

Making the most of research in policy making

NZCER Seminar: Making the most of research in policy making

3 August 2006

Cathy Wylie

Chief Researcher

NZCER

Downloaded from <http://www.nzcer.org.nz/pdfs/15031.pdf>

My first contract research project was one of the most satisfying I've been involved in. In many respects, it framed my expectations of how research could contribute to policy, and it contains within it some of the key principles about what makes for positive relations between contracted research and policy that I'll talk about today. So I want to start with an outline of that project, and the things that made it work well.

It was a study of the education and training experiences of women who were sole parents, in 1979. There was little known about sole parents; the DPB was still a relatively contested entitlement; and the people who asked me to do the work both believed in that entitlement but also saw its potential to sideline women from the world of work and identity outside the home, and leave them with little once their children had left home. The policy makers and I worked as a loose team from the start. I had just completed my PhD thesis in social anthropology, on the ways in which Pākehā in different social positions constructed a sense of individuality, and the differences between those senses of individuality. I knew how to interview people about their lives and the way they made meaning. I knew how to identify patterns in what was said, how to be systematic in interpretation.

What I knew about policy and the world of the public service was limited. But I was curious about it, and I respected the people I was working with. They believed deeply that they had a responsibility to make a difference—not that anyone used that phrase much then. They were seasoned veterans, who were open with their experiences, because they wanted me to understand their world. They also worked together honestly and openly: and this was a team that included local operations and national policy Department of Social Welfare staff, data management, and the Treasury officer whose patch included the department.

The word “scoping” wasn’t much used then either, but that’s what we did: we took the time needed at the start of the project to ensure we communicated, that the research questions were clear, and feasible, and that the project would provide the information and understanding they were seeking. They outlined their needs to me; I asked questions to check I understood. They gave me what background information they had; I went away and digested it and came back with more questions and ideas for the research. We discussed those, and there was a further iteration. The project emerged as a comparison of three groups: women who had never received the DPB; women who had had it and gone off it; and women who were still on it. Operational staff helped me find participants. I worked hard and fast—then, as now, the results were needed to feed into the policy cycle. We had another discussion mid-way through, when I talked about impressions, and asked further questions that would help me in both the remaining interviews and interpretations. I wrote a draft; the feedback was positive and also helped show what needed clarifying. We also discussed the implications of the research. The report was used to support the case for a training incentive allowance, and was also of use in showing both the common situations and issues for sole parents, and the diversity behind the stereotype. It remained for some time a reference in any other research and policy work in the area.

The contract document I had was, from memory, a couple of pages. It specified the agreed sum for the work, the deadlines, and the purpose of the work—and nothing else. It took five minutes of my time to read it, think about it, and sign it.

I’m fortunate that I had that early positive experience of a productive partnership between research and policy. I felt that I worked to my strengths, as did the policy people. I felt responsible for what I had undertaken, and I wanted it to be the best that I could do. Time mattered, but there was just enough of it to do my best. I felt that I could report what I found, and offer an interpretation. I enjoyed the dialogue, and the fact that we could challenge each other. I enjoyed thinking that, in a small way, I had made a contribution that policy makers could use to open doors for others.

I have had other positive experiences since of where the research I have done, or my team has done, has been made the most of by the policy makers who commissioned us. The common elements were:

- early clarity around the core purpose(s) of the research (or evaluation);
- an agreed methodology that was fit for purpose, budget, and time frame;
- seeing the work as a real partnership, rather than as a unidirectional contract;
- capability and continuity within the teams of each partner; and
- respect for the role of each other, and awareness that the roles are different.

There is a growing need for us to grapple with how to make the best use of research in policy making. From our perspective as a research organisation, we are concerned with recent approaches to contract research, particularly with our key client, the Ministry of Education. On

the one hand, the Ministry is indicating that it wants to work in partnership with research providers (and others in the sector); on the other, we are seeing policy makers undermine this aim by giving priority to risk management, leading to not only higher transaction costs but also the greater risk of not getting most value out of the research, and therefore wasting public money. The over-focus on (some kinds of) risk on the part of some policy makers ignores the fact that both researchers and policy makers share a key driver: the desire to improve something through the use of systematic analysis of information that takes account of the particular contexts within which policy operates on the ground. Both policy makers and research[ers] want to see value for money: something meaningful that increases understanding and helps decision making in a particular area.

