



Lancaster University
MANAGEMENT SCHOOL

Lancaster University Management School
Working Paper
2003/058

**A Taxing Problem: The Complementary Use Of Hard And
Soft Or In Public Policy**

Mike Pidd, Joyce Brown and Ceri Cooper

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX
UK

©Mike Pidd, Joyce Brown and Ceri Cooper
All rights reserved. Short sections of text, not to exceed
two paragraphs, may be quoted without explicit permission,
provided that full acknowledgement is given.

The LUMS Working Papers series can be accessed at <http://www.lums.co.uk/publications>
LUMS home page: <http://www.lums.lancs.ac.uk/>

A Taxing Problem: The Complementary Use Of Hard And Soft Or In Public Policy

Joyce Brown¹, Ceri Cooper¹ and Michael Pidd²

¹Inland Revenue, London

²Lancaster University, Lancaster

ABSTRACT

A review of the UK personal taxation system used a combination of hard and soft OR approaches in a complementary way. The hard OR was based on data mining to increase understanding of individual taxpayers and their changing needs within the personal tax system. The soft OR was based on soft systems methodology with two aims in mind. First, to guide the review and, secondly, as an auditable approach for collecting the views of key internal and external stakeholders. The soft and hard OR were used alongside one another, rather than one providing a contextual scheme for the other. The experience reveals that soft OR is much more than common sense and, that, used in parallel, soft and hard approaches have a powerful synergy.

INTRODUCTION

During the last 20 years, an approach to OR, often known as soft OR, has been recognised as legitimate. It is often contrasted with hard OR, though both terms are unfortunate, since the terms soft and hard have other connotations. The approaches are often discussed as if they represent extremes, but the basic argument presented here is that they are useful complements to one another.

The question, though, is how can they best be combined? Mingers & Gill, (1997) provide a wide range of views of the ways in which different fundamental methodologies can be combined. Jackson and Keys (1984) analyse different approaches to OR and place them in a typology based on that of Burrell and Morgan (1979). Most agree that before different methods and methodologies are combined, it is important to think through how this will be done, and that this is best achieved from an understanding of the philosophy and assumptions underpinning different approaches. Though it is difficult to argue against this from a purist viewpoint,

anecdotes suggest that practitioners, whatever their philosophical position, often do combine hard and soft approaches in their work and that this combination may be rather pragmatic. This paper describes a large-scale public policy study in which both soft and hard methods were consciously used and discusses the power of this combined approach. Though presented through a particular case study, the aim is also to show that this complementarity makes intellectual as well as pragmatic sense.

The taxing problem

The Inland Revenue is a major department of the UK Civil Service. Its original remit was to collect the taxes needed by the Government to finance public services. Until recently, its major task was the collection of income taxes from individuals and corporate taxes from businesses. Its remit is now broader and includes a number of other responsibilities such as the payment of tax credits (for example, to working families who are low paid) and the management of the system for collecting student loan repayments. Formerly a department with the more or less single pre-occupation of collecting taxes, it has become an agent of social change with tasks that include helping people to move into and to remain in work.

The Inland Revenue keeps its operations under continuous review trying to find more efficient and effective ways of meeting the goals set for it by Government. The study described here allowed some blue-sky thinking and analysis, set within the context of continuous review. The terms of reference for the overall study were set by the Board of the Inland Revenue and can be summarised as follows.

A study, in consultation with stakeholders, of the scope for modernising the operation of the UK's personal tax system. The study should take account of the current situation, developments elsewhere in the world, the possibilities for simplification and the opportunities provided by new technologies. It should take account of the need for a more 'joined-up' approach to public policy and service provision.

The study team was a small group of tax policy experts that was able to draw on other resources it felt necessary. It chose to draw on the Inland Revenue's own Operational Research resources and those of Lancaster University. This led to parallel work in both soft and hard OR. The hard OR was, mainly, based on data mining to establish profiles of customer groups and was conducted by in-house OR staff. The

soft OR was based on Checkland's soft systems methodology (SSM), which was used to provide a structured and rigorous approach to the study and also as a way of gathering stakeholder views. The way in which the SSM provided an auditable structure for the study is discussed elsewhere (Brown, Checkland & Cooper, 2003).

WHAT IS MEANT BY HARD AND SOFT OR?

This is a question often discussed at conferences in which soft OR work is presented and it is likely that, if four OR people were asked this question, they would provide at least five answers. Rosenhead (1989) and Rosenhead and Mingers (2001) provide discussions of this question, as does Pidd (1996, 2003). Rather than repeat these classifications here, a rather broader approach will be taken instead, based on workshops run with practitioners and academics present.

The INCISM network

Around the start of the millennium, the UK Engineering and Physical Sciences Research Council (EPSRC) established a number of interdisciplinary networks to enable researchers to meet and discuss ideas related to large complex systems. The initiative came, primarily, from the UK computing community, which was concerned at the increased scale, scope and complexity of IT systems. One of these networks, INCISM (the Interdisciplinary Network on Complementarity in Systems Modelling) was established to examine the ways in which hard and soft methods might be combined. Details of INCISM, which operated in 2001/2, can be found on its website (INCISM, 2002).

The INCISM network was based on two dimensions of complementarity, as shown in figure 1. Its initial composition was six partners; these being three university departments (Lancaster, Strathclyde and The Royal Military College/Cranfield University) and three practitioner groups (BT Exact, Shell International and Dstl) - all based in the UK. Thus, in figure 1, the first dimension of complementarity is between the two worlds of the academy and practice. As shown on the figure, these are not distinct worlds, but are rather points on a spectrum, since many academics engage in consulting work and some practitioners act as visiting and adjunct staff in universities. The second dimension of figure 1 is that of hard and soft and, again, these are shown as points on a spectrum rather than as wholly distinct.

