

RESEARCH STRATEGIES FOR POLICY RELEVANCE

Amanda Wolf¹
School of Government
Victoria University of Wellington

Abstract

Research can fail to be policy relevant when too little attention is paid to the “why” and “how” of policy change in the real world, and when relevant information remains elusive due to the complexity of social reality. Five strategies for researchers to consider when conceptualising new research are proposed, each of which addresses something about the mechanism of policy change. The strategies address issues concerned with both the availability of information and the fit of that information with policy argument needs. The five strategies serve to (1) generate new ideas about “what works” or what accounts for policy-relevant effects; (2) accrue ideas about the way mechanisms work for different people and in different circumstances; (3) improve understanding about why and how one mechanism works, and how it works in comparison with other mechanisms; (4) reveal the indirect mechanisms at work in a policy system; and (5) reinforce a realistic view of “causality” that supports timely action.

INTRODUCTION

Social science researchers, policy analysts and policy decision makers have long shared a commitment to better link research with policy. In historical perspective, today’s focus on research into “what works” renews attention to evidence after a long reign of efficiency in the policy spotlight. Evidence-based (sometimes softened to “evidence-aware”) policy is a recent term, but the idea is not new. At least since the 1930s, social scientists have provided input for policy development (Parsons 1995:20). In the 1960s and 1970s, these efforts emerged as a distinct field of enquiry, named policy analysis. Government agencies hired policy analysts in the genuine hope that they would crack the toughest policy nuts, such as poverty and illegal drug use, thus paving the way for effective interventions.

¹ **Correspondence**

Amanda Wolf, School of Government, Victoria University of Wellington, PO Box 600, Wellington, Telephone (04) 463 5712, email: amanda.wolf@vuw.ac.nz

Today's proponents of evidence-based policy are wiser. For instance, Nutley et al. (2003) caution against over-expectation when they add a contextual qualifier in their stock phrase "what works, for whom, and in which circumstances". False promise, however, continues. Proponents of evidence-based policy have made great strides in proposing ways to improve the knowledge base for decisions. However, people oversimplify when they argue that improved social knowledge will lead to better policy outcomes. Not all knowledge will inform better decision making. And not all good decision making will lead to better outcomes. This article addresses the issue from the orientation of research design and methodology – from the strategies researchers use to produce knowledge – rather than that of the knowledge that is produced. In particular, my aim is to improve the chances that research outputs will indeed inform decision making, because they will be more policy relevant.

A semantic confusion needs to be addressed at the outset. I follow Majone's view that "evidence" is not the same as "information", although the two are often conflated. Majone writes that "evidence" is "information selected from the available stock and introduced at a specific point in the argument in order to persuade a particular audience of the truth or falsity of a statement" (1989:10). When the two words are conflated, we may fail to discriminate problems of availability from problems of fit for purpose.

In New Zealand, the "availability" (or information) problem predominates, expressed as either a lack of New-Zealand-specific information, or a lack of ways to make better use of a flood of mostly international information. The common prescription is to set priorities to improve the stock of New Zealand information. However, the "fit-for-purpose" (or evidence) problem needs to be addressed as well. This problem is a subset of the availability problem, as reflected in Majone's definition. More specifically, an evidence problem is apparent when very little of the vast stock of information available makes its way into policy arguments (Shulock 1999), or where the information that is asserted in support of a policy decision is unpersuasive.

Policy decisions, without doubt, should be well informed. Information may be valuable and interesting, but it is not evidence until it serves to make a policy argument, perhaps supporting a problem definition, or explaining or justifying a policy response. Policy researchers, then, might ask how effectively they are producing information that supports good decision making. Policy-relevant research results from a good match, an effective blending of researchers' awareness of policy decision needs and decision makers' awareness of a knowledge base that bears on their decision.

In this article I direct attention mainly to the issue of how to improve the policy relevance of research (where that is the espoused objective of the research in question). I focus at the front end of the evidence challenge, by considering how social science

researchers can better conceptualise studies to improve the policy relevance of their results. Research conceptualisation refers to the initial design or vision of a research project, more like an architect's concept drawing than like a builder's blueprint (Hakim 2000:1). All successful conceptualisations match a "solution" to a client's "problem". In the policy research context, successful research starts with policy developers' and decision makers' ("users") knowledge needs, and hinges on how well these needs are specified. From the researcher's standpoint, the design challenge is how best to provide users with a solution they are happy with.

Research conceptualisation is key to achieving policy relevance. The architectural metaphor readily evokes the core challenge. Good architects have, in addition to technical skill and materials knowledge, a design sense that allows them to avoid bad ideas and, more often than not, provide a design concept for a structure or space that satisfies their commissioners. Similarly, policy-relevant research – relevant by design, and not by accident or serendipity – is underpinned and informed, more often than not, by well-conceived approaches to generating research solutions. Put somewhat differently, policy relevance is not an automatic by-product of "good" research.

Thus, to keep focused at the conceptualisation stage, I assume ideal research conditions. In these cases, researchers are fully able to contribute high-quality research to a well-functioning, realistic policy process within recognised constraints. Thus, I skirt challenges relating to research capacity and capability (technical skills, funding), to government preferences (strategic policy development, choice of desired outcomes), and to logistical or personal factors (awareness of research findings, willingness to consider unpalatable findings, maintaining reasonable mutual expectations between researcher and policy developer). What remains to focus on is the "reflective practitioner's" (Schön 1983) knack for seeing to the heart of what *really* needs to be known, if that knowledge is to qualify as policy relevant.

In the next section I illustrate how research can fail to be policy relevant through shortcomings in the information content of research, and through failures in information fit and extent, even where information quality is high. This review of the problem paves the way for considering five methodological strategies to improve relevance. Each strategy responds to a slightly different challenge facing researchers, but they overlap and are mutually reinforcing. I conclude on a cautiously positive note, for there are no serious hindrances to more widespread uptake of the suggested practices (indeed, all are currently in practice to some degree).