So the purpose of this seminar is to provide policy makers with a seasoned contract researcher's perspective, emphasising the importance of approaching contracts with the aim of ensuring quality work within a real partnership frame. From our experience at NZCER, this is a much better way to manage the risks policy makers perceive. At the same time, it does not increase risks for the researchers, including the longer-term risk of undermining the research capacity and capability in New Zealand, as the recent trends could do.

When the OECD reviewed educational research and development policy in New Zealand in 2001, it said a quantum leap in investment in educational R&D was necessary, and it identified the need for more concentrated strategic or basic research, and increased spending on research, including research capacity and infrastructure. There has been some progress since then, but the base is still vulnerable, and not able to be taken for granted.

I'll take each of the five principles in turn, and describe a best-case scenario and a worst-case scenario in terms of achieving a positive role for contracting research in relation to policy making. These scenarios are drawn from actual experiences of NZCER researchers.

Clarity about the purpose of research

The best-case scenario

The person or people who want the work done have spent some time working out what it is that they want to know, and why. They really do want to know something that is not currently known or understood. That usually means that they are well informed (as a group), and know enough to know if there is further knowledge they need. They usually have a direct intended use for this knowledge, but they also think that individual pieces of research and evaluation should be contributing to broader or deeper understanding that can be used in future policy making as well.

They can articulate what it is that they need to know and understand, and why, and engage in discussion about their needs and purpose that is not defensive. On the one hand, they want the researchers to deeply understand the underlying issues and purpose—they want to know they can trust them with the work. On the other, if the researchers have some additional knowledge or insight, they are open to using that, and sometimes altering their initial purpose or scope. The

policy makers are open to the unexpected, and prepared to revise plans along the way if, in the course of the research, discoveries are made that strongly indicate that there would be greater value in changing direction or focus.

In the best-case scenario, these discussions occur very early in the piece, when the policy makers have decided on their preferred provider, and before the contract has been finalised. Because this discussion has occurred, it is easier for both parties to feel they are part of a shared enterprise, and to continue to act honestly and respectfully together. It means that the researchers start their work with a good grasp of the real issues for the policy makers, allowing them to design useful instruments (and more efficiently), and to bring that understanding to bear in the analysis and framing of results.

It also means that the contract—the legal document—does really reflect what has been agreed.

Some worst-case scenarios

1. A policy initiative has had an evaluation component attached to it because someone understands that this is “required by Treasury”, or believes that this will be more likely to get new funding approved. But the person or people involved do not actually see the need for evaluation, or want any new information, or people involved from outside. So they will go through the motions of designing a Request for Proposal (RFP), but not actually engage with the thinking needed to ensure that new understanding is gained. The resulting work is likely to be superficial, and a waste of public money.
2. Murkiness, disagreement, or a lack of shared understanding among the people involved in designing or approving an RFP, or misunderstandings that result in confused or unrealistic RFPs, or, worse, unwillingness or inability to engage with researchers who seek clarification. Some of these differences in expectations of a piece of research or evaluation may only surface once a report is written, when disappointment is expressed that cannot be resolved because the data that were needed were not collected; or else only resolved with re-analysis and re-writing that researchers may find themselves doing at their own expense.

Both these scenarios usually leave both policy makers and researchers with a sour taste in their mouths. These policy makers’ cynicism about research is usually confirmed. If researchers have stereotypes about policy makers as “bureaucrats” who are more interested in protecting patches or positions, or rolling on regardless of context because they are following a rule, then those stereotypes will be confirmed. Or, even more unfortunately, researchers will develop the stereotypes, and a protective layer of cynicism about not only the way policy is developed, but also its ability to do good. Both researchers and policy makers are left mistrustful of each other. If each side has been able to interpret the contract differently, and the policy maker has ended up taking a rigid reading, both will pay much more attention to contracts in future (but not the thinking, discussion, and agreement that underpin a good contractual relationship), and they will

approach every new contract as something to be designed for the worst possible scenario—the highest degree of risk. This creates its own unnecessary tensions and compliance costs.

