteration 0 and the population relationship score at iteration 50. The x-axis represents the size of the increase in income (figure 10) and the size of the multiplier in the calculation that determines the tolerable loading (figure 11). These graphs show that the outcome is sensitive to the input variable and to the only constant applied in the simulation.

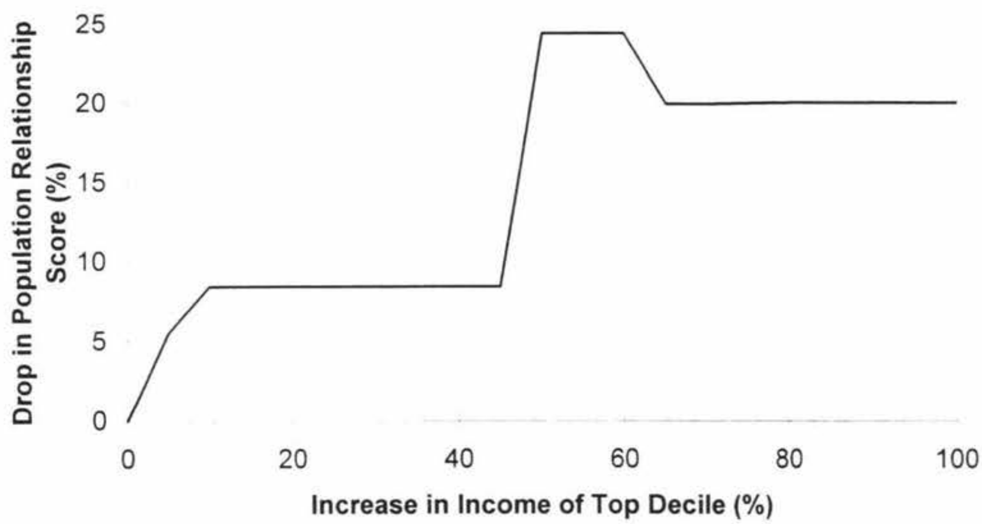

Figure 10. Dependence of size of drop in population relationship score on size of income increase in top decile

Figure 10 shows the effect of the size of the increase in income for different sizes of the increase in income. It shows a distinctly non-linear pattern, with three distinct step changes. Over a large proportion of the range, the loss in relationships is insensitive to the size of the increase in income, but over small parts of the range the loss of relationships is very sensitive to the size of the increase. The relationship is strongly non-linear across a wide range of values, and the system is very sensitive to the size of the income increment. This non-linear response to the input variable is a very clear warning that we are very unlikely to be able to assume any linearity in the system.

Figure 11 shows the sensitivity of the loss in relationships to the size of the loading multiplier, which we might think of as representing the robustness of relationships.

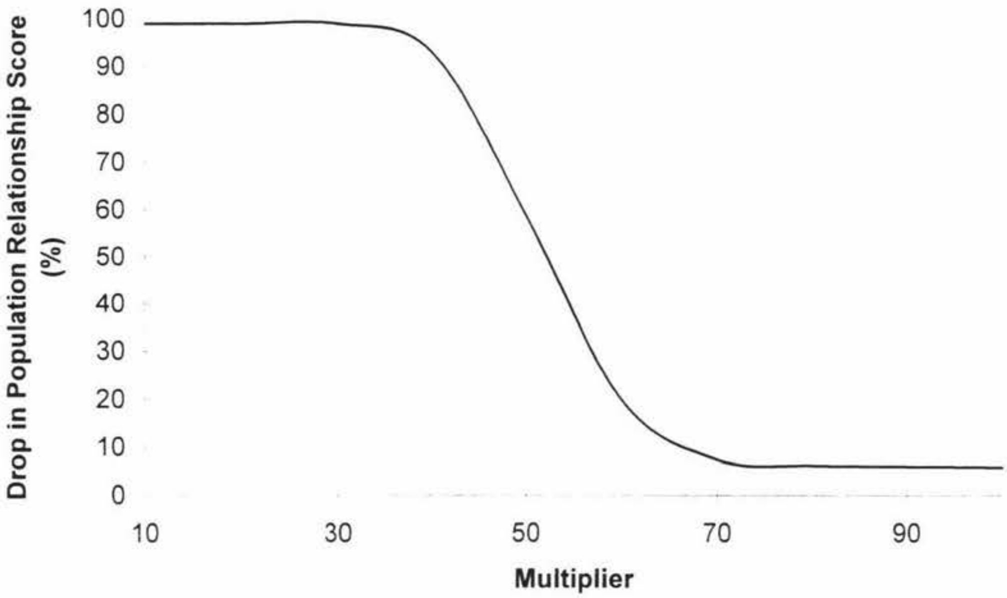

Figure 11. Dependence of size of drop in population relationship score on tolerable loading multiplier

Again figure 11 indicates a non linear relationship. In this case the non-linearity is less concerning. In part this is because our model is based on the assumption that decisions to terminate a relationship are based on loading and the differences between people. It is not based on any mechanism affecting the robustness of relationships. Also, there is more potential for the relationship to be linearised within some ranges: we could assume linearity in each of the ranges between 10 and 40, between 40 and 70 or between 70 and 100.

Nevertheless, both relationships are non-linear, and so we cannot assume that we can simply add the effect of an increase in income to the background noise.

Figure 12 shows the inclusion of both processes in the simulation. It incorporates both an increment in income of the highest income decile and a background random making and breaking of relationships. This plot also shows for comparison the earlier results from figure 8, the income increment alone, and from figure 9, the background noise alone. If the effects of income increment alone and

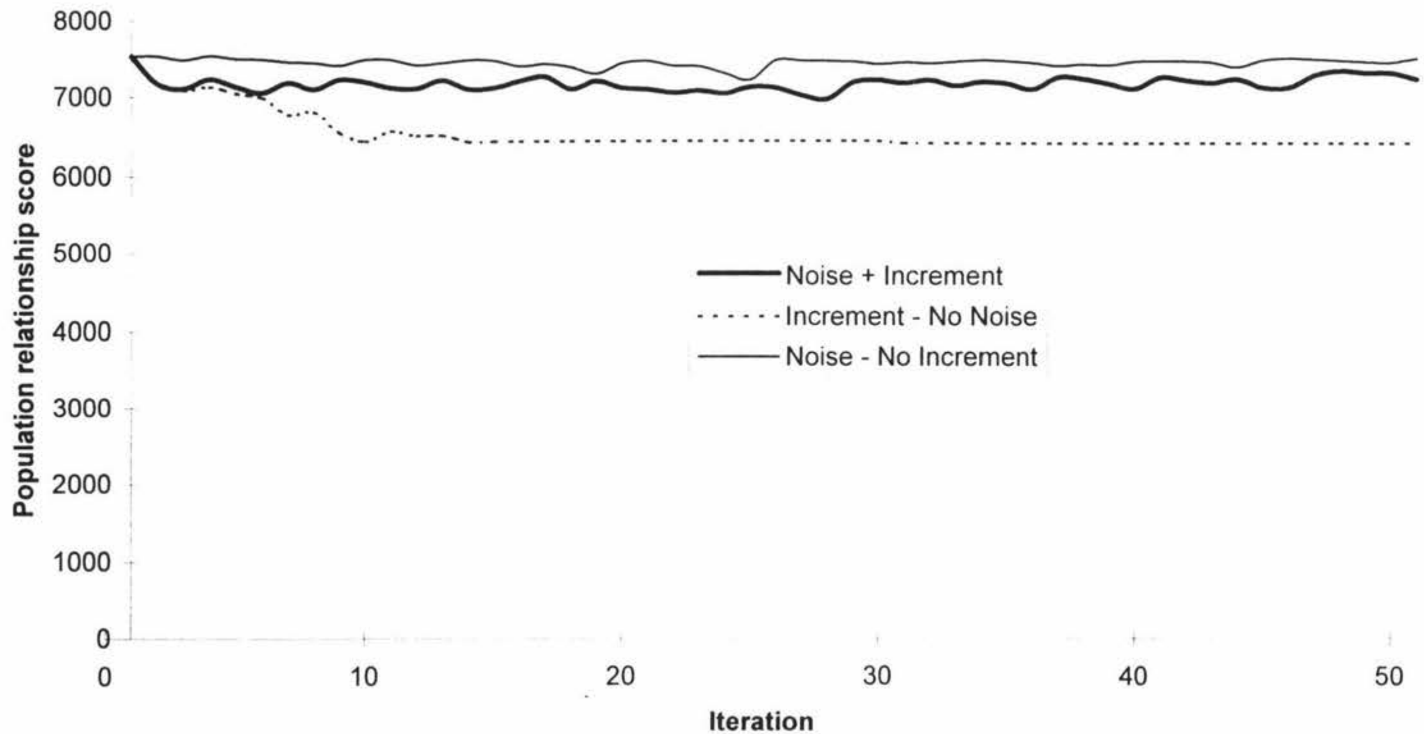

Figure 12. History of population relationship score with and without noise (income increment=100%, tolerable multiplier=60, probability of relationship gain or loss during the iteration=0.01)