THE "RELEVANCE" PROBLEM

Policy-relevant research presents what has been, what is and what is likely to be, in a specific social context, in order to inform policy decisions. Continuing with the ideal scenario, let us assume that users are satisfied when researchers provide (enough) high-quality information, in understandable terms, which they can use to make and justify robust policy choices. In this formulation there are two critical links: one concerns the information content, and the other its contribution to policy choice. Greenberg et al. (2000:367) offer a similar view on critical links in the context of the usefulness of policy experiments by highlighting the reliability of information, the connection to the right users and the actual use of the information in decision making. While recognising the connections between available content and usefulness, weaknesses in each type are separately considered below.

Information Content Limitations: "What" Answered, Not "Why" and "How"

Many policy researchers focus on "what" questions and produce descriptive information. A recent discussion paper on knowledge needs from Statistics New Zealand (2003) provides a summary of policy questions organised into cross-cutting themes, such as population and security. Under the heading "Culture and Identity", for example, are a number of exploratory and descriptive questions, including:

- How do people living in New Zealand identify themselves and to what groups do they feel they belong? (This asks, in essence, What identifiers do people living in New Zealand apply to themselves?)
- What are different groups of New Zealanders' attitudes to "belonging" in New Zealand, and what is their experience of various aspects of life in New Zealand?
- What is the current status and health of the Māori language?

In general, "what" questions require a descriptive answer about a social phenomenon: What types of people are involved, and what are their characteristics? What knowledge, beliefs, values and attitudes do they hold? What is likely to happen? "Why" questions build on descriptive information to investigate the causes of, or reasons for, characteristics or patterns in a social phenomenon. They are directed toward understanding and explaining: Why do people think and act this way? Why does this activity have these particular consequences? Finally, "how" questions are concerned with bringing about change: How can people's attitudes or behaviour be changed? (Blaikie 2000:60-61).

"What" questions, of course, have immense value, both on their own and as precursors for "why" and "how" questions. For example, the National Advisory Council on the Employment of Women (NACEW) coordinated a large-scale survey on childcare. NACEW wanted to investigate the use of childcare and the factors that affect use or

non-use of childcare for labour market purposes. They designed an add-on to the Household Labour Force Survey to provide data needed to find whether a lack of childcare was a barrier to participation in employment, voluntary work, or study and training. In the report, we read that the survey finally laid to rest one concern:

NACEW was initially concerned that parents struggled to juggle a plethora of types of childcare... In fact, for the majority of families this is not now the case, with 80% of pre-school children who use ECE [early childhood education] and care using no more than one type, and very few using more than two types. (NACEW 1999:5)

While descriptive, this information does not require a “why” or “how” answer for policy relevance, and users can confidently assume that “juggling childcare” is a minor facet of the problem.

There are, however, three dangers lurking in an overemphasis on “what” questions, as some further examples will illustrate.

First, researchers may think that more, or richer, descriptive information must necessarily be of greater value than less of it. The danger here is largely one of lost efficiency, as resources that could be shifted to addressing “why” and “how” questions are expended in more “what” questions. We find such cases in the vogue for meta-analyses. Meta-analyses are valued when they reveal conclusions that demonstrably reduce existing uncertainty, but their costs can easily exceed the expected return. (Here, as in the other examples I cite, I am not disregarding the obvious benefits of a research tool, such as meta-analysis – which has a wide range of applications – nor judging the specific research cited, but simply making use of an actual example to make a more general observation.) For example, Nutley et al. highlight a systematic review of 23 research studies to examine whether “job absenteeism is an indicator of job dissatisfaction”. They conclude, “yes” (2003:13). Such a commonsense result is only partly rescued by the qualifier, “stronger association was observed between job satisfaction and frequency of absence than between satisfaction and duration of absence” (ibid.).

A second danger with descriptive research occurs when users put too much faith in the policy relevance of descriptive information. Users may expect that available information will be usable as evidence, and researchers may be prone to assert information as evidence to satisfy this expectation. Returning to the childcare survey, we find that a great deal of the accumulated descriptive information is thin in “evidence” potential. The report ends:

It is apparent from the data that there is a relationship between labour force status, household composition, age of children and the amount and type of ECE and care used. However, it is less clear how these factors interact to influence the decisions parents make about their work and ECE and care arrangements... It is evident that for some parents (particularly sole parents and those on lower incomes) lack of access to quality, affordable ECE and care constitutes a barrier to their participation in paid work, unpaid work, or study/training. These barriers need to be fully understood and addressed in order to give parents options about their ECE and care arrangements and their economic participation. (NACEW 1999:58-59)

The survey has developed a detailed picture, maybe even a nearly “true” picture, which could be improved with advanced statistical examination of the data. However, the picture, on its own, falls short of the understanding needed for policy improvements. Given the nature of the survey, and the success of some of its revelations, researchers (and users) could conclude that the quoted paragraph calls for even more detailed descriptions. Note, in the quoted extract, an expressed lack of knowledge: “However, it is less clear how these factors interact to influence the decisions parents make.” Despite the use of the interrogative “how”, the expression calls for clarity about what interactions occur and with what influences on parents’ decisions.

Even granting that this clarity would suggest explanations for patterns in childcare usage (in the form of complex associations), associations of interactions and decisions would not, on their own, tell us why parents chose as they did. For instance, a pattern might emerge that suggests that parents tend to stop using grandparents as caregivers when their children start school and their incomes increase moderately. Such information does not support causal explanations. Similarly, knowing that there are barriers does not explain what holds them in place, or what to do to reduce them effectively. Researchers might accept that the picture as it stands in the survey results is rich enough to shift to a “why” or “how” focus to pursue a fuller understanding needed for policy improvement.