An agreed methodology that is fit for purpose, budget, and time frame

The best-case scenario

The policy makers use good advice from a department's own research section—and government agencies should have one! Government agencies need a research section that people have to pay attention to—that is, there are clear department guidelines about contracting research, and the role of the research section includes early involvement. The principle of partnership applies here, too—within a department.

The policy makers specify what their research questions are, and what their budget is (knowledge from their research section about what size project they may be looking for and therefore what cost range it is likely to fall in is used). I believe that if this early estimation shows that the questions policy analysts have cannot be well answered within the money they can find for the research or evaluation, either the questions need to be scaled back to something feasible, or the project not embarked upon until further funding can be put together.

Good specification of research questions and estimation of the project size allows researchers to design something feasible, and not waste either their own or officials' time with designing something wonderful but not within budget.

The research questions derive from the careful discussion of the purpose of the research discussed earlier. Research methods that are most likely to provide the kind of data needed to answer the research questions may be mentioned, as may be forms of analysis, but researchers are invited to use their expertise in design.

The time frame allows sufficient time for:

- the clarification of core purpose at the start of the project (but this time will be saved on the contract negotiation);
- setting up the data collection—recruiting participants takes more time now with devolved public services, private providers, and the need for ethics approval and consents. Sometimes it is essential to spend time at the start of a project developing relationships before people will even consider participation. It is important that people trust researchers—not just so that they take part in the research willingly (and hopefully with some interest), but that they will then trust the research team in its analysis and reporting. It is also important to develop research instruments properly, so that the best data can be gathered, and analysis is not left short;
- the data collection (in which things rarely go like clockwork);

- analysis and writing which, in my experience, usually takes all the time allocated for it and more, since this again is not a mechanistic or clerical exercise, or one in which all the work or “deliverables” can be specified. It is not until you see the data as a whole that you can see how far analysis can go. Very often at this stage, experienced researchers then see the potential to dig deeper, and to provide a richer understanding in relation to the policy makers’ core questions.

It has to be said that our desire to understand and our excitement at what we see—the connections experienced researchers can make with other research, and their knowledge of the context—does mean we tend to over-deliver. If we are working to a contract in which we have to specify actual days, then this causes problems: do we stop with the number of days specified in the contract? Clearly, policy makers would prefer that we did not—but the cost is carried on our side.

Analysis and writing involves not just finding patterns and connections, but writing about them in ways that make sense to policy makers and the sector, that take account of different contexts; and

- discussing findings—both expected and unexpected—with policy people and incorporating feedback before finalising a report, or moving on to the next stage of a project. This is particularly important if a new direction becomes evident. I think, for example, of the powerful work on the use of assessment and the need to build professional understanding of both assessment and “what next” to meet children’s learning needs that originated from what was originally an evaluation of a schooling improvement cluster (Vivianne Robinson and Helen Timperley’s work in South Auckland).

Some worst-case scenarios

1. The RFP

- Presents a shopping list of inconsistent or vague research questions, which can be only partially or superficially answered if all are tackled. This shopping list approach is more likely if the policy makers commissioning the work come, as they increasingly do, from a range of areas within an organisation. It’s good to see much more cross-area policy work, but such work needs to be well led and coherent.
- Does not give a budget range or limit. Some policy makers do not understand that there is a difference between a tender for, say, x number of widgets, and asking someone to provide a customised design to match a brief which may be clear, but is still open-ended.
- Asks for outcomes that cannot be provided within the budget allocated. An experienced provider will know this better than one who is not. Nonetheless, it has to be said that even

experienced providers such as NZCER, which aims to provide something feasible for the budget, struggle to break even on many contracts. There is a continuing, and worsening, reluctance on the part of policy makers as a whole to pay what it costs to do good research.

- Having an unrealistic budget may mean that the most attractive proposal is also the most unrealistic, and the risk for policy makers is that the resulting research does not give them what they sought. Over time, seeing value for money simply as the cheapest bid could undermine the further development or maintenance of research expertise, since research organisations with a mix of experienced and newer researchers and established systems of quality assurance may be out-bid, and forced to either reduce the quality of their work, or cut back on their employment of experienced researchers.
- Having an unrealistic budget for the scope of research required may also mean that policy makers ignore their own systems of safeguards; for example, specifications that bidding organisations have quality assurance systems and ethics processes.
- Specifies methods that are inconsistent with the kind of material needed to answer the research questions, or the population of interest, or that cannot be undertaken within the budget or within the time frame.
- Specifies analysis that is inconsistent with the kind of data, research questions, or cannot be undertaken within the budget or time frame.
- Specifies too many milestone reports for the budget or size of work. This specification means either that the research work itself has to be cut back or, more usually, the researchers carry the additional cost, and end up out of pocket.