background noise could simply be added, the process should look like the "income increment - no noise" but with added noise from the random making and breaking of relationships. Instead the combined process shows the that the mean population relationship score is considerably less affected than it would have been without a background process of making and breaking relationships. This demonstrates that we do need to include the background noise process in our simulation model. Having completed the model, it is now time to try putting some questions to it.

Reversibility

We might assume that if increasing income differences has consequences for health, restoring the previous income distribution will reverse the process. If we consider the model simulated here, some relationships become stressed by increasing the income differential and may breach. Reversing the change in income differential would not be expected to restore those lost relationships. But a reversal might allow a greater proportion of new relationships to survive, and the lost relationships might be restored through the make and break process.

Figure 13 shows the output of a simulation in which the income of the top decile is doubled at the first iteration, and halved again at the fiftieth iteration. This restores the original income distribution at the fiftieth iteration.

The output shows that the level of population relationship score remains the same following restoration of the original income distribution at the fiftieth increment.

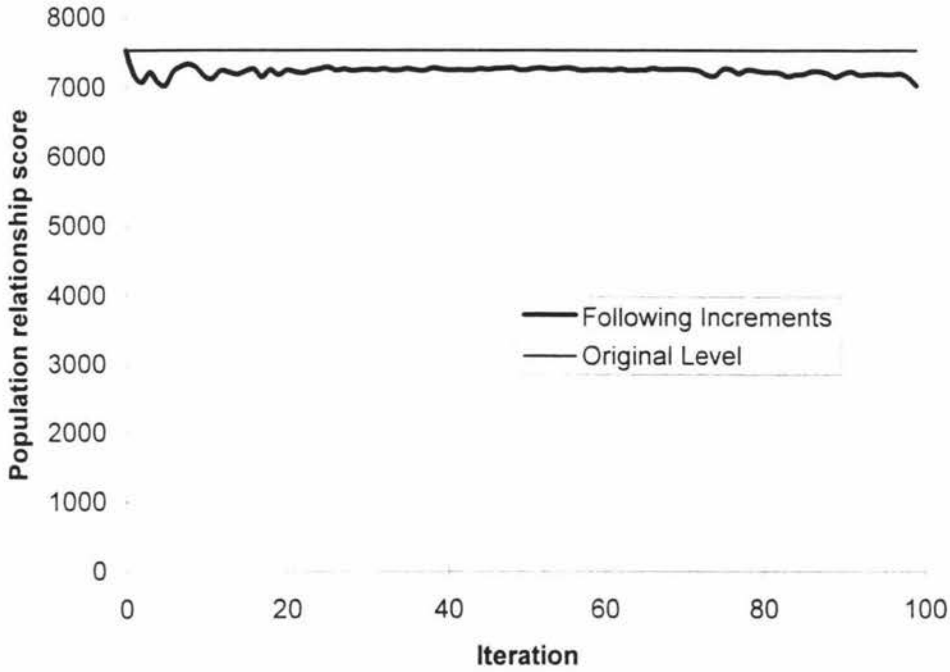

Figure 13. History of population relationship score with change in income inequality introduced then reversed (income doubled at t=1, halved at t=50, tolerable multiplier=60, probability of relationship gain or loss during the iteration=0.01)

Suddenness of change

Some writers have indicated that the rate of change may be important in determining health outcomes, and that the shock of rapid change may overwhelm the resources of individuals and communities.

To test this, I ran a simulation in which the income inequality was increased gradually over ten iterations. The result was compared with the previous result, resulting from a single step change. As before, the single step change was a doubling of the income of the top decile. The gradual change consisted ten steps in which the income of the same group was increased by 7.2% in each iteration. This compounds over ten increments to give a total increment of 100%. The result of this slow increase compared with a sudden increase is shown in figure 14.

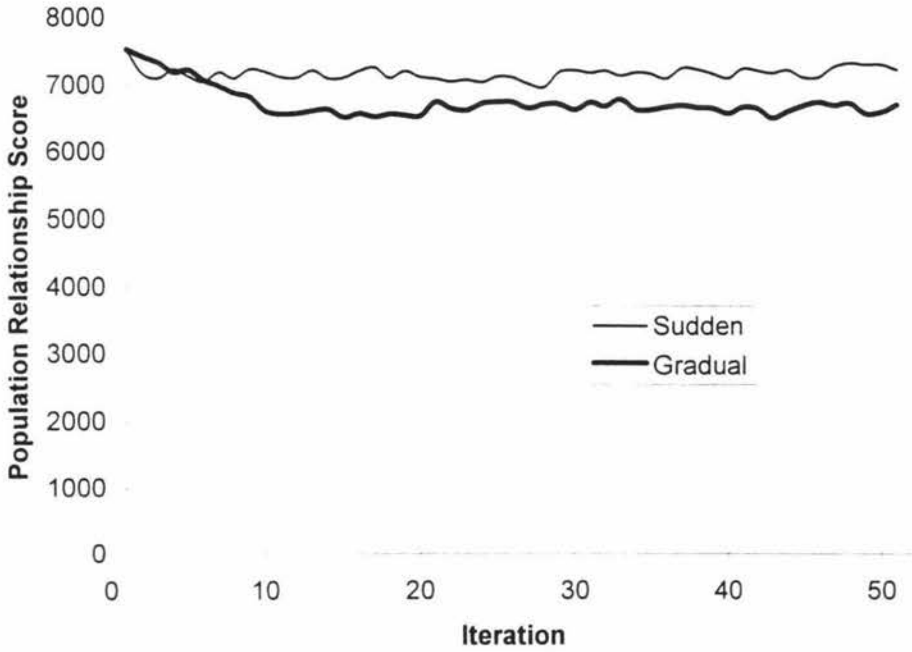

Figure 14. History of population relationship score comparing sudden increase (income doubled at t=1) and gradual increases (income increased 7.2% at each of ten intervals) (tolerable multiplier=60, probability of relationship gain or loss during the iteration=0.01)

Gradually increasing income inequality substantially increases the loss in relationships over the effect if the same change in income distribution is applied suddenly. The system responds anew to each increment, rather than to the overall size of the change.

Distribution of loss of loading

The previous graphs used the Population Relationship Score as an indicator to summarise the overall change in relationship function in the whole population. The following tables use the individual histories to identify which individuals have lost a relatively high proportion of the functionality in their relationships. The numbers in the tables are a count of the number of individuals that have lost more than 10% of their initial loading.

We might expect that the group most exposed to loss would be the highest income group, as their relationships that extend outside the top decile will be the first relationships stressed by income inequality. One epidemiological study found that the mortality effects of income inequality extended through all quartiles (Ben-Shlomo et al., 1996). This would lead us to expect to find disrupted social relationships in both high and low income groups.

We saw earlier that a background noise of making and breaking relationships has the effect of reducing the loading lost following an increase in income inequality. One possible protective mechanism behind this might be that newly appearing relationships act as a baffle, dampening the outward spread of relationship breakdowns. This would not act to protect the most exposed relationships at the source of the relationship breakdown chain, but might act to protect relationships in lower deciles. Table 1 shows the effect of background noise for relationships of high and lower income individuals.