In the third and most serious type of information content failure, users assume that a “what” result carries its own “why” and “how”. In an example of this failure, California spent US\$5 billion to reduce class sizes following an evaluation of an experiment called STAR (Student/Teacher Achievement Ratio) run in Tennessee, which showed significant achievement improvements in early grades associated with reduced class sizes. But the investment failed to change Californian children’s achievements (Hirsch 2002). Californian school officials might have paused to consider what STAR (as evaluated) actually showed: the improvement in achievement was not accidental in the circumstances of the STAR experiment, but that does not mean the effects will be repeated in new circumstances. The experimental outcomes (in

retrospect) appear to have depended on other factors, in addition to class size, that were present in Tennessee but absent (or different) in California. Sadly, the multi-million dollar STAR study lacked a theoretical interpretation of its own findings, which might have saved California taxpayers billions. In essence, the study showed “what” but not “why”. An effect was described, but the study did not try to tease out why smaller class sizes were more effective, nor did it look at how student achievement changes were brought about.

Information-Fit Limitations: The Constraints of Diversity and Complexity

The evidence-based policy mantra of “what works, for whom and in which circumstances” will clearly be challenged as the “whom” increase in their diversity and the “which circumstances” in their complexity. New Zealand is a small country, and research capacity, even at its ideal best, is likely to remain very limited relative to the diversity and complexity of the things we would like to know more about. As researchers know well, small effects, which nevertheless may contain significant policy information, can too easily remain undetected at the scale of research that is usually possible in this country. In the face of diversity and complexity, policy questions are inevitably simplified and available information may fall far short of policy relevance. In this situation, the information failure is that the available information fails to be sufficiently fine-grained for policy relevance. Still, unlike the limitations canvassed in the previous section, this is not so much a pathology of research design as it is a straightforward question of scope, available resources and, sometimes, natural limitations (as when a population of interest remains too small to be studied with statistical robustness).

Limitations due to necessary simplification in the research questions addressed can be distinguished from limitations inherent in the complexity of a situation. In the latter case, understanding of social phenomena at a level suitable for “evidence” requires specific strategies designed with complexity in mind. The authors of a study of the housing needs of mental health clients (also referred to as “tangata whai ora” in the report) express the dilemma of reasonably full information coinciding with relatively poor understanding:

The complexity of the issues facing people who experience mental illness makes it difficult to identify the factors related to housing that pre-eminently impact upon wellbeing... Overall the group interviews highlighted a wide range of factors that compromise the capacity of consumers/tangata whai ora to sustain independent living. The most significant finding, however, is that it is a combination of factors rather than any one factor on its own that creates situations where consumers / tangata whai ora may not be able to manage living independently in a sustainable way. (Peace et al. 2002:64-67)

In view of this, the authors show that the solution is not more fine-grained information, but a different way of viewing the information that exists. They foreshadow a solution by proposing a “sustainability framework” to help researchers and policy makers “conceptualise the inter-linked and complex factors that affect the lives of consumers/tangata whai ora” (p. 67, and Part Three of the study report). The writers continue:

From this understanding, new styles of integrated policies and services may be developed that work to integrate goals, always focused on the full range of needs of the consumer/tangata whai ora rather than on isolated aspects of their lives. (p. 67)

This is just one example of a widespread challenge to research design. It should be seen as an opportunity to seek out new strategies for creating evidence regarding complex phenomena of policy interest.

FIVE STRATEGIES

The issues raised in the previous section have some ready answers. Researchers should focus more on answering “why” and “how” questions, as well as “what”. Researchers should concentrate less on simple explanations and more on better understanding of complex causation. These are methodological aims. Just as prescriptions to “close knowledge gaps” lead to lists of knowledge priorities, so methodological gaps call for priority research strategies. Research strategies concern how research questions are answered and, thus, whether the research will be policy relevant. In this section I offer five research design strategies. Each works to bring the researcher and user into more productive relationships. Each also focuses, as it must to meet evidence needs, on the mechanisms of policy change.

Strategy 1: To complement inductive and deductive methodologies researchers can explicitly attend to “abduction”, which gives rise to initial hypotheses

This strategy aims to short-circuit the lengthy and expensive process of acquiring descriptive information about a social phenomenon and then “testing” the reliability or validity of this information. In a policy context, it is not a strategy to be used in isolation. Yet it would help to narrow the gap between descriptive and explanatory information, and thus to increase the evidence-readiness of information about how policy variables affect desired outcomes.

Most social scientists have long grown accustomed to associating the terms “inductive” or “deductive” with contrasting research strategies. Inductive strategies are those that are as premise-free as possible, and seek to discover common patterns in the social phenomena studied; they may go further, and establish tentative general explanatory

conclusions about those patterns. Deductive strategies, in contrast, are explicitly grounded in theory and seek to establish instances in which social phenomena exhibit or fail to exhibit features expected by some theory or other.

“Abduction”, a term introduced by Charles Peirce (1998 vol.5:145), is the process of adopting an explanatory hypothesis (which may then be “tested” to see if it helps to explain patterns observed in social phenomena or empirical data). It refers to the instinctual processes by which a thinker (such as a researcher) uses other information to narrow the otherwise infinite possible causes and explanations to formulate plausible hypotheses. Abduction, which Peirce also calls “retroduction”, is “inference to the best explanation”, through “reasoning backwards” from consequent to antecedent (Wirth n.d.). In essence, according to Peirce, a surprising fact is observed; the thinker reasons that if some hypothesis, *H*, were true, then the fact would be a matter of course. Hence there is reason to suspect that *H* is true (Peirce 1998 vol.5:189). Ideas about *H* come from the researcher’s own experience (e.g. an experienced lab scientist might make guesses about a new substance in a beaker) or from research designed to generate possible hypotheses.