2. Setting up a project that is doomed to achieve at best partial success or use

- The policy people are well connected with what is happening “operationally”; there is good monitoring through various kinds of reporting, and thus policy people with operational people can create useful pictures of what is happening, and why, in order to improve an initiative as it rolls out. In such a case, there may be little that researchers can add that is timely. (And researchers don’t like feeling that their work has not been of some use—that they have wasted theirs’ and others’ time.)
- Setting up an evaluation as an afterthought. This means that no baseline or “before” data can be collected, and the researchers cannot work with those implementing an initiative or policy from the start to design instruments to capture the right kind of data, and so that these people will feel some ownership of the evaluation (and its use). A sense of ownership through this joint work also helps consistency of data across different sites.

- Wanting to know the impact of a new initiative when it is still in the setting-up phase. All that that kind of evaluation will often yield is the issues involved in early implementation, which are by now well known, and one would hope impressed upon newer policy makers by their more seasoned colleagues.
- Wanting to know the impact on, say, a major outcome like student achievement, when no consistent data (before or after) are available, or the cost of collecting such data or designing appropriate tests may be more than the budget allows.
- Wanting to know impact data on major outcomes too soon, before it is likely that things have shifted in the practices that are likely to result in changes in the major outcomes.

Some of these problems stem from political pressures: the desire to be able to report positive results within a political timeframe. Accepting that these pressures are unlikely to wane, it seems important that policy makers at least try to provide Ministers with more realistic measures of likely progress, and the reasons why they are realistic, so that Ministers feel they are on reasonably safe ground. The growing use of “theories of action” is therefore welcome, though—as always—as tools to be used to think with, and adapted to contexts, rather than new rulers used mechanistically.

3. Problems around analysis

We are seeing more interest now in more sophisticated analysis, which we welcome. But this is often not matched by the budget or time frame for the research, including the translation of that work into policy or lay language.

There are three main issues here:

- a) A desire for statistical “proof” of causality; either for a policy or initiative as a whole, or individual elements of it.

I can understand this desire in terms of the rational framework within which people try to make policy, and need to justify expenditure. Policy makers need to feel on sound ground in deciding where they are going to be able to get the “best bang for the buck”, to be able to compare, say, class size and laptops for every student and whole-school professional development side by side.

But the reality is that no single statistical study of policy impact or evaluation of an initiative is going to be able to provide definitive proof, or be able to definitively isolate the action of one aspect from all others. To go into all the reasons why would take another seminar, or two. I just want to mention a few reasons here.

- Our statistical methods are designed for an idealised world, and rely on some assumptions that cannot always be met, particularly in the social world (this includes the random controlled experiment).

- To statistically tease apart aspects that usually occur together is difficult to do unless you have large samples, and where we are looking at, say, the impact of schools on students, nested samples. They also need to include a large range of factors—since we know enough now to know we have to understand not just that something “works”, but *how* it works, and whether it works differently in different contexts, and if so, *why*. Very few New Zealand studies have such samples, or include the factors needed to develop the deeper understanding that policy makers (and others) want. Indeed, very few international studies are so comprehensive.

Policy makers seem to me to know this at one level; but not all accept it. This can mean that they dismiss analysis that is not of a certain kind, or that they give more weight to some forms of analysis than others, even if the measures used in that analysis are thin, or not well related to their intended subject. They can also dismiss consistent patterns coming from a number of studies using qualitative work or simpler (“descriptive”) statistics, because they do not match up to some ideal form of analysis. Yet the concept of consistency across studies is part of the bedrock of statistics.

- b) Not wanting to hear that something is not working as intended—not seeing the evaluation or research as the opportunity to fine-tune or change.
- c) Not wanting any reference to other studies or parallel endeavours—which limits the usefulness of the research or evaluation, particularly if there are findings or understandings emerging from the work that would be applicable elsewhere in the government agency, or increasingly, in cross-agency work.