Table 1. Effect of income and noise on the number of individuals affected by a high loss of relationship score following a sudden change in income inequality.

		Individuals with a high loss of relationship score	
Income	Number of individuals	Background noise	No noise
Low & Medium Income	67	3	16
High Income	33	10	15
Total	100	13	31

High income individuals are significantly more likely to lose a high proportion of their loading, no matter whether there is a background noise of relationship making and breaking or not. The addition of background noise is protective for both high and low income individuals, but the protective effect is greater for the low and medium income individuals.

We also saw that the effect of making changes in income inequality slowly was to increase the overall loss in relationship loading. It is possible that this effect is not evenly distributed across individuals of high and low incomes. Table 2 shows the effect of making changes suddenly or slowly.

Table 2. Effect of income and rate of change of income inequality on the number of individuals affected by a high loss of relationship score in an environment with background noise.

		Individuals with a high loss of relationship score	
Income	Number of individuals	Sudden change	Slow change
Low & Medium Income	67	3	13
High Income	33	10	17
Total	100	13	30

Again, high income individuals are significantly more likely to lose a high proportion of their loading, no matter whether changes are made suddenly or not. Making changes suddenly is protective for both high and low income individuals, but the protective effect is again greater for the low and medium income individuals.

Many individuals with a high level of relationship loading have a larger number of relationships. It is possible that this results in a smaller increment in the loading onto their other relationships when redistribution is necessary, and a consequent reduction in the tolerance required of remaining relationships. If so, an initially high level of relationship loading might be protective against the loss of relationships cascading through the population following an increase in income inequality. The effect of an initially high loading score under varying background noise conditions is shown in table 3.

Table 3. Effect of initial relationship score and noise on the number of individuals affected by a high loss of relationship score following a sudden change in income inequality.

		Individuals with a high loss of relationship score	
Initial relationship score	Number of individuals	Background noise	No noise
Low-Med Score	67	7	17
High Score	33	6	14
Total	100	13	31

The effect of an initially high loading score with different rates of change of income inequality is shown in table 4.

Table 4. Effect of initial relationship score and rate of change of income inequality on the number of individuals affected by a high loss of relationship score in an environment with background noise.

		Individuals with a high loss of relationship score	
Initial relationship score	Number of individuals	Sudden change	Slow change
Low-Med Score	67	7	21
High Score	33	6	9
Total	100	13	30

There is no significant difference in the proportion of relationships lost by individuals with low and medium initial loading scores compared with groups with high initial loading scores.

This completes the demonstration of how we might run simulations to ask questions of the model. It does not represent an exhaustive exploration of all of the possibilities presented by the model. Rather it demonstrates some of the possibilities for exploring the available data at both an aggregate and an individual level. We now go on to demonstrate how we might interpret the results from these simulation runs.

Chapter

7

Discussion

Once we have assembled results from simulation runs, the methodology proceeds as we would using conventional techniques. We interpret the results. While the process of interpretation will be the same as we use for traditional approaches, the vocabulary that we use will be different, as it will reflect the dynamic nature of our model.

Where we have evidence available, we should compare the results with the available material and assess its performance. In doing so, we need to recall the design of the model, as this will determine the range over which we expect the model to reflect the target.

Results of the simulation runs

The first runs tested the behaviour of the model under a range of parameters. In these runs the model was stable, as it moved to a stable outcome following a perturbation. This outcome is consistent with the network of relationships in a real population. The stability of the model held under a wide range of income increment step sizes, and under a range of parameter values for the sensitivity of individual relationships to differences between the individuals. Although the model was stable in the ranges tested it was strongly non-linear in relation to both of these variables.

The model retained its stability with the addition of background noise. In this case, noise represents a normal process of making and breaking of relationships that occurs in a population independently of change in the external environment. We saw in figure 12 that the effect of adding a normal background process of making and breaking relationships was to prevent a larger loss or relationship loading. While at an individual level the immediate experience of breaking a relationship is unpleasant, the break-up may in the longer run be seen as a positive outcome. There is no equivalent positive perspective placed on a high rate of relationship breakdown at the population level, where high and increasing divorce rates are more usually interpreted as indicative of moral and social breakdown.

In contrast, the result in figure 12 suggests that there may be benefits to a high acceptability of breaking relationships. The rates of making and breaking relationships must be approximately equal, unless there is a trend either toward everyone having no relationships or alternatively toward everyone having a relationship with everyone else. A high rate of breaking relationships therefore allows a correspondingly high rate of making new relationships. These may insulate against a cascading breakdown when external drivers stress some relationships. This insulating effect of noise is likely to work through new relationships relieving stress in the system and damping the cascade of load distribution. It may be that a higher rate of making and breaking relationships is necessary to reduce the overall extent of lost relationships in a population in times of change.

This protective effect of the background noise was found as a by-product of the model development and testing process, and illustrates how the process of development and testing can itself be fruitful. Once the model was completed I applied two specific queries to illustrate how a simulation model might be used to answer questions.

The first of these asked whether, having made changes in income distribution, reversing these changes would restore the original level of relationship function in the population. The simulation tested this through doubling the income of the top decile at the beginning of the simulation run, and halving it again at the fiftieth iteration. If the process were linear, driven by income distribution as an input, we would expect that it would track income inequality through both widening and reducing income inequality. The existing models of income distribution, social cohesiveness and health are implicitly linear, so application of these would suggest that the process of loss of relationships is reversible simply by restoring the original income distribution. Figure 13 suggests that this may not be the case. The relationship score drops following the increase in income but there is no restoration of the population relationship score in the period following the reversal of the increase in income inequality at the fiftieth iteration.

The result is counterintuitive, in that we tend to assume that processes are reversible, without necessarily considering that this reversibility depends on the linearity of the process. As we rarely address non-linearity in analysis, so we also rarely consider its potential consequences. If we consider this particular process solely in terms of its linearity at the level of the smallest system component, we would not expect the process to be reversible. There is, after all no reason to believe that restoring similarity would, in and of itself, repair relationship breaches.

If this is the case, there must be an entirely different process by which social cohesiveness can be built up. We cannot assume that social cohesiveness will develop without some process to build it. It may be that reducing income inequality is not sufficient to reverse the loss of relationships, but is necessary to enable other processes to restore social cohesiveness. The model cannot tell us anything about what those other processes might be, as it is not designed to do this.

If we compare this result from the model with a finding from longitudinal studies, the result given by the model appears to be counter to the experience of the United Kingdom during the Second World War. In this period, social cohesiveness increased while income inequality was reduced (Wilkinson, 1996). It is, however, possible to identify other concurrent processes that were likely to have built a sense of identity, unity, and common purpose characteristic of social cohesion.

The second illustrative question asked whether a more gradual introduction of the change in income inequality might lessen the relationship effects by reducing the size of each shock. Rather than reducing the effect of the change, the loss of relationships responded to each step increase in income inequality, with the result that the same overall change had a much greater effect when introduced slowly than when it occurred suddenly. Effectively the system is reacting rapidly to the incidence of each change, rather than to the overall size of the change, and gradual change becomes a drawn out series of shocks. This is reminiscent of Durkheim's

(1951) suggestion that suicide rates are a result of change itself, rather than depending on the exact nature of the change.

This simulation result is again perhaps difficult to understand if we consider it from an individual level. We might expect that individuals would adapt more easily to a number of small changes than to one large change. But if we consider the effects on an organisation of changes, such as restructuring, that are implemented over an extended period, these results are perhaps less surprising. In common with this model, organisations work as systems of individuals, and actions directly affecting a small number of individuals have much wider effects throughout the organisation. Drawn out changes can be disruptive whether in an organisation or in our model population.

The final question was whether particular groups are more exposed to losing relationship loading than others. It made no significant difference whether individuals began with a high or a low level of relationship loading, both were equally at risk.