Researchers practise abduction as a matter of course, but making this mode of thinking explicit heightens the chances that researchers will use its strengths. Researchers who are familiar with mixed-method approaches will note that abduction is a related “dialogue of ideas and evidence” (Ragin 2000). While abductive reasoning can well serve all policy-relevant research, it may be particularly helpful in situations of complexity and diversity (Wolf 2002). Morçöl (2002), taking into account recent “post-Newtonian” views of complexity, cognitive theory and other aspects of the new sciences, can be read as an argument for abductive reasoning in policy research.

There are three distinct ways for researchers to draw on abduction and its outputs. First, researchers can be more reflective about how they think their way to plausible explanations. They can include the search for and selection of hypotheses as a distinct research task. Second, researchers can use abductive reasoning, as it is embodied in a small set of methodologies which have been developed to highlight the emergent properties of abduction. Charles Ragin’s *Fuzzy Set Social Science* (2000) presents a methodology (which Ragin classifies as “retroductive”) to discover complex causality. Abduction has a role in Q-methodology, which is an approach to discover emergent properties in people’s subjectivity (Brown 1980). Morçöl (2002) adds agent-based modelling and repertory grid techniques to the two methodologies already cited, as suited to the “new mind for policy analysis”, allowing and creating the possibility for new policy-relevant understandings to emerge (pp. 253-255).

Third, researchers can draw on the abductive thinking of research subjects. For example, experienced social workers excel at “reading” situations. For good reasons,

we constrain their ability to act on their own hunches alone, yet we have much to learn from systematic study of these judgements. Abductive methodologies work with the “whole” of a phenomenon of inquiry. The researcher simultaneously “discovers” and “brings forward” salient aspects or combinations of aspects (such as arguably are present in the social workers’ thinking) in a manner that allows further scrutiny. In this way, hunches that stimulate policy-relevant research arise from the successes and failures of current and past policies. Policy-relevant research more often needs to start with the policy questions themselves, which are often best generated by users and others who have “experienced” policy.

Strategy 2: To strike a balance between strong information with no theory behind it and strong theory with no information behind it, researchers can develop and articulate middle-range theories

Ray Pawson (2002) reinforces many ideas presented throughout this paper. In this section I wish to focus on one element in particular of his “realist synthesis” strategy. Middle-range theories represent the idea content (as distinct from the factual content) of empirical data, but fall short of grand theory. In essence, they present theories of “what works, for whom, and in which circumstances”. (Pawson refers to “CMO propositions, where C is context, M is mechanism, and O is either a positive or negative outcome.)

In designing research to develop middle-range theories, researchers look closely at the core mechanism of an intervention. A “mechanism” comprises the essential elements that do the basic causal work to change institutional conduct or people’s behaviour in some way (Bardach 2000:77). Mechanisms describe what is doing the work. As well, researchers look to outcomes for measures of the work done by the mechanism. Pawson’s “context” incorporates both the circumstances in which, and for whom, a mechanism works, as well as elements of a broader context, such as societal norms. Middle-range theories create strong links between research and policy. According to Pawson, “Since policy making is itself a conceptual, conjectural, and self-revising process, then well-grounded, middle-range theory building is the most apposite source of inspiration” (2002:213).

Pawson’s work emerges from the traditions of scholars who question the usefulness of policy research and of careful observers of the methodological challenges of making sense of messy policy realities. He argues that while there are infinitely many “cases” of policy interventions – certainly one for each person/policy pair – there are only a small number of mechanisms for change. Thus, researchers can assume some underlying mechanism of change is behind policy in large numbers of interventions, each of which is a “case” of the mechanism. Researchers can extract the “emerging propositions” by reviewing the cases (Pawson 2002).

Following Merton, Pawson praises middle-range theory – theory using “concepts that are concrete enough to explain the specifics of any intervention, but abstract enough to confederate our understanding of the different policy domains” (2002:214). He uses the intervention of “naming and shaming” (applied to such cases as hospital mortality report cards, car crime indices and car safety reports) to illustrate how realist synthesis aims to account for differences in the effectiveness of an intervention across all instances of both success and failure. In querying all instances, Pawson’s stance is abductive: new ideas, as they emerge, help to make sense of even more “cases”. His methodology shows how to systematically refine and synthesise ideas (theories).

The basic device is a series of comparisons of cases of naming and shaming, which draws out, from the often-sketchy available data, details of context, mechanism and outcome (CMO). To further guide the researchers, a series of generic questions can be set up regarding the CMO variables as manifested in instances of success and failure, as illustrated in Table 1.

Table 1 Success and Failure in Naming and Shaming

	Success	Failure
Identification	Performance / behaviour in question is observed, classified, rated, etc.	Identification or classification of observed behaviours is inappropriate, measures are over- or under-discriminate, etc.
Naming	Information on failing / deviant party is disclosed, disseminated, etc.	Disclosure is poorly managed, over-restricted; people wrangle about the meaning of the information
Public sanction	Broader community acts on disclosure in order to shame, censure, control	Wider public go beyond shaming (humiliation, vigilantism) or fall short of shaming (apathy or sympathy)
Recipient response	Behavioural change follows sanction, with subject being regretful, re-integrated, etc.	Individual or institution accepts “shame” as a badge; rejects and continues to adopt a perverse response

Source: adapted from Pawson 2002.

Pawson’s method starts with some general questions, such as:

- What to disclose and how loudly to shout?
- Who can be moved to shame and who can deliver the shaming sanction with authority and magnitude?
- What are the kinds of problems for which the emotional challenge of shaming will result in behavioural changes?

He ends with ideas or hypotheses that can be further explored empirically, but that can also directly inform policy development. For instance, Pawson's initial results suggest that shaming appears to work better when the responsible bodies have "watchdog credentials" and a range of sanctions at their disposal (2002:225). Further, the cases suggest a fruitful angle for researchers to investigate in order to fine-tune theory (ibid.). All else being equal, would shaming work best when the responsible party was quick with penalties, or when they kept the penalties in clear view of the shamed without actually using them?