Seeing the work as a partnership, rather than as a unidirectional contract

The best-case scenario

As an example, when the research team has developed draft instruments, the feedback they get is timely and informed, and the questions are improved as a result—more likely to make sense to the participants, and to cover anything that has changed in the policy settings between the proposal and the fieldwork. If any issues arise during fieldwork or analysis that indicate some change may be needed to what was initially agreed, there is thoughtful discussion of the implications of the change. In projects with several phases, each phase is mapped out in the initial stages of working together, but finalised closer to the time, so that the results of the previous stages can be fully utilised.

Policy makers treat the contract as a two-way commitment, and understand that their best leverage with the researchers is not the words in the contract, but the researchers' own drive to do good work—and therefore win more work in future. Experienced researchers know that their reputation depends on the quality of each piece of work.

When the project is finished, there is also thoughtful discussion of the implications of the work in relation to both the particular policy or aspect of work that lay behind it, as well as other related policy work or issues. The commissioners want the work used by others, so they broker opportunities with other policy makers, and support presentations to the sector.

Some worst-case scenarios

The client does not want to know that there are difficulties with the RFP or their approach. This may mean they choose proposals that will not work, leading to the end results of frustration and mistrust and poor use of public funds mentioned earlier.

It can also mean that they:

- accept feasible proposals only grudgingly;
- are haughty in discussions of what is to go into the contract;
- may use a contract that is designed for widgets or the purchase of a measurable output rather than research, and thus waste researchers' time in the initial discussions, as the researchers try to get something which is liveable and does not impose unnecessary compliance costs on them;
- insist on unnecessary milestones, read the contract literally, want everything planned and specified at the initial stages, and insist on sticking to that plan even if what emerges from the research or evaluation indicates that some other approach or focus may be more fruitful to follow;
- use the contract as a stick, will not enter into any joint problem solving if the need emerges, provide poor feedback which does not result in an improved report; and
- want to treat what results as their own private property, with little sharing within their own agency, let alone with the sector.

Despite this, it is possible for something worthwhile to be achieved, if the researchers involved are very good, and tenacious. However, their experience will leave the researchers unwilling to work with this person or group again. That might not matter if we had a surplus of good researchers—but we don't. In a small country, we need to use researchers wisely, and since we don't have much money for research, we need researchers who can move quickly—can use their accumulated experience of a given area as well as their research skills. We can't afford to squander a scarce resource because of poor work by policy makers.

Such situations are more likely where policy makers are not open to “bad news”, where they have narrow conceptions of their own work and others’, where they are highly “risk-averse” (and therefore see risk everywhere), and where they understand much less about research (and policy in the real world) than they think, but are not open to deepening their understanding.

Indeed, if you want a gauge of the quality of a policy maker, look at the way they work with researchers, the work that results, and what comes out of that work.

Problems around use of reports

NZCER has six reports that were accepted by various sections within the Ministry of Education in 2004 or 2005 and as far as we know from feedback, were found useful—but that are still waiting public release. We find this embarrassing to explain to the participants of the research (let alone journalists). Here is why we think timely release of research is important:

- It indicates that the Minister and government agency really do understand that in a decentralised public service, they need the sector to be well informed, since in the final analysis, it is the sector—at the individual school or class level—that will take the action that may improve a situation, and allow the Minister to make progress towards a government outcome. The Crown may own the copyright to a report, but it actually needs the information to be open.
- People get suspicious of reports that remain unreleased.
- People grow cynical about the value of research, and they become less willing to take part in research. It can be hard enough to encourage time-stretched people to participate in research—and we cannot afford to narrow willing participation.
- Others cannot build on the new knowledge that was created. This makes both research and policy making less efficient and effective. It may also mean wasting public funds, particularly if another policy maker commissions a study that covers much of the ground of the one that was not released.

In my ideal world, the contract between the policy client and researchers would include a time frame for public release of findings within two months of the final report acceptance. (Public release simply means to make available and notify the sector, and allow researchers to share findings with participants; media releases may or may not be relevant.)

Capability and continuity within the teams of each partner

Pretty self-explanatory at one level. But there are some issues here.