The loss of loading was more concentrated in the higher income group than in lower income groups. This difference was significant whether or not there was a high level of background noise, and whether or not the income changes were made suddenly. High levels of background noise and making changes suddenly were most protective for lower income groups. Under conditions of low noise or gradual change individuals in all groups suffered high losses of loading. This is consistent with the reported (Ben-Shlomo et al., 1996) mortality effects of income inequality appearing in all deciles.

Network model variations

As we create models, their design is not fixed. In the process of working with a model we can trigger other ideas about the processes that we are exploring. These can either add to the foundation provided by the existing model, or they can trigger ideas of alternative models giving approaches. The following illustrate some possibilities for these.

Groups as components

The theoretical components from which this model is constructed are entirely individual, drawn from psychological knowledge of individual behaviour. Unlike the more usual application of knowledge about individual behaviour, the model attempts to accumulate individual behaviours to form a population system. In this system model a relationship trajectory can be tracked for each individual and the modelling essentially models individuals. But in this analysis each individual is embedded in the system rather than being isolated from the system.

Being based on the behaviour of individuals, the model represents a psychological orientation in its analysis. In doing so, it perhaps maintains the largely artificial boundaries between social sciences. This orientation has influenced the choice of individuals and their relationships as components in the model, so the model does not incorporate group entities. But the construction of this model using exclusively individual level components does not represent a denial that group interactions may also be an important determinant of the behaviour of the system. These could also be incorporated in the model.

For example, the model described is two dimensional, having only one layer of essentially identical individuals. More complex social structures could be incorporated by assembling a multi-layered model in which other social structures might be represented as layers, with linkages between layers as well as within layers. Groups might then be associated with individuals between layers, and with other groups within a layer and between different group layers.

Alternatively, some writers have noted that groups may themselves be important individual entities in social networks, and may even substitute for individual relationships (Felton & Shinn, 1992). An organisation might be modelled as a different type of individual, with a set of characteristics suited to its nature.

The population

This model used a population of one hundred individuals. It does not represent a closed population of one hundred individuals. The relationship findings that were used to generate the processes in the model were developed in much larger populations, and are unlikely to reflect processes at work in a small, isolated population. This population size was chosen for convenience, as it was large enough to generate a system effect, while being small enough for simulation runs to be completed in a reasonable time. Each doubling of size of the population increases the run time fourfold, as the number of elements in the matrices varies as the square of the size of the population. The effects appearing in this model also appeared in a smaller population of fifty individuals. While it is possible that different patterns may appear in larger populations, this is unlikely as the smaller system is likely to represent a larger operating under the same rules. This is due to the nature of patterns arising in complex systems: the patterns in a large system reflect the patterns that appear in its subsystems.

The individuals in the base population used here are all essentially identical, except for their income and location. We could increase the number of individual characteristics modelled. Other simulations could incorporate other aspects of individual differences such as gender or age differences.

The simulations here were all based on a single, randomly generated population. We could also manipulate distributions to reflect features of interest. Rather than using a population with randomly generated features, we might assemble a population that models features of interest in real populations. For example, we might manipulate the population to represent physically isolated groups with only a few relationships linking these groups.

At this stage of its development, the model has individuals in fixed locations. It is unlikely that we would retain this as the model was developed, as it excludes a process that is potentially important. The creation of wealthy enclaves has been suggested as potentially important in generating the effects of income inequality on social cohesiveness. The simulation has showed that the high income group was the most exposed to losing relationships following an increase in their income. Where people are able to move closer to their friends this may add to the tendency to form enclaves. Mobility could be added by allowing the individuals to move to minimise the distance to their friends.

Other network approaches

Other techniques exist in the social sciences for the study of networks of relationships (Wasserman & Faust, 1994). These provide tools for the analysis of relationship data where this has been collected in the course of research. While the techniques are generally applicable, they are necessarily limited to data sets which can be collected in practice. In this case, we can rarely assemble a complete set of relationships in a population. This model gives us a means to generate population data, and network analysis techniques may be useful in the analysis of data generated by the model. In the results I have used only the Population Relationship Score to indicate the system state, but the techniques for social network analysis offer alternatives.

The presentation of one model here does not imply that this is the only possible model for considering how income inequality and health might be related. The network modelled here is only one possible combination of fragments of theory from psychology. Other models might better model other facets of relationships. Different features of relationships have been modelled by other writers. Both Granovetter (1973) and Wallace (Wallace, Fullilove, & Flisher, 1996; Wallace & Wallace, 1997) have used models of relationship networks to investigate aggregation processes. But, unlike this model, their models were concerned with the role of relationships in the transmission of information.

Granovetter's (1973) model is the simpler of the two. It characterises relationships only as 'strong' or 'weak'. The model incorporates only one rule; if there is a strong relationship between individuals A and B, and between B and C there must be some relationship between individuals A and C, or individual B will find the situation intolerably uncomfortable. No corresponding rule exists for weak relationships. The aggregate consequence of this simple rule is that strong relationships tend to form people into isolated cliques. Weak relationships are the essential links between otherwise isolated groups. Granovetter (1973) notes that the implication is that weak ties are essential for both social mobilisation and for the transmission of innovation.

Wallace and colleagues (1996) model is much more complex, but retains an elegance despite this complexity. Their model reveals a social network model with a mathematical similarity to models for the coding of information to maximise transmission of information. This model represents the social network as a communication channel, and behaviours as information codes. Wallace suggests that the nature of the network and the constraints of the environment may tend to adaptively select behaviours which maximise the effectiveness of the communications channel.

In both of these models the network is dominated by the characteristics of the relationships linking individuals, rather than by the characteristics of individuals.

A vocabulary of dynamics

Working through this example has demonstrated some alternative ways of talking about ideas in psychology. At a fundamental level this calls on us to decide whether we will talk about a phenomenon in terms of stochastics and randomness, or whether we will talk about it in terms of deterministic processes and complexity.

In developing the example, I have used a vocabulary of dynamic systems. This vocabulary has run through the formulation of the model, through to the discussion of results. At the earliest stages, the language was concerned with processes, and how the outcome of these processes in response to change. I have considered whether the processes were reversible, the alternative being that they exhibit hysteresis. More generally I have considered the linearity of the processes, and the consequences of their non-linearity. The model exhibits a strongly non-linear characteristic in relation to the input variable. It produces stable outcomes, but these outcomes move to different stable points or attractors. This different vocabulary reflects different ways of thinking about phenomena and processes. In the final chapter I will comment on what working through the example suggests about the way that dynamic systems methodologies might lead us to think about health psychology.

Chapter

8

Conclusion

The last few chapters have been concerned with the example; the development of a model and the presentation and discussion of the results from it. The experience gained from this gives us the opportunity to assess the benefits and disadvantages of dynamic systems simulation.

Benefits of dynamic systems and simulation approaches

In the first chapter I identified three distinctive features that might indicate a significant dynamic systems component, and where a dynamic systems approach might be worthwhile. In the example I focussed particularly on aggregation and dynamics, but the third feature, multidimensionality, was also exhibited by the example.

The example exhibits two aspects of aggregation. Firstly, it requires analysis to take place at two levels of aggregation, as income inequality and social relationships exist at two different levels of aggregation. Secondly, there is evidence suggesting that the process of aggregation may itself be an important factor. The example demonstrates that integrating individuals into a system, rather than considering them as socially isolated groups or individuals, gives us an aggregation process. A closer look at the example suggests reveals that these two aspects of aggregation are different facets of the same process, and both are addressed when we take a systems approach.

In the example, the model is based on some observations about individuals and their relationships. Engaging these observations as rules governing behaviour in the model system generates a response in the whole system. Had we considered the effect of these observations on an individual by individual basis, we could not have found the patterns that we did.

A dynamic systems approach of assembling individuals into a complete system aggregates the individuals into larger groups, without making the assumption that we can simply add the individuals together. When we do this, and identify some simple rules for the operation of the system, rich patterns can emerge from these assemblies of simple components (Holland, 1998). These patterns are a consequence of the complex interaction of components, and will only arise as a result of the aggregation process.