Strategy 3: To avoid leaps of faith in working with policy hypotheses, research can continually query assumptions using intervention logic

Pawson urges empirical attention to the core mechanism of an intervention as a means to draw out the "idea content" and to fine-tune explanations of the observed effects of policies. Identifying the timing or threat of penalties is a step to further enquiry. In a related (but not identical) fashion, "intervention logic" supports conceptual attention to the core theory or policy hypothesis of a mechanism. Policy hypotheses are propositions about the way policy brings about a behavioural change and, therefore, a desired outcome. Starting with the idea and not the facts, one can examine penalty as "penalty", not as a set of penalty observations. By focusing at the level of ideas, intervention logic can provide researchers with a tool for querying assumptions (as well as for its better-known functions in managing for outcomes) (Baehler 2002). As researchers know, a key to solid empirical research is to first work through the assumptions in a hypothesis, systematically and neutrally.

Leaps of faith are easy to make when working with policy hypotheses: an hypothesis can assert a causal mechanism without specifying one. For instance, many policies provide information services to clients. The core logic is that, being informed, the client will exhibit a changed (for the better) behaviour. However, it is not only possible that the information can change the "wrong" behaviour (e.g. the fear of some that sex education leads to promiscuity), but that we may also fail to understand the chain in the causal connections between receipt of information and behavioural change. For example, a client has to understand and internalise the information, to acknowledge that it applies to their situation, to consider the implications of the new information for their day-to-day behaviour, to be persuaded that a change is good, and then to make the change (which must also be shown to have been the "right" change).

A given policy intervention can support a range of hypotheses. For example, there are several possible mechanisms through which increasing police presence on roads (intervention) can reduce average traffic speed (effect). Each mechanism, together or in combination, may be active in the observed effect. It may be that all drivers, when they see a police cruiser, slow down a little bit, thus lowering by that little bit the average

traffic speed. Or, the presence of police cruisers on the roads may decrease average traffic speed if police stop and ban the worst speeders from driving, while everyone else drives at their customary speeds. Research that seeks to identify and compare hypotheses about mechanisms will often be structured around the assumptions in competing hypotheses.

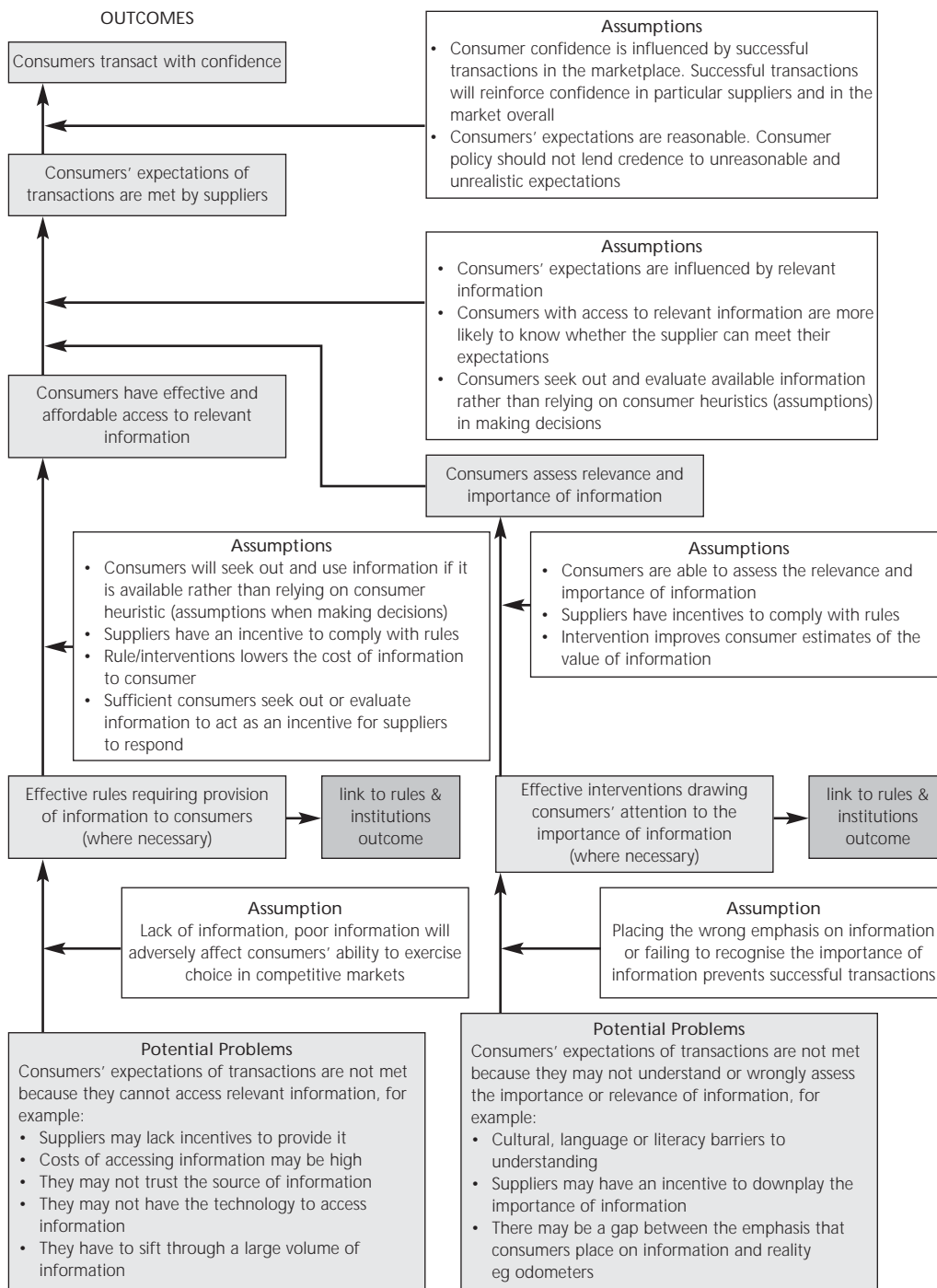
Figure 1 presents an intervention logic diagram, prepared by the Ministry of Consumer Affairs (2003:71), that is an outstanding assumption-querying framework. As presented it is quite generic, but it lends itself to tailoring in specific situations. For example, according to the New Zealand Medicines and Medical Devices Safety Authority (Medsafe), under current laws consumers do not always have the knowledge and skills required to distinguish between the valid claims made by reputable distributors of dietary supplements (herbal remedies and the like), and the sorts of extravagant claims made for some products (2003:11). In New Zealand, specific therapeutic claims can only be made in respect of a registered medicine. Medsafe details many instances of illegitimate claims, such as unsubstantiated claims, false information, omissions or discrepancies due to translation error or failure to provide translation, aggressive marketing, inadequate information, and borderline products sold as dietary supplements (2003:56-62). However, although some of the products with faulty claims are associated with serious illness or death, illegal claims do not necessarily lead to bad outcomes. (Nor, for that matter, does the presence of accurate information ensure safe use of products.) It can be surprisingly difficult for researchers to understand how a policy variable functions in a complicated system, such as how a claiming rule functions in a consumer's information environment.