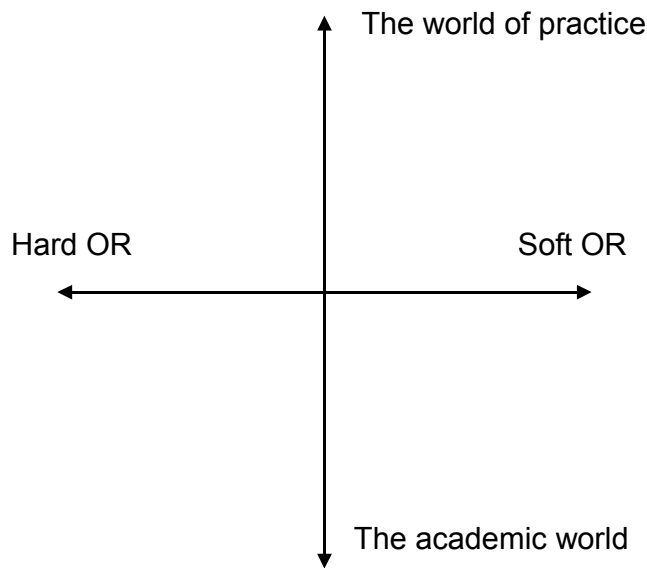


Figure 1: Complementarity in INCISM

Thus, the INCISM network was established to see what complementary insights would emerge from a mixed group of academics and practitioners who met regularly to discuss aspects of hard and soft OR. The second INCISM meeting was devoted to a discussion of the differences between hard and soft approaches. It became clear that there are many different ways, even in OR, in which the terms are used and it is important to understand these.

Looking behind hard and soft OR

One INCISM participant has consistently argued (Checkland, 1981) that the terms soft and hard are based on important philosophical differences. Checkland (op cit) draws a distinction between approaches based on different ontologies (assumptions and theories about the world) and the consequential epistemologies (theories of knowledge about that world). In Checkland's terms, hard OR stems from a positivist ontology that leads to a functionalist epistemology. In broad terms, someone adopting a positivist ontology assumes that the world can be known through experimental interaction with "the world out there". This is a very powerful view that underpins much physical science and economics. For someone adopting a positivist stance, a model is likely to be seen as a would-be representation of the real, external world. Experiments on and manipulations of that model are expected to yield useful insights for the real world if the model is valid. Thus model validity is a major concern if this

positivist stance is adopted, since the model is regarded as a would-be representation of the real, external world (Pidd, 1996).

Similarly, someone adopting a wholly positivist stance might regard the world as really consisting of entities that interact and form systems that can be observed and manipulated. This can lead to a functionalist epistemology in which the entities must be examined in terms of their function within such a system. This yields very useful insights, but can lead to a very conservative view in which the only legitimate role for the entities is the part they play in achieving the objectives of the system they are assumed to compose. Thus, in these terms, hard OR is an essentially analytical approach that allows options to be compared within some defined set of objectives. It offers powerful ways to investigate how a goal should be achieved, but is limited when considering what that goal might be and why it is important. A positivist ontology combined with a functionalist epistemology is extremely powerful and brings many insights, yet there are parts that this combination cannot reach.

By contrast, Checkland (op cit) argues that soft OR stems from a phenomenological ontology that leads to an interpretive or constructivist epistemology. As previously, these terms need to be unpacked. A phenomenological stance is one in which the observer recognises her own position and the fact that she uses her mental processes when experiencing the external world. Thus, the ways in which she interacts with the external world and the mental process that she uses as she does so, are important. This does not imply a position in which there is held to be no real or external world whatsoever, merely a view that different people may, legitimately, see it differently and may add to it. Phenomenology has many definitions, but for present purposes, it is the study of the way that things appear to us as humans. In these terms, a model represents the way that an individual or group sees something, which leaves the question of model validation suspended for the time being.

An interpretive epistemology (sometimes known as social constructivism) is one that regards individuals as learning when they try to make sense of their experiences and assumes that this is done through a process of mental construction. This construction is based on the learner's context, beliefs and prior knowledge. As we interact and act with other people we socially construct our everyday world, or at least aspects of it. In these terms, the role of OR, in particular of soft OR, is to help

people to make sense of their worlds and to construct helpful accounts of it. A model, in these terms, can be a useful device to support debate between people as they learn how best to act in a particular situation. This view of soft OR allows the possibility of rational analysis that supports asking what should be done and why, rather than being limited to advise how something should be done. This is achieved by providing procedurally rational (Simon, 1976) decision support to participants as they explore the what and the why questions.

Pragmatic aspects of hard and soft OR.

As well as this philosophical view of the terms hard and soft OR, the INCISM meeting also explored some practical and pragmatic implications of the terms as they are used in everyday OR practice. Table 1 captures some of these practical and pragmatic aspects, which represent the ways in which active theorists and practitioners view the differences. Some of these aspects clearly overlap with one another, as will become clear.

	Hard OR	Soft OR
Methodology used	Based on common sense, taken for granted views of analysis and intervention	Based on rigorous epistemology
Models	Shared representation of the real world	Representation of concepts relevant to the real world.
Validity	Repeatable and comparable with the real world in some sense	Defensibly coherent, logically consistent, plausible
Data	From a source that is defensibly there in the world, with an agreed or shared meaning, observer independent	Based on judgement, opinion, some ambiguity, observer dependent
Value & outcome of the study	Quantification assumed to be possible and desirable. From option comparison based on rational choice.	Agreement (on action?), shared perceptions. Informing action and learning.
Purpose of the study	For the study: taken as a given at the start For the model: understanding or changing the world, linked to the purpose.	For the study: remains problematical For the model: a means to support learning

Table 1: Practical aspects of hard and soft OR

The first aspect shown in table 1 refers to the role of methodology in OR. As used here, methodology is taken as the study of how to go about an investigation and has two aspects. First a set of principles that underpin what we do and, secondly, the methods and procedures that we choose to use – based on those methodological principles. As can be seen from table 1, methodology is typically based on taken for granted views of analysis and rationality in hard OR. By contrast, in soft OR, methodology needs to be based on careful consideration and reflection – on a thought-through epistemology. Thus, there is a much greater stress on self-awareness in soft OR, for the consultant needs to think very carefully about her role so as to be aware of what she is doing in the particular social context of the study. It is usually assumed, in hard OR, that there is no real need to justify the methods and approaches in use, since they are taken to rely on unproblematic assumptions about external reality based on objective rationality. By contrast, there is a danger that soft OR could drift off into sloppy and purely relativistic thinking, were it not to be grounded in a careful consideration of methodology. It is precisely this rigorous concern for methodology that made soft systems methodology attractive for the tax policy study discussed later. The Inland Revenue study team were determined to use an approach that could be audited and that was defensible.