The patterns that arise are not only cross-sectional, they also develop longitudinally. Exploring the system through simulation can give us a history of the what happens as the process develops. While we have not done so in this example, it is possible to track the whole process through the system and detect patterns in its spread and development. The simulation model is not simply a means to aggregation, it also produces a dynamic model.

In the example we were particularly interested in the effects of a change in income inequality. The effects of changes develop over time, and we need methods that can address this time component. Statistical tools to do this require a series of twenty or more data points (Gujarati, 1995), and it is often impracticable in the social sciences to meet these data requirements. Even if they were practicable, many of these tools are concerned exclusively with finding patterns in data, and are disconnected from theory (Gujarati, 1995).

Theorising about the effects of a change, or the unfolding of a process in time is challenging. We need to understand the dynamics of a phenomenon to understand these, but have few ways to conceptualise dynamics in social systems. This inability to theorise about dynamics can generate a reluctance to explore how situations develop over time.

The processes lying behind phenomena drive their dynamics, so to understand the dynamics we need to explicitly consider these processes. A dynamic systems approach forces us to do this, as the processes form the basis for dynamic models. This approach gives us the means to theorise about dynamic phenomena. I noted earlier that the process of aggregation produces a dynamic model. One reason is that the rules that we use to build the aggregated model relate to the processes that link the individuals. The strengths of dynamic systems models for conceptualising both aggregation and dynamics are founded in the focus on the processes involved.

Understanding dynamics is important in understanding the response to change, but it is also important in understanding how normal processes unfold over time. This is important in health psychology, as health effects are often a result of long term processes. Linear correlational models lead us to expect that extreme outcomes are the consequence of extreme inputs. In dynamic systems extreme outcomes do not depend on abnormality or extremes in characteristics or inputs. Rather they can be the result of processes operating in a normal range. In the example, the processes of relationship making, maintenance and breaking does not include any processes that are abnormal or extreme. Despite this, we see that the consequences of the process operating in the system can be extreme for some individuals, with a number losing all of their relationships. For some of these, the loss of all relationships is a consequence of their extreme income position. For others it is a consequence only of their position in relation to surrounding relationships. Dynamic systems approaches allow us to explore how effects can develop in a system in the long term. In non-linear systems these effects can appear quickly if the system moves to a new stable point.

The income inequality and social relationships example also has an element of multidimensionality, as it includes elements in economic and social dimensions. Again, the model addresses interactions in these two dimensions through being specific about the process involved. In that process, economic factors drive increasing income inequality. This results in increasing differences between some linked individuals, as a result of their differing incomes. In turn, increasing differences has consequences for social relationships. This process directly links factors in economic and social dimensions and integrates dimensional layers into one picture.

I originally identified three features for which dynamic systems approaches may be appropriate. A dynamic systems approach accesses all of these by considering the process lying behind the phenomenon. This highlights another advantage of dynamic systems approaches; they give us another way to build theories in health psychology.

We currently use theory building techniques and a style of theory that is consistent with the static theories that we commonly work with. In these, we often use diagrams to help clarify our theoretical ideas, but while these can be useful in helping us to clarify our ideas, they can also lure us into some ways of thinking about theory in health psychology (Spicer & Chamberlain, 1996). The form diagram that we use is essentially activity on arrow. The arrows implicitly require some process through which one variable and its corresponding construct are correlated with a second variable and corresponding construct. But drawing constructs in boxes invites us to explore and unpack the constructs rather than to attempt to unpack the process: arrows do not make inviting candidates for unpacking. This contrasts with systems diagrams, which focus attention on the processes that lie behind the correlation.

Using dynamic systems modelling there is no short cut to lure us away from doing the theoretical work. Using conventional theory building techniques, a good model has substantial overlap between its theoretical formulation, its quantitative formulation and its statistical formulation. The model depends on this overlap for its construct validity and its internal validity. But the overlap can allow us to drift between these formulations, and in doing so we can be distracted from doing the theoretical work by doing our thinking at the quantitative or statistical level (Spicer & Chamberlain, 1996). Building a dynamic system model forces us into a theoretical mode in which we consider the processes involved and their interactions.

When we are looking into dynamic processes, we need knowledge on which to base our models. This does not have to be drawn from previous work in dynamic processes. For example, the observations used in building the model for the example were drawn from previous work using conventional research methods. I do not propose dynamic systems methodologies as a replacement for existing methodologies. To the contrary they depend on evidence collected by those methodologies.

To build a dynamic systems model, we particularly need information on processes. When we focus on processes, we find that we begin to ask different questions. Returning again to the example, the model is predicated on knowledge about the processes surrounding relationships. It includes a redistribution process, in which people respond to relationship breakdowns by redistributing their time and effort in other relationships. There is relatively little data on this redistribution process. The data that is available is a by-product of a different question, in which researchers asked what people do to recover after a relationship breakdown. Focussing on processes can lead us to different questions in conventional research. It may be that the exercise of attempting to build a dynamic systems models itself leads us to ask better questions using other methodologies.

A problem with traditional methods is that they have generated a large pool of fragments of theory, but offer us few means to assemble those fragments. One way that we have attempted to cope with this fragmentation has been to add more and more facets and constructs to flowcharts and more and more variables to the data in the hope of a complete picture appearing. This has resulted in an ever growing set of pieces, and the impression that the complete picture is ever more elusive.

In contrast, a dynamics systems approach can allow us to assemble some fragments of theory into a single model, as we did in the example. A dynamic systems approach can give us a means to assemble knowledge fragments, and to explore the resulting systems. This can give us a way to incorporate findings from previous research into our theorising.

The above has described some of the benefits of taking a dynamic systems approach. Dynamic systems have specific advantages for phenomena where aggregation, dynamics or multiple dimensions are prominent features. These are common features in health psychology. Beyond these specific advantages, dynamic systems approaches have some more general advantages. They offer an alternative way to develop theory, that concentrates on the processes underlying phenomena. This focus on processes produces theories that are integrative. It can also lead us to ask different types of questions in other research, as it depends on information about processes to allow us to build dynamic system models.

As we have seen, there are a variety of ways to approach dynamic systems in practice. The example demonstrates only one of these, simulation. Simulation of dynamic systems itself has some advantages.

In the example, the process of developing the simulation model itself generated an unexpected result, that there may be a protective effect of a normal background of making and breaking relationships. This demonstrates one advantage of simulation modelling; the development of the model allows us to interact with the system and to learn from that interaction. As we begin to learn from this process we may find that we develop a more intuitive understanding of the particular system, and of dynamic systems more generally.

Simulation studies have some other practical advantages over experimental or demographic studies. They can be relatively inexpensive. The simulation work here was carried out on inexpensive equipment using inexpensive software. Modifications to the experimental design and size of the sample can be made at no cost and quickly applied, and we can approach the model with new questions as we think of them. Simulation gives us the means to investigate long term dynamics, and to access systems with a large number of components. It is a safe way to experiment, as the individuals studied are virtual rather than real, and so there are no human or animal ethics concerns.

Limitations of simulation and dynamic systems approaches

The above has presented some of the benefits and advantages of simulation and of dynamic systems approaches more generally. This is not to suggest that these methods are not without their own disadvantages.

One of the most important of these are concerns about how closely a simulation model represents the target. Simulation models may fail by producing patterns that do not exist in the target, or by failing to produce important patterns exhibited by the target (Doran & Gilbert, 1994). The very difficulties that lead us into simulation modelling can also make it difficult to make comparisons to fully test the model against the target system.

In the example I have attempted to ensure that the model represents the target by assembling the system model based on evidence from other studies. While this increases the chances that a model will be good, a model does not have to be assembled from accurate components. For an example of this we might consider the earlier example of the notation of drum music. The musical notation is not an accurate representation of what is played. The difference is obvious in the performance of a live drummer compared with a drum machine. But despite its inaccuracy, musical notation remains a perfectly good model of drum music.

The only way to judge the model is through the outputs that it produces, and whether these match what we can observe in the target system. Although we are unlikely to be able to completely test the model against the target system, it is quite likely that we can observe some aspects of the target. Where we can, we are able to judge how well the model has represented the behaviour of the target. This allows us to build up a broad impression of how useful the model is as we develop experience with it.