A researcher interested in supporting an improved policy for information on dietary supplements can use the flowchart in Figure 1 to fine-tune a research design. Good policy research conceptualisation begins in situ with actual circumstances. However, it also needs a sense of desired outcomes and a general idea of how proposed policy interventions might "work" to achieve those outcomes. While there are different ways of expressing end objectives, one view is that when the policy is "fixed", consumers will have "better access to comprehensive, reliable and objective information" about dietary supplements (Ministerial Advisory Committee on Complementary and Alternative Health, Ministry of Health 2003). The Ministry's flowchart puts this in terms of a generic objective – consumers transact with confidence – and defines two sorts of potential problems: consumers' expectations of transactions are not met because:

- they cannot access relevant information
- they may not understand, or may wrongly assess, the importance or relevance of information.

Both of these problems are further detailed, presenting to the prospective researcher a rich array of matters to probe. In addition, the flowchart presents 14 "assumptions"

Figure 1 Ministry of Consumer Affairs, Flowchart 2: Meeting Consumers' Expectations: The Information Dimension



Source: Ministry of Consumer Affairs 2003:71.

that relate to the causal chain linking information and the outcome, “consumers transact with confidence”. Each of these assumptions is fertile ground for researchers seeking a way to link research and policy development.

Strategy 4: To avoid artificially segmenting reality, researchers can seek to understand the system and the variables of behavioural change into which a policy will be fitted

In a number of circumstances, researchers appropriately focus on isolating a single variable (or a small number of variables) at a time. However, it does not follow that the results of the research should be transferred directly into a policy prescription. Where the causes of a matter of policy concern are complex and multiple, variable-at-a-time policy responses may be ineffective or counter-effective. Media stories dramatise the systemic nature of policy, raising, for example, the claim that when police are engaged in speeding or drink-driving blitz campaigns, unsolved burglaries will necessarily rise. This directs attention, as it were, to the bulge in the balloon and to the forces applied elsewhere that have produced the bulge.

Fortunately, in a number of areas, researchers and users are starting to get systems frameworks in place, which is a precondition for system-level research and policy development. An example is the Department of Labour’s Human Capability Framework (n.d.). Informative, systems-based analyses are increasingly routine. (Witness the increased interest in policy networks and increased attention to operational “risk-scanning” for a very different manifestation of systems thinking.) Nevertheless, even as systems awareness has increased, its connections to policy action lag (Stewart and Ayres 2001:80). Policy research needs to do more than look at variables that can be altered by public policy.

Systems approaches offer new ways to zero in on the self-organising and adaptive capacities of systems. Systems thinking builds on an understanding of the phenomenon of interest as a subset of more general processes and relationships. A key challenge is to shift a long-established orientation to policy variables to one of systemic relationships. The tools of government action, reflected above in Bardach’s definition of “mechanism” as the “essential elements that do the basic causal work to change institutional conduct or people’s behaviour”, reflect a behaviourist model of government action. Yet systems approaches suggest that, although behavioural change may still be the “final cause” of policy change (following Aristotle, behavioural change is what the policy change is for), the efficient cause(s) (which make the behaviour change) may be system based. For example, to combat childhood obesity, it is clear that the at-risk child must engage in behaviours such as eating better and exercising more. But what will move the child to do this? Can a policy be designed that would insinuate itself between the child and the chippies? Note that we do not necessarily need to

understand why the child is at risk of becoming obese in order to effectively counteract the behaviours in question.

To the extent a policy problem is a manifestation of a set of interrelated (systemic) elements, at least some of which reflect past policy choices, there is a “problem situation” with many causes (Stewart and Ayres 2001:84). Thus, interventions may not always most effectively be applied to behaviours via incentives and sanctions, and so on. Instead, policy intervention may be better directed to act on the patterns of communication, influence and exchange in a system context. Policy-relevant research at a systems level is still in its infancy, but offers intriguing possibilities.

While somewhat tangential to many social policy areas, scholarship on natural disasters is illustrative. Human activity has sufficiently intersected with the natural world as to be accorded causal status in some disasters. For instance, people who build houses on steep terrain may both contribute to and suffer from rain-induced slips. At the same time, some disasters, such as the expected bursting of a tephra dam on Mt Ruapehu, are natural equilibrations that are redefined as disastrous. Moving to the social sphere, the greater interest in systems methodologies, such as soft-systems analysis (Checkland 1999), classical dynamic modelling and complexity reading (Hutchinson 2002), while challenging, promise commensurate returns.