The second aspect shown in table 1 refers to the role of models in both hard and soft OR. This has already been touched on earlier, but bears repetition here. Underlying truly hard OR is a view that a model is a would-be representation of some aspect of the real world. There is no need to assume that the model is complete or fully detailed, for many writers argue that simplification is inevitable in modelling and some argue that it is desirable (e.g. Powell (1995) and Willemain (1994)). Modelling, in these terms, is an activity in which technical methods and insight are used to develop an external representation that is intended to provide useful insights into that which is being modelled. The data mining described in the tax policy study led to models that were held to be representative of UK taxpayer groups and their interaction with the current tax system. By contrast, in soft OR, a model is taken to be a representation of concepts relevant to understanding and working in the real world. Modelling in soft OR is a process of learning and shaping leading to an interpretation of how things might be. Thus, in the tax study, the models developed were abstract

representations of the features held to be desirable and necessary in a future tax system.

The third aspect of table 1 refers to model validity, a topic briefly introduced earlier. If a model is intended, as in hard OR, as a would-be representation of the real world, then it must be possible to compare it, in some way or other, with that real world. Without such a comparison, which can amount to a Turing test, what faith can there be that the model is valid and can be trusted? Of course, even in hard OR, this argument is on very shaky ground if the models are of systems as they might be, not as they are. In such cases, there is no referent system against which the model is to be compared. In the tax study, the models resulting from the data mining were held to have face validity by the tax policy experts, being in accord with their experience. Thus validation is sometimes problematic in hard OR and in some communities, notably in defence simulation, the idea of model credibility is used instead (Balci, 1987). Taking this idea further, in soft OR it is better to ask whether a model is defensibly coherent, logically consistent and plausible. For example, in SSM, the soft approach used in the tax study, conceptual models are usually expected to comply with known theory about the behaviour of physical systems. Thus, they must be self-maintaining through control mechanisms and their performance must be measurable, conceptually at least. This was how the credibility of the conceptual models developed in the tax study was assessed, in addition to their plausibility or face validity.

The fourth aspect of table 1 is the role of data in the work being done. A positivistic view of data, underpinning much hard OR, is that data come from a source that is defensibly there in the world (it is not just arbitrary), that they have an agreed or shared meaning (possibly based on known theory – e.g. in statistical method) and are independent of observer bias. Such assumptions need not be limited to purely quantitative data, but could also apply to qualitative data, for example the rules to be applied when collecting taxes. By contrast, things are not so simple in soft OR. Data are regarded as based on judgement and opinion; they may contain some ambiguity and are definitely observer dependent. It important to realise that even quantitative data can be ambiguous and observer dependent. For example, the number employed for the cost of capital in investment decisions or in stock control is quantitative, but is based on judgement and opinion. It is also highly ambiguous and may well be

observer dependent. In the tax study, the data used in the hard OR came from the Inland Revenue's records of UK taxpayers; the data for the soft study were collected in workshops and interviews with stakeholders.

What of the value and outcome of the study or intervention, shown as the fifth element of table 1? It is usually the case that hard OR aims to produce a tangible product in the form of recommendations, system design or change in the everyday real world. Many OR consultants sell their services on just this basis and some charge for their time as a percentage of audited savings that result from their work. This is, then a very appealing view that can easily be justified, or not, by a comparison of costs and benefits. In soft OR, things are not so simple, since the stress is on helping people to agree in situations where there may be disagreement and conflict about objectives as well as about what should be done. It may be that, once this agreement has been reached, it is possible to engage in some hard OR to decide exactly what should be done. It is fair to say, though, that even very hard OR projects may result in learning and may be used as a device to help people think through their objectives. In the tax study, the outcome of the hard OR was a set of models that represented archetypical taxpayer groups and the ways in which they interact with the tax system. The soft OR resulted in agreed recommendations of the ways in which the system might be changed.

Finally, table 1 shows that the intended purpose of soft and hard OR studies may differ. Perhaps this should have been discussed before the other aspects, but it is simpler to understand at this stage. In a hard OR study, the terms of reference for the study are agreed as quickly as possible at the start of the work and the aim is to meet those terms of reference. This assumes that the people drawing up those terms are clear about what needs to be done and why it needs to be done, but they wish to find the best way to do it. Similarly, the purpose of the modelling is to achieve as good a fit as is possible between the real world and the model, to enable the model to be used as a vehicle to see what would happen in the real world if particular actions were taken. In soft OR, things are very different, for the model is used as a vehicle to support the learning of the participants in the study. Further, the purpose of the study itself is something that is open to question throughout the engagement. It should be clear, though, that even in a hard OR study the terms of reference may be renegotiated while a soft OR study needs to reach agreement on its aims.

THE HARD OR IN THE TAX STUDY

The study team wanted a better understanding of the types of customers served by the Revenue, to better appreciate their needs of the personal tax system. To do this, they used OR techniques to examine the types of individual served by the personal tax system and the heterogeneity in this customer base. Underlying this was a basic question. What would be required to meet customer needs, and what system(s) will be required if their needs were diverse rather than homogeneous? This issue had become even more pressing with the changing role of the Revenue.

The study team took as their prime focus the impact of the personal tax system on people in employment. In the UK employees pay tax on their earnings under a system known as Pay As You Earn (PAYE). In this system employers deduct tax each week or month from wages and salaries using code numbers (reflecting the individual circumstances of each employee) and tax tables, both of which are provided by the Inland Revenue. The employer then forwards the tax deducted to the Inland Revenue. Employees who are taxed through PAYE usually have very little direct contact with the Inland Revenue, since their employers do most of the work of calculating and deducting the tax on their earnings. Only those employees who have higher earnings, additional income from other sources or particularly complex affairs have to fill in a tax return at the end of the year.