Research organisations need to be able to muster good teams—most primary research now is done in teams—and provide them with work that engages them. As I mentioned, good social researchers are still scarce in New Zealand. I’m not sure that we have the right policy settings to offer good career structures. At NZCER we count ourselves extremely fortunate to have a

purchase agreement that provides us with an opportunity to mentor and apprentice new researchers.

Good statisticians are even scarcer, and our need of them is ever increasing. But neither statisticians nor experienced social researchers are on the list of desirable occupations for would-be migrants to New Zealand.

We see an irony in the fact that in some contracts, we are now being asked to not only specify all the research team, but to get client acceptance of any changes we might make in that team. Ensuring we can deliver quality research seems to be our responsibility as an organisation—and how we do it should therefore be up to us. As noted earlier, it is deeply in our interest to always provide good quality.

The irony for us is that the new approach does not cover changes in client teams. And this is, for us, one of the big risk factors for bringing about good and useful research. In the good-case scenario, those leaving have enough time to hand over carefully, and to share discussions with their replacement and the research team leader so that there is shared understanding. But sometimes that has not been possible, or the replacement wants to set their own stamp on a policy or project.

An increasing risk factor for research as more policy makers commission their own research, and seek advice from their own research section too late, or not at all, is the variable understanding of how research works, and what can validly be expected of good researchers. Related to this is the uneven employment of experienced researchers within government agencies, and the variable use that is made of them. There needs to be good research and evaluation expertise within government agencies for them to commission good research—but their role needs to be clearer, and they need to have sufficient authority within their organisation so that their advice is taken.

Respect for the role of each other, and awareness that the roles are different

Neither researchers nor policy makers want to feel they have been told what to do, or scorned because their everyday world is different from the other's. I've shuddered to see policy makers' eyes glaze over when a researcher takes the moral high ground with them. I confess that I've been guilty of this myself. I've had the good fortune, however, to know what it feels like from the policy maker's perspective to be lectured as if something was simple to change, was up to the individual policy maker (and not the politicians, or a group of colleagues with their own internal quarrels and concerns, or a cross-agency group), or was **the** thing that really would make all the difference. I think researchers are gradually understanding that the work they do is part of a complex process and picture, and are having more realistic expectations of being able to make a difference, or see their work directly translated into specific changes. Though this expectation may have been fed by the public sector framework of judging the worth of things by their

immediate utility—FRST, for example, asked researchers to report what use had been made of their work.

I've also felt my own eyes glaze over when a policy maker has told me how to do research, or refused to engage in respectful discussion on any differences in understanding of the respective merits of different approaches to methodology and interpretation.

To put somewhat differently some earlier points, there is no point in commissioning research if policy makers do not want to use research expertise, or do not want an independent analysis and interpretation.

I think we do need more discussion of what we understand by “independence”. I don't interpret it as meaning that I can do my own thing, regardless of consequence, and that I can ignore valid feedback and comments from clients, or the peer review and critique I get within my own organisation and from other researchers.

But it does indeed mean to me that I want to tell it as I see it. That also means being as clear as I can about how I got to this point, and why it matters; and it means describing it in ways that I think might reach people, rather than alienate them. When I “speak truth to power” I want to be sure I have a valid picture of what truth might be (in a given set of circumstances), and I want power to listen. Research doesn't have a strong lobby group behind it: so each time we speak, we are reliant on making only those claims we can sustain with our collective work.

In conclusion

I hope I've drawn your attention to some fundamental issues around the provision of research that is useful in policy work. I've spoken about how we can improve the existing system of contracts that arise from an RFP. I hope that what I have shared today may be used to do that: to steer this framework away from such heightened sensitivity to risks that it is becoming more likely to end up with poorer quality research—and thus make policy makers even more risk averse because of bad experiences. Some of these issues are around how policy makers work, including how well they use their own organisation's research expertise.

We need to make this framework work to support partnership of the kind I have described here, make the most of our different expertise, and therefore more likely to yield real value for public money. And more likely to allow us to sustain, and grow, our research capability and capacity. But going further, we need to be thinking, as I know some policy makers are, about whether this contract framework is the best way to get useful research, more productive relationships, and ensure a sustainable pool of good quality researchers and statisticians.