Strategy 5: To align researchers’ and users’ standards for research excellence, both parties can be more cognisant of the need to strike, and make explicit, a balance between “good enough” research rigour and “good enough” information for timely, acceptable action

The final strategy is simple to state, but cuts to the core of the “two cultures” problem first noted (in a different context) by C.P. Snow. Weimer and Vining (1999:29-30) frame the two-cultures divide in terms of the conventional objectives of each. Policy, they note, has as its main objective bringing about a change for the better (or stopping a change from leading to worse outcomes). Policy research has as its objective estimating the impacts of changes in variables that can be altered by public policy.

Thus, policy-relevant research requires bridging the gulf between the conventional objectives of policy research and of policy. Users readily accept that the information is never all in when a decision needs to be made, yet the evidence-based policy norm idealises 100% certainty from research. At a recent policy seminar, a senior policy manager caricatured an academic researcher spending years tracking one research quarry, whereas a typical policy analyst pursues 20 in half an hour. Perhaps it is time for hybrid sport.

The NACEW childcare survey will serve as an illustration once again. It was administered as a 12-minute add-on to the Household Labour Force Survey, and its report runs to 191 pages, of which 80 are tabulated data. One may ask if the survey results would still be “good enough” if certain questions had been omitted, and the corresponding resources redirected to providing a “good enough” picture of an additional target.

In addition, we might look again at the various activities under way to improve the knowledge base for policy decisions. Might we more closely target the “must have” information, and reserve resources for “just-in-time” efforts as circumstances unfold that call for new and different understandings?

SUMMARY AND CONCLUSIONS

Research can fail to be policy relevant when too little attention is paid to the “why” and “how” of policy change in the real world, and when relevant information remains elusive due to the complexity of social reality. Five strategies for advancing the policy relevance of research are proposed, each of which addresses something about the mechanism of policy change. Each suggestion has been supported by examples, even if successful practice is not yet widespread. Table 2 summarises the value of each strategy as a means of increasing the chances that users will find the evidence they need for decision making as it bears on the mechanisms of policy change.

Table 2 Summary of “Mechanism” Focus in Five Strategies

Strategy	Mechanism Focus
1. Seek “abductive” insights through heightened attention to the possibility of their emergence when using specific research designs and questions	1. Generates possible ideas about “what works” and/or understanding of policy-relevant effects
2. Balance information and theory by developing middle-range theories	2. Accrues ideas about the way mechanisms work, “for whom” and “in which circumstances”
3. Avoid leaps of faith in working with policy hypotheses by using intervention logic to query assumptions	3. Accrues understanding about why and how a mechanism works
4. Avoid artificially segmenting reality, using systems approaches	4. Reveals indirect mechanisms in policy
5. Seek to balance “good enough” research rigour and “good enough” information for timely, acceptable action	5. Reinforces pragmatism – one need not be certain to act well

Together, the five suggested strategies amount to a shift in focal length, more than a shift to “new” research designs altogether. To close this paper, I similarly wish to shift my focus, from research design to researcher (supported by the user), and their

orientation to the policy relevance of research. Research becomes relevant to policy when its knowledge content fits the evidential requirements of a policy argument. To increase the chance of good fit, researchers need to become more policy-aware, working from real policy questions, seeing more clearly where the evidence will slot in. Users must accept that “evidence” is the name for information in an argument, and, so, it is theirs to commission or select. They need to look with more evidence-aware eyes on the information and incipient policy arguments provided by researchers.

I will start by posing caricatures of researchers and users in the sphere of policy relevance. Researchers are happiest pursuing esoteric questions of interest to themselves. They believe their results to be so profound as to have an obvious place at the heart of decision making, and are stunned when users ignore them. Users want clear, instant answers to problems as they see them, and cannot bear researchers who claim that it is not as simple as they see it. Neither has much patience for the other’s point of view, which leads to the obvious prescription to work on finding and cultivating the common ground between them. I suggest three means for this.

First, researchers must engage more proactively with users, and vice versa. Researchers need to develop their designs to shine research light where users see the shadows of evidence to lie, not only where they have a curiosity to look. Conversely, users might point researchers toward the unexplored terrain that interests them. In short, researchers can elicit clearer needs from users, and users can attune themselves to what ideas and information researchers have. Such moves will decrease tendencies to oversimplify research problems, and to hang unrealistic hopes on research results. Researchers who are more policy-aware will find themselves with meatier challenges to inspire and reward them. Users who are more evidence-aware will steer researchers to questions most begging to be answered.

Second, researchers should acknowledge that their concepts and theories do not always answer policy questions (Shulock 1999). As Pawson notes, “research must follow the metre, vernacular, and content of policy decisions” (2002:227). Knowledge “of” (a policy variable, for instance) is not the same as knowledge “for” (creating change in that variable, for instance). Even further, knowledge “just because” (it is a curiosity worth investigating) is not the same as knowledge “so that” (a better policy can be set). Researchers may need to become more innately interdisciplinary, in the sense of being guided by applied problems. Researchers need to learn to think in terms of policy change and policy mechanisms, regardless of their particular disciplines. After all, economic, legal and political theories are all addressed to social behaviour, which in the real world is an undifferentiated mass, merely seen through different lenses. In the social sphere, academic subdivisions (economics or politics, for instance) turn on the definitions of terms and not on subdivisions inherent in the social phenomena under study (Streeten 2000:11).

Third, a notion of working from the middle, or finding a suitable balance between competing forces, is a recurring theme in this article. I have highlighted middle-range theories, the interplay between ideas and evidence, a balance of “good enoughs”, avoiding leaps of faith, and attention to the relationships between variables. Striking a balance between a rock and a hard place is easy enough to prescribe, but elusive to achieve. Yet, like many skills in the policy arena, it is as much a habit of mind as it is a natural ability. Typical habits may resemble those needed for “bicycle repair” (fixing things, particularly things that have gone wrong in a predictable system) or those for an “endless game of Monopoly” (engaging in strategic interaction, particularly where the behaviour of others is unpredictable) (Stone 1997:259). Since I have assumed an ideal scenario, we can ignore instances of poor thinking habits (lazy, superficial and incompetent) for either bicycle repair or Monopoly. Some policy questions, of course, more closely resemble the extremes of bicycle repair or Monopoly. Yet, while pure types have their strengths, there is a happy medium where thinking is at the same time both grounded and interactive.