Data mining

The main "hard" OR approach used was data mining. This included basic data analysis (population counts, means, etc), web/link analysis to understand linkages between events (e.g. the types of PAYE codes issued to individuals through time as their circumstances changed) and cluster analysis based on Kohonen self-organising maps (Kohonen, 1990). The aim of the cluster analysis was to place individuals, as far as is possible, into homogenous groupings based on input variables chosen by the analyst. Used in this way, the data mining formed part of the problem structuring for the study and provided an evidence-base that identified natural segments of the personal tax customer and their needs from the system. Thus, it is not just soft OR methods that can be used for problem structuring, some hard techniques can also be

drawn upon for this. Pidd (1977, 1996) argued that, problem structuring often involves more than soft methods and may include preliminary data analysis.

The data mining used a random sample of just over 52,000 taxpayers based on the 1999/2000 tax year. For each of these taxpayers, data were extracted near to the end of December 2000, including age, employment history, sources of income and indications of their interactions with the tax system, such as the PAYE codes issued and repayments made in the previous year. The idea was to gain a snapshot of the taxpayer characteristics and circumstances at that time. Extensive data validation and manipulation was carried out prior to clustering since, as is usually the case (Pidd, 2002), the available data were incomplete and needed some manipulation and analysis prior to the cluster analysis. This showed that some of the data were flawed and, where possible, these were replaced by proxies. For example, age was not always present in the dataset, but a person's National Insurance Number usually contains information that allows their age to be estimated.

Basic analysis – population characteristics

The first aim was some overall understanding of the sample and an appreciation of how well the current system was working for different types of taxpayer. As an example, one driver for the study was a concern that the current system might not be meeting the needs of a changing workforce. The current tax system assumes that taxpayers are relatively homogeneous and enjoy stable employment, which may cause problems if a high proportion of people are switching employment during a tax year or because of an increase in portfolio workers. However, the data mining showed the situation to be less extreme than this. For example, only about 5% of employees (excluding directors who accounted for 3% of the sample) had more than one job during the year, and as many as 40% had not changed jobs/main source of income within the last 5 years.

Cluster analysis

The cluster analysis was done with the Acustar software (Package developed in-house by EDS) and employed Kohonen self-organising maps (op cit), an approach based on neural clustering. The software was chosen because of the quality of output produced which is particularly good when explaining results to business experts. The approach

is iterative, starting with initial clusters set up by the software which allocates a vector to each cluster based on random values for each input variable. The analyst specifies the number of clusters, N , into which the input data set is to be organised, the number of training cycles to be used (individuals being added to a cluster during each iteration based on the minimum distance between the individual and the cluster vector), and the extent to which clusters and neighbouring clusters are updated during the training. The extent to which clusters are updated also varies during training as specified by analyst. As clusters are updated, the cluster vector changes to reflect the individuals allocated to that cluster. In turn, this may cause the cluster to attract different individuals during later training cycles. Towards the end of the training cycles the map hopefully converges in that there is very little updating of the vector as individuals are generally attracted towards the same cluster, and the cluster vector converges towards the vector means for the individuals within it.

Thus, clusters form as the training progresses, leading to a cluster map, with similar clusters placed contiguously. The map is actually a torus.

It is not sensible to try to develop a map from a single attempt. Hence, the process of clustering was gradual and, at each attempt, important parameters such as the size of map, training cycle, inclusion of particular variables were amended as felt appropriate. In most cases, a 5 x 5 cluster map was used although on occasion this was thought too large, particularly when clustering on more homogeneous sub-groupings of the population (e.g. Company Directors), when a 3 x 3 map was used.

This clustering was part of an attempt to understand how heterogeneous was the customer base and, in turn, a wish to understand how well Inland Revenue processes mapped on to the clusters. Hence, once clusters were formed they were examined in the light of how individuals in the cluster interacted with the Revenue (e.g. Number of changes to PAYE codes issued per annum, etc). This showed how well or how badly the processes of the PTS worked for the various customer groups.

The data mining identified several segments with very simple affairs for whom the system works very well. However, other groups were not so well served. For example, transient workers, usually young people, whose affairs tended to be more complex and for whom PAYE may not operate very accurately. As expected, the clustering identified other groups with complex affairs for whom, again, PAYE

was not so well suited. The clustering, and the attempt to understand the clusters in the light of Inland Revenue processes, provided a useful basis for steering the SSM work in systems design.

The results of the data mining were presented at workshops with Inland Revenue business experts and statisticians to obtain their interpretation of the clusters. This was valuable in helping to interpret clusters and to ensure that the analysis was meaningful.

Conclusions from the data mining

The main conclusion from this data mining is that the vast majority of taxpayers have very simple tax affairs and the system appears to work very well for them. But the system works less well for some segments and, as is discussed later, the SSM part of the project sought to identify ways in which the needs of these customers could better be met. Another valuable insight from the cluster analysis was the extent to which traditional segments (e.g. Company Directors) were actually heterogeneous groups containing people with very diverse characteristics and needs. Indeed some Directors had much more in common with employees with very simple affairs than with other Directors who have much more complex affairs.

Perhaps the major benefit of the data mining was that it provided solid, quantifiable evidence of the current operation of the personal tax system in the UK. It moved the debate away from subjective statements of interest and demonstrated that, for many UK taxpayers, the system works well – though this does not mean that these people like paying income tax! It provided objective evidence of types of taxpayer, their needs of the tax system and the ways in which they interact with it.