We also need to be cautious when using numerical modelling. Many simulation techniques use approximation methods to obtain solutions. These can be unstable, or worse can produce wrong answers. An unstable failure of a numerical routine is seen when statistical programmes stop processing after failing to find a solution after a certain number of iterations. These problems can exist, even where the processes are well-behaved mathematically and well understood. Some methods can produce answers to known accuracy, but other methods cannot.

Some other forms of simulation produce models that, while they have proved reliable in practice, leave a residual of doubt because we cannot understand their reasoning or determine their accuracy. Solutions found through neural network models suffer this problem. We do not know how they produce a result and it is difficult to place trust in a simulation where we do not understand what it is doing. There is an element of this in the non-linear behaviour of the system in the example. When we vary the size of the income step, we observe sudden steps in the response of the system. While we may understand these as the system hitting a trajectory that takes it to a different attractor, this gives us very little insight as to what is happening inside the system as we meet these transitions.

The above problems with simulation relate to the mechanics of simulation models and the components that go to make them up. An entirely different problem relates to the interface with the user. Models are only useful if they can help us to understand the target system, and a simulation model is unlikely to do this if it is difficult for users to relate the model back to the target system. Numerical models and computers do not have the feel of the social and psychological systems in a health context, and so do not necessarily make particularly user-friendly aids to understanding health psychology. This is particularly apparent in the example, in which the numerical nature of the final realisation is visible. In the case of the example, this has happened because I have exposed the details of the construction of the simulation model to illustrate the process. Careful interface design may ensure that numerical routines in a realisation would not be visible to the user, any more than the numerical nature of the data on a compact disc is visible to someone listening to music.

Part of our discomfort with numerical models and dynamic models is the close association of systems analysis tools with the physical and engineering sciences. It would be entirely reasonable to regard an attempt to realise psychological and social systems in physical and engineering terms with scepticism. This use of simulation modelling is not an attempt to do this. Rather it is an attempt to widen the techniques available in health psychology. While the techniques may have been developed first in the physical sciences, they are not exclusively associated with these disciplines. Their use simply offers us one way to think about some types of pattern that we might find.

The above all relate particularly to computer simulation methods of simulating dynamic systems. Dynamic systems approaches also have their own limitations. They will not be applicable to all situations. In many cases static models may be perfectly adequate. Where they are, they are likely to be the most suitable tools available. Even where a dynamic model can say something about a static problem, it will be much more awkward to use than a suitable static model.

More fundamentally, some writers have questioned whether formal dynamic systems concepts are useful at all in the social sciences (Puddifoot, 2000). Here I have used a very general definition of a system, and social systems clearly meet this definition. Whether or not dynamic systems approaches are useful depends on our experiences of using them, as the use of dynamic systems models is at an exploratory stage. In the case of the example, developing the model has proved interesting. It has led us to consider the processes operating in the network of relationships and the possibilities that the economic environment might disrupt these. Although the model might not produce clear-cut, definitive results, it does suggest some things that we might look more closely at.

Even if we accept that systems conceptions are valid, we do not know the nature of the systems that we are dealing with. This model assumes that social behaviour is deterministic, and that outcomes are determined by the individuals states and the environmental inputs. This is characteristic of dynamic systems approaches, as these are deterministic (Alligood et al., 1996). The alternative is that there is an element of chance, and the system is stochastic. Again we have little experience to point to one or other of these. That said, potential sources of entirely random noise in social systems are not readily apparent. People tend to believe that they have reasons for why they do things, and psychologists tend to act as though they believe this too.

The place of dynamic systems approaches

I have noted above the benefits and difficulties of using dynamic systems approaches. In claiming potential benefits, I am not saying that traditional theoretical and statistical methods are wrong. Existing methodologies give us a powerful tool set for understanding health psychology, and the use of these has been very important in maintaining scientific rigour. But if we are to be creative and versatile craftspeople we need to know the limitations of our tools, and to familiarise ourselves with a wide range of tools for a wide range of situations. Dynamic systems tools have the potential to extend the tool set used in health psychology allowing us to access systems problems and processes. In doing so, they may extend the reach afforded us by existing approaches.

References

Allen, R.E. (Ed.). (1990). *The concise Oxford dictionary of current English. (8th ed.).* Oxford: Clarendon Press.

Alligood, K.T., Sauer, T.D., & York, J.A. (1996). *Chaos - an introduction to dynamical systems.* New York: Springer.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286,* 509-512.

Ben-Shlomo, Y., White, I.R., & Marmot, M. (1996). Does the variation in the socioeconomic characteristics of an area affect mortality? *British Medical Journal, 312,* 1013-1014.

Berkman, L.F., Glass, T., Brissette, I., & Seeman, T.E. (2000). From social integration to health: Durkheim in the new millenium. *Social Science and Medicine, 51,* 843-857.

Berkman, L.F., & Syme, S.L. (1979). Social networks, host resistance and mortality: a nine-year follow-up study of Alameda County residents. *American Journal of Epidemiology, 109,* 186-204.

Brogan, W.L. (1974). *Modern control theory.* New York: Quantum Publishers.

Carroll, D., Davey Smith, G., & Bennett, P. (1996). Some observations in health and socio-economic status. *Journal of Health Psychology, 1,* 23-39.

Chattoe, E. (1996). Why are we simulating anyway? Some answers from economics. In K.G. Troitzsch, U. Mueller, G.N. Gilbert, & J.E. Doran (Eds.), *Social sciences microsimulation.* Berlin: Springer.

Coburn, D. (2000). Income inequality, social cohesion and the health status of populations: the role of neo-liberalism. *Social Science and Medicine, 51,* 135-146.

Cohen, S., & Wills, T.A. (1985). Stress, support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310-357.

Doran, J.E., & Gilbert, G.N. (1994). Simulating societies: an introduction. In G.N. Gilbert & J.E. Doran (Eds.), *Simulating societies: the computer simulation of social phenomena*. London: UCL Press.

Durkheim, E. (1951). *Suicide* (J.A. Spaulding & G. Simpson, Trans.). New York: Free Press. (Original work published 1897)

Elgerd, O.I. (1967). *Control systems theory*. New York: McGraw-Hill.

Evans, R.G. (1994). Introduction. In R.G. Evans, M.L. Barer, & T.R. Marmor, *Why are some people healthy and others not?*. New York: Aldine de Gruyter.

Fehr, B. (1996). *Friendship processes*. Thousand Oaks, California: Sage Publications.

Felton, B.J., & Shinn, M. (1992). Social integration and social support: moving "social support" beyond the individual level. *Journal of Community Psychology, 20,* 103-115.

Ferrie, J.E., Shipley, M.J., Marmot, M.G., Stansfeld, S., & Davey Smith, G. (1995). Health effects of anticipation of job change and non-employment: longitudinal data from the Whitehall II study. *British Medical Journal, 311,* 1264-1269

Franzoi, S.L. (1996). *Social psychology*. Madison: Brown & Benchmark.

Geronimus, A.T., & Bound, J. (1998). Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *American Journal of Epidemiology, 148,* 475-485.

Gleick, J. (1987). *Chaos*. London: Penguin.

Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology, 78,* 1360-1380.

Gujarati, D.N. (1995). *Basic econometrics (3rd edn.)*. New York: McGraw-Hill.

Hilborn, R., & Mangel, M. (1997). *The ecological detective: confronting models with data*. Princeton: Princeton University Press.

Holland, J.H. (1998). *Emergence: from chaos to order.* Oxford: Oxford University Press.

House, J.S., Landis, K.R., & Umberson, D. (1988). Social relationships and health. *Science, 241,* 540-545

Kaplan, G.A., Pamuk, E.R., Lynch, J.W., Cohen, R.D., & Balfour, J.L. (1996). Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *British Medical Journal, 312,* 999-1003.

Kawachi, I., & Kennedy, B.P. (1997a). Health and social cohesion: why care about income inequality. *British Medical Journal, 314,* 1037-1040.