In New Zealand, the quality of thought exhibited by minds that are at the same time grounded and interactive is referred to collectively as *nous*. *Nous* means intelligence, common sense or gumption. It derives from the Greek association of the term with the highest sphere accessible to the human mind. Political or policy *nous* is savvy, but in relation to the context, the people and the issue, not necessarily in a self-serving way. So, too, is research *nous* savvy in relation to context, people and issues: it is policy aware.

REFERENCES

- Baehler, Karen (2002) “Intervention logic: A user’s guide” *Public Sector*, 25(3):14-20.
- Bardach, Eugene (2000) *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving*, Chatham House Publishers, New York and London.
- Blaikie, Norman (2000) *Designing Social Research*, Polity, Cambridge.
- Brown, Steven R. (1980) *Political Subjectivity: Applications of Q Methodology in Political Science*, Yale University Press, New Haven and London.
- Checkland, Peter (1999) *Systems Thinking, Systems Practice: Includes a 30-year Retrospective*, J. Wiley & Sons Ltd, Chichester.
- Department of Labour (no date) Human Capability Framework, www.dol.govt.nz/human-capability.asp.
- Greenberg, David, Marvin Mandell and Matthew Onstott (2000) “The dissemination and utilization of welfare-to-work experiments in state policymaking” *Journal of Policy Analysis and Management*, 19(3):367-382.
- Hakim, Catherine (2000) *Research Design: Successful Designs for Social and Economic Research* (2nd ed.), Routledge, London.
- Hirsch Jr., E.D. (2002) “Classroom research and cargo cults” *Policy Review*, 115 (Oct) www.policyreview.org/OCT02/hirsch_print.html.

- Hutchinson, Iris (2002) *The Craft of Policy Analysis in the Presence of Complexity*, Master of Public Policy research paper, Victoria University of Wellington.
- Majone, Giandomenico (1989) *Evidence, Argument, and Persuasion in the Policy Process*, Yale University Press, New Haven.
- Ministerial Advisory Committee on Complementary and Alternative Health, Ministry of Health (2003) *Complementary and Alternative Medicine: Current Policies and Policy Issues in New Zealand and Selected Countries*, discussion document, www.newhealth.govt.nz/maccah/htm.
- Ministry of Consumer Affairs (2003) *Creating Confident Consumers: The Role of the Ministry of Consumer Affairs in a Dynamic Economy*, Ministry of Consumer Affairs, Wellington.
- Morçöl, Göktug (2002) *A New Mind for Policy Analysis: Toward A Post-Newtonian and Postpositivist Epistemology and Methodology*, Praeger, Westport, Connecticut.
- NACEW (National Advisory Council on the Employment of Women) (1999) *Childcare, Families and Work: The New Zealand Childcare Survey 1998: A Survey of Early Childhood Education and Care Arrangements for Children*, www.nacew.govt.nz/fldPublications/labour_report.pdf.
- New Zealand Medicines and Medical Devices Safety Authority (Medsafe) (2003) *Submission to the Health Committee on its Inquiry into the Proposal to Establish a Trans-Tasman Agency to Regulate Therapeutic Products*, Wellington.
- Nutley, Sandra, Huw Davies and Isabel Walter (2003) "Evidence-based policy and practice: Cross-sector lessons from the UK" *Social Policy Journal of New Zealand*, 20:29-48.
- Parsons, Wayne (1995) *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*, Edward Elgar, Aldershot, United Kingdom.
- Pawson, Ray (2002) "Evidence and policy and naming and shaming" *Policy Sciences*, 23(3/4):211-230.
- Peace, Robin, Lynne Pere, Kate Marshall and Susan Kell (2002) "It's a Combination of Things": *Mental Health and Independent Housing Needs: Part 4, Group Interviews*, Ministry of Social Development, Wellington.
- Peirce, Charles Sanders (1998) *Collected Papers of Charles Sanders Peirce*, Charles Hartshorne, Paul Weiss and Arthur Burks (eds.), 8 vols., Harvard University Press, Cambridge, Massachusetts.
- Ragin, Charles (2000) *Fuzzy Set Social Science*, University of Chicago Press, Chicago and London.
- Schön, Donald A. (1983) *The Reflective Practitioner: How Professionals Think in Action*, Basic Books, New York.
- Shulock, Nancy (1999) "The paradox of policy analysis: If it is not used, why do we produce so much of it?" *Journal of Policy Analysis and Management*, 18(2):226-244.
- Statistics New Zealand (2003) *A Social Statistics Programme for New Zealand: Discussion Paper on Information Needs – Issues and Gaps*, Statistics New Zealand, Wellington.

- Stewart, Jenny and Russell Ayres (2001) "Systems theory and policy practice: An exploration" *Policy Sciences*, 34(1):79-94.
- Stone, Deborah (1997) *Policy Paradox: The Art of Political Decision Making*, W.W. Norton, New York.
- Streeten, Paul (2000) *What's Wrong with Contemporary Economics?*, www.vanzolini.org.br/seminariosp2000/paulstreeten.pdf.
- Weimer, David L. and Aidan R. Vining (1999) *Policy Analysis: Concepts and Practice* (3rd ed.), Prentice-Hall, Upper Saddle River, New Jersey.
- Wirth, Uwe (no date) *What is Abductive Reasoning?*, www.rz.uni-frankfurt.de/~wirth/index.html
- Wolf, Amanda (2002) "Diversity research: 'Medium-n' social science methodologies for policy analysis" in *Proceedings of the Sociological Association of Aotearoa New Zealand Conference*, 5–7 December, Christchurch.