THE SOFT OR

The soft OR approach in this study was based on Checkland's Soft Systems Methodology (SSM) (Checkland, 1996, 2000), which was used for two purposes. The first was as an overarching methodological guide; an issue discussed in detail in Brown, Checkland and Cooper (2003). Secondly it was used to gain understanding of how different stakeholders regard the personal tax system in the UK; an aspect discussed here in more detail. The SSM work was carried out by Inland Revenue staff, academic staff from Lancaster University (including Checkland) and external

consultants well versed in SSM. The SSM was the basis for a series of consultations with internal and external stakeholders who might be affected by changes to the personal tax system. They were invited to comment on their opinion of its current operations and also to suggest how it might be changed.

Workshops

SSM formed the basis for a series of workshops and interviews with a range of stakeholders. Each workshop aimed to elicit stakeholders' requirements and ideas by drawing comparisons between how they viewed the current tax system and what they would like to see in their ideal tax system. The idea, as shown in figure 2, was that this comparison would generate ideas for change. As with many problem solving approaches, this stage of the work allowed the divergent elicitation of many possible options for the future. However, it is important to follow this with a convergent phase in which fewer main themes are examined in detail. SSM was used to support both the divergent and convergent phases.

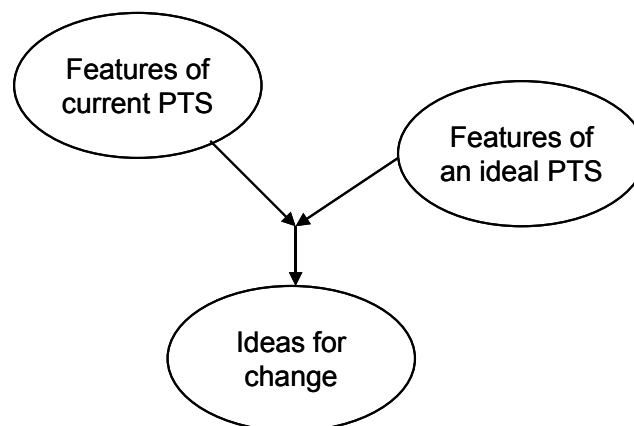


Figure 2: Generating ideas for change using SSM

Three broad groups of stakeholders were consulted through a series of workshops and interviews.

- External customer workshops were held with people such as employees, tax credit recipients, pensioners, employers and accountants.
- Internal (Inland Revenue) stakeholder workshops were held with groups such as operational staff, IT experts and tax policy experts.
- Individual interviews were held with senior members of the Inland Revenue and representatives from other Government Departments.

With 14 workshops in prospect, which were to be conducted by a range of people including members of the Inland Revenue team, Lancaster University staff and external consultants, a consistent approach and format was needed.

A typical workshop began with an introduction to the study, its background and its aims and objectives. Once participants were introduced to the study, the workshop focused on eliciting their requirements and ideas. As mentioned earlier and captured in figure 2, this was based on a comparison of participants' views of the current system and how they would like it to be.

Such a comparison can be actually rather difficult in practice. Particularly as workshops covered a broad range of stakeholder groups, each with varying degrees of knowledge and experience of the personal tax system. Hence, each workshop used a core metaphor – the tax system as a car. Participants were asked what type of car, for them, represented the tax system and its features. If they suggested an old, barely reliable banger, this suggested that their experience of the tax system was less than positive. This simple device ensured that the discussions in the workshops remained at a strategic level and kept away from detailed issues. Once settled into this mode of thinking about the current tax system, participants discussed the features they would like to see in an ideal personal tax system. By drawing comparison between the two, participants were encouraged to come up with ideas for change that would deliver their ideal system.

These ideas were expanded, using pro forma, in everyday language, designed to elicit the components of a root definition in SSM. Following each workshop, the ideas generated were converted into root definitions via PQRs, CATWOEs and 3Es, providing a clear definition and an appreciation of the nature and scope of the ideas generated by participants. The understanding gained through this proved a useful precursor to the following stages of the SSM study. Using this approach proved to be very successful in encouraging participants to contribute their views and ideas. The workshops generated over 90 distinct ideas for changing the current system: some were aimed at improvements within the current system, whilst several proposed fundamental changes to the way the personal tax system is operated.

Clustering and reducing the range of ideas

The workshops produced far too many ideas to analyse in the time available and there was considerable overlap. Thus, it made sense to cluster and prioritise them. This is a common difficulty in any problem solving that includes both divergent and convergent phases. Ideas are generated in the divergent phase, but these must be evaluated and some will be selected, which requires a convergent phase. This divergent to convergent shift was accomplished via a three-stage process. This included clustering together ideas that were extensions of, or variations on, other ideas. For example, a number of workshops raised ideas around improving the linkages between Inland Revenue computer systems, allowing easier access to taxpayers' records and a more complete picture of taxpayers' affairs. These ideas were grouped together to form a single cluster.

The clustered ideas were categorised as to their feasibility, excluding for example those that were already planned or too ill defined, and evaluated as to the likely positive impact of ideas on the tax system. Clusters with a minor impact were excluded in favour of those expected to have a high impact, leaving ten major themes to take forwards.

Working with the major themes

A core concept of SSM is its use of root definitions, to capture the essence of an idealised system. The analysis team developed root definitions for each of these themes. A root definition in SSM has six elements captured in the CATWOE mnemonic (Checkland, 1981, 2002).

Since it is not always straightforward to go from a theme to a root definition, an intermediate step was employed in this study. This is often known as PQR, based on the following formulation for a proposal: that it would 'do P, by Q, to achieve R'. Hence, three fundamental questions were asked of each major theme. To illustrate this, consider the proposal to provide a single interface linking the various Inland Revenue computer systems in use.

P: What activity will be done? *A system to give staff access to a single, coherent, summary record of taxpayers' affairs.*

Q: How will it be done? *By providing a single IT interface which accesses all Revenue records for an individual, presenting that information in a single view and*

allowing single edit revisions to those records.[ie. A system in which, when data is updated on one record, the relevant changes are reflected in all the records holding that data rather than each record having to be updated individually.] .

R: Why will it be done? *To enhance customer service, provide proactive advice to taxpayers, improve efficiency and, over time lead to more reliable, up-to-date records*

A CATWOE follows from these answers.