Kawachi, I., & Kennedy, B.P. (1997b). The relationship of income inequality to mortality: does the choice of indicator matter. *Social Science and Medicine, 45,* 1121-1127.

Kawachi, I., Kennedy, B.P, Lochner, K., & Prothrow-Stith, D. (1997). Social capital, income inequality, and mortality. *American Journal of Public Health, 87,* 1491-1498.

Kennedy, B.P., Kawachi, I., & Prothrow-Stith, D. (1996). Income distribution and mortality: Cross sectional ecological study of the Robin Hood index in the United States. *British Medical Journal, 312,* 1004-1007.

Kennedy, B.P, Kawachi, I., Prothrow-Stith, D., Lochner, K., & Gupta, V. (1998). Social capital, income inequality, and firearm violent crime. *Social Science and Medicine, 47,* 7-17.

Lynch, J.W., Davey Smith, G., Kaplan, G.A., & House, J.S. (2000). Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *British Medical Journal, 320,* 1200-1204.

Lynch, J.W., & Kaplan, G.A. (1997). Understanding how inequality in the distribution of income affects health. *Journal of Health Psychology, 2,* 297-314.

Maton, K.I. (1989). Community settings as buffers of life stress? Highly supportive churches, mutual help groups, and senior centers. *American Journal of Community Psychology, 17,* 203-232.

McLaren, N. (1998). A critical review of the biopsychosocial model. *Australian and New Zealand Journal of Psychiatry, 32,* 86-92.

Morgan, D.L. (1990). Combining the strengths of social networks, social support, and personal relationships. In S. Duck & R.C. Silver (Eds.), *Personal relationships and social support.* London: Sage Publications.

Morris, J.K., Cook, D.G., & Shaper, A.G. (1994). Loss of employment and mortality. *British Medical Journal, 308,* 1135-1139.

Nowak, A., & Lewenstein, M (1994). Dynamical systems: A tool for social psychology. In R.R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology.* San Diego: Academic Press.

Pamuk, E.R. (1985). Social class inequality in mortality from 1921 to 1972 in England and Wales. *Population Studies, 39,* 17-31.

Pappas, G., Queen, S., Hadden, W., & Fisher, G. (1993). The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. *The New England Journal of Medicine, 329,* 103-109.

Phillips, D.C. (1992). *The social scientist's bestiary.* Oxford: Pergamon Press.

Phillimore, P., Beattie, A., & Townsend, P. (1994). Widening inequality of health in Northern England, 1981-1991. *British Medical Journal, 308,* 1125-1128.

Pierce, G.R., Sarason, B.R., & Sarason, I.G. (1990). Integrating social support perspectives: Working models, personal relationships, and situational factors. In S. Duck & R.C. Silver (Eds.), *Personal relationships and social support.* London: Sage Publications.

Podder, N., & Chatterjee, S. (1998). *Sharing the national cake in post reform New Zealand: Income inequality trends in terms of income sources.* Paper presented at the Annual Conference of the New Zealand Association of Economists, Wellington 1998.

Power, H.M., & Simpson, R.J. (1978). *Introduction to dynamics and control.* London: McGraw-Hill.

Puddifoot, J.E. (2000). Some problems and possibilities in the study of dynamical social processes. *Journal for the Theory of Social Behaviour, 30,* 79-97.

Putnam, R.D. (1995). Tuning in, tuning out: The strange disappearance of social capital in America. *PS: Political Science and Politics, 28,* 664-683.

Rodgers, G.B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies, 33,* 343-351.

Sampson, R.J. (1991). Linking the micro- and macrolevel dimensions of community social organization. *Social Forces, 70,* 43-64.

Sampson, R.J., Raudenbush, S.W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277,* 918-924.

Sarafino, E.P. (1998). *Health psychology: biopsychosocial interactions.* New York: Wiley.

Scruton, R. (1996). *Modern philosophy.* London: Mandarin Paperbacks.

Spicer, J., & Chamberlain, K. (1996). Developing psychosocial theory in health psychology: Problems and prospects. *Journal of Health Psychology, 1,* 161-171.

Sprecher, S., & Fehr, B. (1998). The dissolution of close relationships. In J. Harvey (Ed.), *Perspectives on loss: a sourcebook.* Philadelphia: Brunner/Mazel.

Soobader, M., & LeClere, F.B. (1999). Aggregation and the measurement of income inequality: effects on morbidity. *Social Science and Medicine, 48,* 733-744.

Statistics New Zealand. (1999). *New Zealand now: Incomes.* Wellington: Statistics New Zealand.

Susser, M. (1994). The logic in ecological: I. The logic of analysis. *American Journal of Public Health, 84,* 825-829.

Thoits, P.A. (1995). Stress, coping, and social support processes: Where are we? What next? *Journal of Health and Social Behaviour, Extra Issue,* 53-79.

Vallacher, R.R., & Nowak, A. (Eds.) (1994a). *Dynamical systems in social psychology.* San Diego: Academic Press.

Vallacher, R.R., & Nowak, A. (1994b). The chaos in social psychology. In R.R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology*. San Diego: Academic Press.

Walberg, P., McKee, M., Shkolnikov, V., Chenet, L., & Leon, D.A. (1998). Economic change, crime, and mortality crisis in Russia: Regional analysis. *British Medical Journal, 317,* 312-318.

Wallace, R., Fullilove, M.T., & Flisher, A.J. (1996). AIDS, violence and behavioral coding: Information theory, risk behavior and dynamic process on core-group sociogeographic networks. *Social Science & Medicine, 43,* 339-352.

Wallace, R., & Wallace, D. (1997). Socioeconomic determinants of health: Community marginalisation and the diffusion of disease and disorder in the United States. *British Medical Journal, 314,* 1341-1345.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge: Cambridge University Press.

Wilkinson, R.G. (1990). Income distribution and mortality: a 'natural' experiment. *Sociology of Health and Illness, 12,* 391-412.

Wilkinson, R.G. (1992). Income distribution and life expectancy. *British Medical Journal, 304,* 165-168.

Wilkinson, R.G. (1996). *Unhealthy societies.* London: Routledge.