- **Customers:** *Staff, Individuals.*
- **Actors:** *Inland Revenue and its IT partners.*
- **Transformation:** *Takes and updates information from different systems and presents it in coherent ways*
- **Weltanschauung.:** *It would allow staff to operate more efficiently, would improve customer service. Further, treating an individual's tax affairs as a single entity is a good thing. The change could also lead to a paperless office and may improve compliance.*
- **Ownership:** *Inland Revenue.*
- **Environmental constraints:** *the IT resources available, the legislation on data sharing and resources to run such a system.*

Finally, it is fundamental to SSM that such idealisations should be grounded in ways that allow their performance to be measured and controlled. This is usually captured in the notion of the 3 Es as follows (Checkland and Scholes, 1990), using the same example as before.

	Fundamental issue	Examples of performance measures
Efficacy	Does the system do what it is supposed to do? (Relates to P)	Allows single edit revision Delivers information required reliably (i.e. correct information)
Efficiency	Are minimum resources used? (Relates to Q)	Minimal resources to use and maintain
Effectiveness	Are the high-level aims of the system being met? (Relates to R)	Improves customer service Accurate and coherent customer data available

Activity modelling

In most SSM, the idea is to identify changes that are systemically desirable and culturally feasible (Checkland, 1981). Usually, this requires activity modelling, using systems concepts. In this study, this was done by developing the generic model of the

UK personal tax system shown in figure 3. It is important to realise that, in SSM, a model is not intended as a would-be representation of the real world; rather it is a vehicle to support debate – in this case, amongst the tax policy experts of the Inland Revenue. The generic model identifies those activities that must be present in any modern personal taxation system in the UK. For example, activities 1.1, 1.2 and 1.3 relate to the need to align the tax system with political priorities.

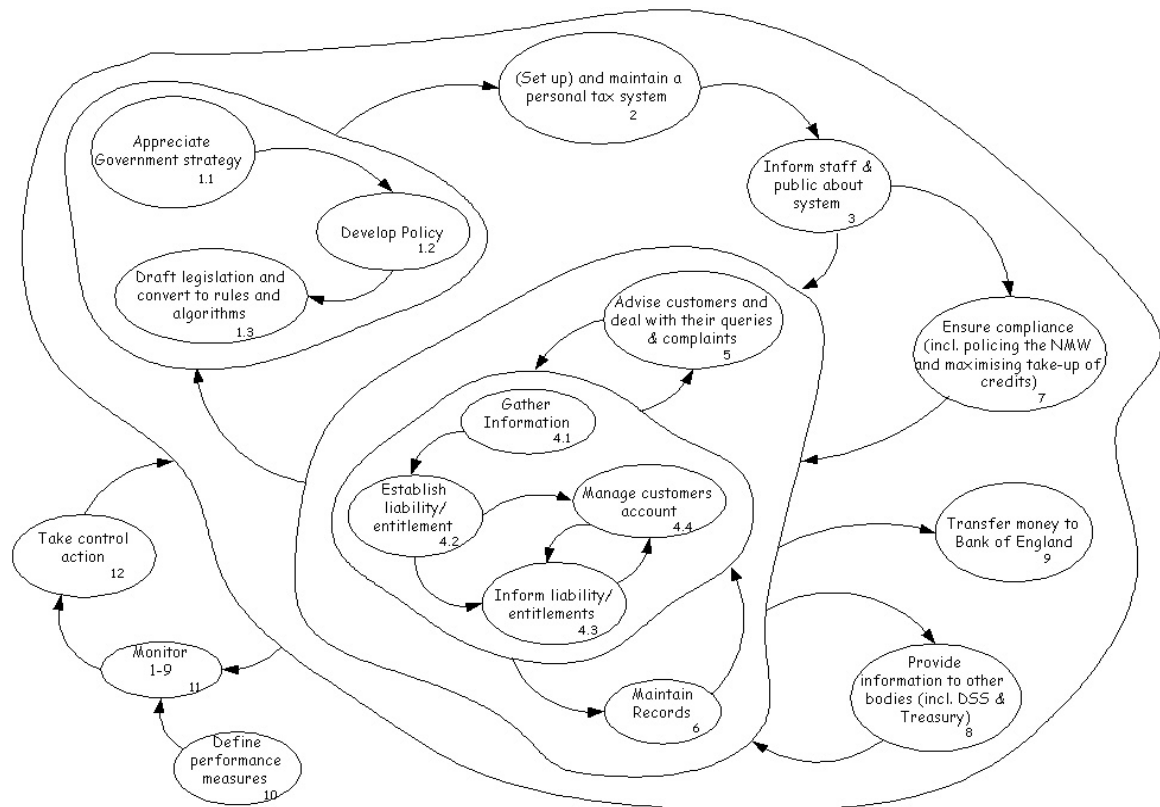


Figure 3: Activity model of a Generic UK tax system

The ten root definitions were reduced down to the four that potentially had the largest impact on areas of concern to the study, such as current system strains and issues of great concern to emerge from workshops. Whilst the others were not ignored, they were regarded as less pressing. Activity models, based on variations of the generic model, were then developed for the four systems encapsulated within the root definition.

These systems were thoroughly evaluated, in terms of their impact on customers, how well they supported the various roles the department now plays and the extent to which the system exhibited the ideal features stakeholders described in

workshops. They were also tested against possible future scenarios for the department to measure their robustness and adaptability to change. Finally, the implications of implementing and operating the proposed changes were elicited by drawing comparisons between activity models for the current and alternate systems. All this combined to give a rich account, and understanding, of the possible changes and to support further debate within the department.

COMPLEMENTARITY

Some general aspects

The preceding makes it clear that both the hard and soft streams of work contributed a great deal to the outcomes of the study in their own rights. SSM provided a consistent basis for the workshops, a methodology for analysing and interpreting the data they produced and a framework for taking the study forward. Similarly, the data mining provided learning on the operation of the current system, offered further insights into the needs of customer segments and ensured that these were objectively, rather than anecdotally, composed. However, the complementarity gained from combining the two further enhanced the value OR brought to the study. Using SSM to structure the process of carrying out the study also contributed to maximising the benefit to be gained from complementarity. It proved to be a valuable tool in bringing together the various strands of work within the study and in surfacing opportunities for combining the hard and soft OR approaches.

Throughout the course of the study, the two approaches complemented each other in a number of ways. For instance, amalgamating the findings from the data mining and SSM enriched the team's understanding of customers, their requirements and problems they may face with the personal tax system. Combining the two approaches allowed an extended coverage of the customer base. Data mining provided detailed information on the employee population, allowing the investigation of some quite small subgroups or segments within this population that could not practically have been represented in the SSM consultation. For example, some individuals have several directorships, and, therefore, fairly complex tax affairs; a segment given the name 'Career Directors'. On the other hand, the SSM consultation exercise allowed access to a much broader range of stakeholders in the personal tax system, including employers, tax agents and Inland Revenue staff as well as employees.

Complementarity was also achieved where the approaches overlapped through merging the different perspectives offered by each approach. Data mining identified current behaviour and taxpayer characteristics; for example, pensioners who had simple tax affairs (i.e. one stable source of income) would require a personal tax system involving only minimal, straightforward contact with the Revenue. By contrast, SSM consultations identified customers' needs based on the views, opinions and experiences of workshop participants. For example, consistent with the data mining evidence, some pensioners consulted wanted simplified forms to complete in line with their simple affairs. But others took this concept further, suggesting that Revenue should assume continuity of a taxpayer's affairs, with pensioners having to fill in forms only when their circumstances changed. Hence, in this way, data mining indicated some of the requirements of the system, and SSM allowed for the investigation of stakeholder requirements and provided some of the potential systems appropriate to them.

However, the complementary use of hard and soft OR means more than this, for it offers a synergy in which the whole is greater than the sum of the parts. Running the two streams of work in parallel opened opportunities for interaction between the approaches. Feeding outcomes from the hard into the soft and vice versa, was used repeatedly in the study both to enhance understanding and validate outcomes. As an obvious example, emergent findings from the data mining helped the team to understand how the different groups might view the current tax system, which informed the workshops. For example, the data mining highlighted the different pensioner segments. This was used in the pensioner workshop to ensure that the needs of those with simple affairs were not overlooked when considering the needs of pensioners with more complex affairs. Similarly, issues emerging from workshops, such as the burden of completing tax returns by those with simple affairs, could be investigated through data mining, which allowed the identification of customer groups most affected by this.

Another illustration of such complementarity was in testing the robustness of alternative systems developed through the SSM consultation. This testing was done by assessing the effect of the proposed changes on the customer groupings that emerged from the data mining. As well as checking for any unintended effects, this allowed the alternatives to be compared in terms of delivering the customer

requirements expressed in workshops and derived from the data mining. The results of the data mining allowed the quantification of any issues, and, where possible, enhancements to the models were made. For example, one system developed to ensure taxpayers with multiple and/or frequently changing sources of income paid the right tax in year, also had the effect of increasing the compliance burden on employees in single stable employment. The results of the data mining allowed the team to compare the relative size of each of the segments, helping to resolve this trade off.

Progressing the data mining and SSM in parallel however, also had its drawbacks. One obvious way that the soft and hard techniques could have complemented each other would have been to use the customer segments from the initial cluster analysis to construct the stakeholder groups for consultation. This was not done due to the time pressure to get SSM workshops underway. Yet, it was possible, later, to check the composition of the workshops to ensure that all the significant customer segments were adequately covered. Synthesising the material in this manner added a great deal of value to the project, providing a more detailed understanding of the problems raised and alternate systems considered.

Table 1 listed some ways in which hard and soft approaches differ and it seems appropriate to return to these in the context of the study of personal taxation.

Methodology used

Table 1 pointed out that methodological considerations are rarely an issue in hard OR, whereas they loom large in soft OR. This was certainly the case in the study described here. The data mining, though properly done, proceeded in a common sense way, adjusting the operation of the software and its parameters as necessary until requisite groups were obtained. By contrast, time was spent informing the Inland Revenue study team about the assumptions that underpin SSM as the soft approach used. This took the form of preliminary workshops with Inland Revenue staff, including tax specialists and OR workers. The aim was to allow them to understand the type of knowledge that is produced by an SSM study so as to manage expectations.

Models

The result of the data mining was a series of insights into the needs and behaviours of groups of taxpayers as discussed earlier. Since the data were believed to be substantially accurate, there was a reasonable belief that the resulting groups and insights were valid, forming a reasonable representation of taxpayers' behaviour. Thus, these models were regarded as representations of the ways in which people actually behave in the real world. The models that resulted from the SSM (see Brown, Checkland and Cooper, 2003) were activity, or conceptual, models that captured the essence of the activities thought necessary in any reform of the tax system that incorporated that particular proposal. That is, the SSM models represented concepts that the study team regarded as relevant to the real work of taxation, but were not intended as models of that world.

Validity

Data mining is regarded as a defensible approach to investigate relationships in a defined data set using established procedures. Thus, it is assumed that if another person followed the same procedures and used the same methods, they would produce similar groupings and insights. That is, the data mining models are broadly repeatable and are, in some sense, comparable, with the real world. There can be no such guarantee in the case of SSM, since one of its core assumptions is that the analyst is part of the study. Thus a different analyst or study team might produce different system designs. However, what is important, is that the conceptual models are defensibly coherent, logically consistent and plausible. This can only be checked with the co-operation of experts in the area being studied, in this case taxation, and here these formed the core members of the study team. However, a different team might have produced other models that were just as coherent, consistent and plausible.