Appendix - Matlab Listings

Linearset.m

```
%SIMULATE THE LINKS BETWEEN INDIVIDUALS OF A PRECALCULATED POPULATION

%The first set is calculated using linearset.m

%This allows me to vary the model as a test of linear behaviour

%Load the stable starting point from previously stored run

%This is the stable starting point.

load     c:\asue\thesis\LinearitySet\linearset     Relationship     Population
NoOfIntervals NoOfRelTypes Loading Income Position

IncomeDistance=zeros(Population);

PhysDistance=zeros(Population);

%Calculate the social distances from the Incomes and the physical distances

%from the position matrices

for n=1:(Population-1);

 for m=(n+1):Population;

 IncomeDistance(n,m)=round(abs(Income(n)-Income(m)));

 PhysDistance(n,m)=sqrt((Position(1,n)-Position(1,m))^2+(Position(2,n)-
Position(2,m))^2);

 IncomeDistance(m,n)=IncomeDistance(n,m);

 PhysDistance(m,n)=PhysDistance(n,m);

 end

end

NetworkSize=hist(Relationship,(NoOfRelTypes+1));

InitialNetworkSize=NetworkSize;

InitialLoading=sum(Loading);

InitialIncome=Income;

TotalLoading=InitialLoading;

%PERTURB THE DISTANCES BETWEEN PEOPLE BY ADJUSTIING THE INCOMES OF A GROUP

%Increase the top 10% "income bracket" Income by 100%

for n=1:Population;

 if Income(n)>=14

 Income(n)=2*Income(n);

 else

 end

end
```

```
%Calculate the new income distances and tolerable loadings immediately
%following the social distance perturbation
for n=1:(Population-1);
 for m=(n+1):Population;
 IncomeDistance(n,m)=round(abs(Income(n)-Income(m)));
 IncomeDistance(m,n)=IncomeDistance(n,m);
 end
end
MaxTolerable=(60*Relationship./((IncomeDistance+1).*(PhysDistance+1)));
for m=1:(Population-1)
 for n=(m+1):Population
 if Loading(m,n)>MaxTolerable(m,n)
 Relationship(m,n)=0;
 Relationship(n,m)=0;
 Loading(m,n)=0;
 Loading(n,m)=0;
 else
 end
 end
end
NetworkSize=hist(Relationship,(NoOfRelTypes+1));
NewLoading=sum(Loading);
TotalLoading=[TotalLoading;NewLoading];
%For NoOfIntervals points adjust all of the loadings by scaling them up by the
initial loading
%divided by the new loading.
x=[0 1];
Count=2;
while Count<=NoOfIntervals
 for m=1:(Population-1)
 if NewLoading(m)~=0
 Scale=InitialLoading(m)/NewLoading(m);
 for n=(m+1):Population
 Loading(m,n)=Loading(m,n)*Scale;
 Loading(n,m)=Loading(m,n);
 if Loading(m,n)>MaxTolerable(m,n);
 Relationship(m,n)=0;
```

```
 Relationship(n,m)=0;
 Loading(m,n)=0;
 Loading(n,m)=0;
 else
 end
 end
 else
 end
 end
 NetworkSize=hist(Relationship,(NoOfRelTypes+1));
 NewLoading=sum(Loading);
 TotalLoading=[TotalLoading;NewLoading];
 x=[x Count];
 Count=Count+1;
end
%Now collect up the results into a two worksheet which I can take straight
into SPSS.
%One will have the loading history.
History=TotalLoading';
wk1write('c:\asue\thesis\LinearitySet\History.wk1',History)
%The other has columns representing
%Income
%Position
%the total change in loadings/initial loading
ChangeLoading=(TotalLoading(1,:)-
TotalLoading(NoOfIntervals,:))./TotalLoading(1,:);
%the total change in each type of relationship / initial number in the
relationship
ChangeNetworkSize=(InitialNetworkSize-NetworkSize)./InitialNetworkSize;
Outfile=[InitialIncome;Position;ChangeLoading;ChangeNetworkSize];
Outfile=Outfile';
wk1write('c:\asue\thesis\LinearitySet\Outfile.wk1',Outfile)
```

NoiseIncr.m

```
%SIMULATE THE LINKS BETWEEN INDIVIDUALS OF A PRECALCULATED POPULATION

%The first set is calculated using linearset.m

%This allows me to vary the model as a test of linear behaviour

%Load the stable starting point from previously stored run

%This is the stable starting point.

load    c:\asue\thesis\LinearitySet\linearset    Relationship    Population
NoOfIntervals NoOfRelTypes Loading Income Position

IncomeDistance=zeros(Population);

PhysDistance=zeros(Population);

%Calculate the social distances from the Incomes and the physical distances

%from the position matrices

IncomeDistance=IncDist(Population,Income);

PhysDistance=PhysDist(Population,Position);

NetworkSize=hist(Relationship,(NoOfRelTypes+1));

InitialNetworkSize=NetworkSize;

InitialLoading=sum(Loading);

InitialIncome=Income;

TotalLoading=InitialLoading;

%PERTURB THE DISTANCES BETWEEN PEOPLE BY ADJUSTING THE INCOMES OF A GROUP

%Increase the top 10% "income bracket" Income by 100%

for n=1:Population;

 if Income(n)>=14

 Income(n)=2*Income(n);

 else

 end

end

%Calculate the new income distances and tolerable loadings immediately

%following the social distance perturbation

IncomeDistance=IncDist(Population,Income);

MaxTolerable=(60*Relationship./((IncomeDistance+1).*(PhysDistance+1)));

%Test whether each relationship has a tolerable loading for closeness and
type.

%And break any which have an intolerable loading.

[Relationship,Loading]=TestandBreak(MaxTolerable,Loading,Relationship,Populati
on);

NetworkSize=hist(Relationship,(NoOfRelTypes+1));
```

```
NewLoading=sum(Loading);

TotalLoading=[TotalLoading;NewLoading];

%For NoOfIntervals points adjust all of the loadings by scaling them up by the
initial loading

%divided by the new loading.

x=[0 1];

Count=2;

while Count<=NoOfIntervals

 Count

 [Relationship,Loading]=BackgroundNoise(Loading,Relationship,Population);

 for m=1:(Population-1)

 if NewLoading(m)~=0

 Scale=InitialLoading(m)/NewLoading(m);

 for n=(m+1):Population

 Loading(m,n)=Loading(m,n)*Scale;

 Loading(n,m)=Loading(m,n);

 if Loading(m,n)>MaxTolerable(m,n);

 Relationship(m,n)=0;

 Relationship(n,m)=0;

 Loading(m,n)=0;

 Loading(n,m)=0;

 else

 end

 end

 else

 end

 end

[Relationship,Loading]=TestandBreak(MaxTolerable,Loading,Relationship,Populati
on);

 NetworkSize=hist(Relationship,(NoOfRelTypes+1));

 NewLoading=sum(Loading);

 TotalLoading=[TotalLoading;NewLoading];

 x=[x Count];

 Count=Count+1;

end

%Now collect up the results into a two worksheet which I can take straight
into SPSS.
```

```
%One will have the loading history.

History=TotalLoading';

wk1write('c:\asue\thesis\LinearitySet\HistoryLinearSteady.wk1',History)

%The other has columns representing

%Income

%Position

%the total change in loadings/initial loading

ChangeLoading=(TotalLoading(1,:)-
TotalLoading(NoOfIntervals,:))./TotalLoading(1,:);

%the total change in each type of relationship / initial number in the
relationship

ChangeNetworkSize=(InitialNetworkSize-NetworkSize)./InitialNetworkSize;

Outfile=[InitialIncome;Position;ChangeLoading;ChangeNetworkSize];

Outfile=Outfile';

wk1write('c:\asue\thesis\LinearitySet\OutfileLinearSteady.wk1',Outfile)
```

IncDist.m

```
function IncomeDistance=IncDist(Population,Income)

for n=1:(Population-1);

 for m=(n+1):Population;

 IncomeDistance(n,m)=round(abs(Income(n)-Income(m)));

 IncomeDistance(m,n)=IncomeDistance(n,m);

 end

end
```

PhysDist.m

```
function PhysDistance=PhysDist(Population,Position)

for n=1:(Population-1);

 for m=(n+1):Population;

 PhysDistance(n,m)=sqrt((Position(1,n)-Position(1,m))^2+(Position(2,n)-
Position(2,m))^2);

 PhysDistance(m,n)=PhysDistance(n,m);

 end

end
```

Testand Break.m

```
function
[Relationship,Loading]=TestandBreak(MaxTolerable,Loading,Relationship,Populati
on)

for m=1:(Population-1)

 for n=(m+1):Population

 if Loading(m,n)>MaxTolerable(m,n)

 Relationship(m,n)=0;

 Relationship(n,m)=0;

 Loading(m,n)=0;

 Loading(n,m)=0;

 else

 end

 end

end
```

BackgroundNoise.m

```
function
[Relationship,Loading]=BackgroundNoise(Loading,Relationship,Population)

%This subroutine creates a background noise of making and breaking
relationships

%This would not be necessary if the system was linear. As it is not,

%the principle of superposition does not apply. This assumes a exponential
relationship

%between the likelihood of a relationship breaking, or being made and the
strength of the

%relationship. That is for strength 1, the chance is 1/Base, for strength 2,
the chance is 1/Base^2

%and so on.

Base=100;

for n=1:(Population-1)

 for m=(n+1):Population

 if Relationship(n,m)==0

 Chance=(Base^5)*rand;

 for j=5:-1:1

 if Chance<(Base^j)

 Relationship(n,m)=j;

 Relationship(m,n)=j;
```

```
Loading(n,m)=19*rand+1;
Loading(m,n)=Loading(n,m);
else
end
end
else
Chance=Base^Relationship(n,m)*rand;
if Chance<1;
Relationship(n,m)=0;
Relationship(m,n)=0;
Loading(n,m)=0;
Loading(m,n)=0;
else
end
end
end
end
```