Data

Table 1 suggested that a characteristic of hard OR is an assumption that data is "From a source that is defensibly there in the world, with an agreed or shared meaning, observer independent." This was certainly the case in the hard OR data mining. As mentioned earlier, great care was taken to validate the data that formed the samples on which the data mining was conducted. Where possible, uncertain items were

triangulated against other sources, to ensure that, when used, the data was as clean and reliable as possible. Thus, the team could be sure that the data source was defensible and that anyone with knowledge of the tax system would form the same understanding of its meaning. By contrast, data in soft OR is “Based on judgement, opinion, some ambiguity, observer dependent.” This was certainly true of the soft OR, which came from careful listening at SSM-based workshops of stakeholders. To reduce the risk of mistakes, more than one team member was at each workshop and their notes and interpretations were compared carefully. But it remains the case that this data, and the way in which it was expressed as SSM PQRs and CATWOEs, was observer dependent.

In his subversive critique of quantification, Boyle (2001) suggests that activities such as data mining, which are essentially impersonal, need to be complemented by interaction with the people involved. “Screeds of data about customers and how often they buy from you is not the same as a real measurement of ‘loyalty’.” (op cit, p149). Hence, in this study, the quantification of data mining was complemented by the SSM-based series of workshops in which real people discussed their particular needs of, and gripes about, the personal tax system. It is important to recognise that this complementarity is not just the simple combination of quantitative and qualitative approaches. Boyle (op cit) rightly argues that statistics representing populations can often be misleading and that direct personal experience is essential if the numbers themselves are to have any validity. Equally, it is true that any personal investigation, in this case based on SSM workshops, can focus on only a very limited number of people, whereas the data mining focused on a large sample from the population of taxpayers. In essence, neither approach will reveal all that there is to know. Taken together, their insights will be complementary.

Value & outcome of the study

Some form of quantified evaluation that allows option comparison is the hoped-for outcome of many hard OR studies. On the other hand, some form of agreement stemming from shared perceptions that will eventually inform learning and action is what comes from soft OR. What was the outcome from this study of personal taxation? Though it may be several years before the outcomes are seen in practice, it seems that both outcomes are evident. The data mining provided quantifiable

estimates of taxpayer segments – for example, as mentioned earlier, 40% of taxpayers had remained in the same job for the last 5 years. This can clearly inform any decisions about the need to provide different treatment for different segments. Thus, even hard OR can lead to improved shared perceptions. The soft OR, though, clearly led to shared perceptions about the types of response that would be needed from the Inland Revenue were it to implement the ideas for improvement from the various stakeholders. Together, the potential impact on changes to the tax system is very large.

Purpose of the study

A different tack is needed when discussing the purpose of the OR component of this study, for it was intended to be a complementary mixture of hard and soft from the start. In a sense, the purpose of the study was as in the hard column of table 1: it was taken as given to the study team by the Board of the Inland Revenue. However, in another sense, this was not true. One interpretation of the terms of reference would have led to the study team to move, as quickly as possible, to statements of what changes were needed and how they should be implemented, using computer systems or whatever seemed appropriate. However, the team wisely chose to act otherwise by trying to understand the taxpayer segments, the ways in which current systems worked for them, and the preferences of those segments as uncovered in the workshops.

Thus, though the terms of reference were taken as given, the interpretation of those terms was crucial to the success of the study. In this way, the team allowed themselves to learn as the work proceeded. They learned about how data mining and SSM may be used in such work and they learned about the current operation of the tax system and how stakeholders wish it to be in the future. They also learned what conceptual elements would be needed in any attempt to develop systems to implement any changes that may be agreed.

ACKNOWLEDGEMENTS

Thanks are due to the other members of the team who were helpful and supportive throughout. These are Brian Mace, Mary Aiston, Angela Walker and Ian Casey of the

Inland Revenue; Peter Checkland and Mark Westcombe from Lancaster University; and Steve Clarke, Mike Hayes and John Poulter, who were the consultants employed.

REFERENCES

- Balci O. (1987) Credibility assessment of simulation results: the state of the art. In Proceedings of the conference on methodology and validation, pp. 19-25, Orlando, Florida.
- Boyle D. (2001) The tyranny of numbers: why counting can't make us happy. Flamingo, London.
- Brown J., Cooper C. and Checkland P.B. (2003) Practising the practice of SSM in public policy (provisional title). In preparation.
- Burrell G.C. and Morgan G. (1979) Sociological paradigms and organisational analysis: elements of the sociology of corporate life. Heinemann Educational, London.
- Checkland P.B. (1981) Systems thinking, systems practice. John Wiley & Sons Ltd, Chichester.
- Checkland P.B. (2000) Systems thinking, systems practice (revised edition). John Wiley & Sons Ltd, Chichester.
- Checkland P.B. and Scholes J. (1990) Soft systems methodology in action. John Wiley & Sons Ltd, Chichester.
- Jackson M.C. and Keys P.A. (1984) Towards a system of systems methodologies. *Jnl Opl Res Soc*, 35, 473-486.
- Kohonen T. (2000) Self-organizing maps. Springer-Verlag, Berlin.
- Mingers J. and Gill A. (1997) Multimethodology: towards theory and practice for mixing and matching methodologies. John Wiley & Sons Ltd, Chichester.
- Pidd M. (1977) OR methodology. In *OR for managers*, S.C. Littlechild (ed). Philip Allan, Oxford.
- Pidd M. (1996) Tools for thinking: modelling in management science. John Wiley & Sons Ltd, Chichester.

Powell, S.G. 1995, "The teacher's forum: six key modeling heuristics", *Interfaces*, Vol. 25, No. 4, pp. 114-125.

Rosenhead J.V (1989) *Rational analysis for a problematic world*. John Wiley & Sons Ltd, Chichester.

Rosenhead J.V. and Mingers J. (2001) *Rational analysis for a problematic world revisited*. (2nd Edition). John Wiley & Sons Ltd, Chichester.

Simon HA (1976). From substantive to procedural rationality. In: Simon HA (1982), *Models of bounded rationality Volume II: Behavioural, Economic and Business Organisation*. MIT Press: Cambridge, MA, pp 424-443.

Willemain, T.R. 1994, "Insights on modeling from a dozen experts", *Operations Research*, Vol. 42, No. 2, pp. 213-222.