

Mixed-method evaluation: a way of democratically engaging with difference¹

mixed-method

At this dawning of the new millennium, evaluation is widely practised in such important societal domains as economic development, education, environmental management, family welfare, and community well-being. But evaluation rarely fulfils its own potential to be a powerful contributor to the health and vitality of these and other critical domains of social life.

This, I argue, is because evaluation is commonly viewed as an activity practised at a distance - apart from the contentious fray of political decision-making about resources and priorities, in a space somehow shielded from special interests and advocacies, like a little air bubble safe from contamination by bias and disease alike (Mark, Henry & Julnes 2000).

Yet, as a long line of argument has demonstrated, values, ethics, and politics are intertwined with, rather than separable from, epistemology and knowledge generation (Schwandt 2002). Evaluative results cannot be generated from, nor represent a neutral position within, the politics of the contexts in which we work. So, like it or not, the practice of evaluation itself either sanctions and reinforces, or alternatively challenges and disrupts, key dimensions of these contexts, notably:

- who has the right to be heard about what
- what counts as legitimate knowledge
- how decisions are made – who participates, what happens publicly and what happens behind the scenes
- what factors or criteria are valued in making decisions, and who gets to determine these
- the ways that people relate in a given context – with trust or suspicion, respect or disregard, reciprocity or selfishness, caring or neglect.

Given the understanding that evaluation and its context are mutually constitutive – notably with respect to values, norms, knowledge legitimisation, and relationships of power and influence – the question becomes: what kinds of values, norms, knowledge claims and power relationships do we as evaluators wish to influence? What ends do we wish our evaluative practice to advance? Whose interests do we wish to serve?

Jennifer C Greene



Jennifer Greene is a Professor in the Department of Educational Psychology at the University of Illinois in Urbana-Champaign, Illinois, USA.

These questions point to the intersection of evaluation with democratic theory, for the ideals of democracy provide the most inspiring and inarguably the most defensible normative agenda for evaluation. Ideas about democracy have surfaced in multiple evaluation discourses in recent years, anchored in the historic work on democratic evaluation by Barry MacDonald and colleagues at CARE, including Saville Kushner – work that has also been advanced by a number of activist-scholars here in Australia, including Robin McTaggart, Stephen Kemmis, Yoland Wadsworth and Marie Brennan. In the US, Ernie House's work on deliberative democratic evaluation champions the democratic ideal of social justice via the principles of inclusiveness and rational deliberation.

Like these theorists, I wish to forge an evaluation practice that serves broad democratic ideals of participation, justice and equity by meaningfully engaging with people in their lived daily experiences. In my evaluative work, I am especially concerned with ensuring that the interests of all legitimate stakeholders are included, particularly those who are traditionally left out of the conversation. I am a supporter of rational deliberation, but a sceptic about the possibility or even desirability of consensus. Rather, I strive to find ways for diverse stakeholders to talk and engage in dialogue with one another toward greater mutual understanding, respect, tolerance, and acceptance.

I am not sure that evaluation, even with fully-fledged, formalised deliberative forums (see McNeil 2002) can ever settle important public debates. I am, however, committed to a vision of evaluation as one important site for engaging with the differences that are important in these debates – differences of perspective, experience, values and political ideology, and differences of privilege, power, prestige and possibility. I believe there is no more important political challenge in today's western democracies than finding ways to genuinely understand and respect the differences that currently besiege our everyday lives.

Mixed methods, diverse perspectives

One approach to democratically engaged evaluation is to adopt a mixed-method way of thinking. Mixing methods in evaluation connects to democratic evaluation through their shared valuing of diverse perspectives and multiple ways of knowing. Good mixed-method evaluation actively invites diverse ways of thinking and valuing to work in concert toward better understanding. In good mixed-method evaluation, difference is thus constitutive and generative (from Greene, Benjamin & Goodyear 2001).

Let me sketch three interrelated dimensions of evaluation practice and discuss how a mixed-method enactment of each invites democratic engagement with diversity:

- how we practise our craft or how we use our methodologies
- how we think about or theorise the phenomena we are studying
- how we position ourselves in our work.

My final point will address the issue of how we position our work in society.

Mixed-methods in practice

Today, evaluators of all stripes routinely use a variety of methods in the broad service of better understanding. Available evidence suggests we do so pragmatically, in response to the character of our evaluation charge and the demands of our context, rather than for philosophical reasons or on explicitly ideological grounds (see Greene & Caracelli 2002).

Even so, I contend that mixed-method evaluation practice inherently engages the challenges of human diversity and does so with profound respect for human difference – and thereby invites democratisation into the evaluation enterprise. Let me support this contention with a brief (hypothetical) example.

BECOME A MEMBER TODAY!

▶ **and enjoy all the benefits ...**

Membership forms are available from the website
<www.aes.asn.au> or refer to page 44 of this publication.

AUSTRALASIAN
EVALUATION
SOCIETY INC.

A community development evaluation

A community development organisation is partnering a local city government to sponsor a series of sports leagues for young teenagers in the area. The idea is to provide an attractive and healthy alternative for teenagers' time, and to lessen the allure of gang membership and other anti-social activities. Coaches and officials for these teen sports teams are being specially recruited from groups often victimised by teenage mischief and crime – ethnic shop owners, elders/retirees, residents of poor neighbourhoods.

A local foundation is supporting this initiative by funding an evaluation of it. The foundation is particularly interested in the meaningfulness of this program experience for its various participants – youth and adults alike. A perfect opportunity for a qualitative case study evaluation!

But the savvy and creative evaluation team, realising the limits of any single methodology, opts to try to locate their case studies in a broader context of city demography. They seek to assess program meaningfulness not only on a case-by-case level, but also in terms of neighbourhood or community characteristics. They plan to do so by using census data, survey data (if available), or perhaps even spatially via GIS. So the evaluation will include selected qualitative case stories, but placed in the context of more standardised or spatialised descriptions of the character of city neighbourhoods.

We could reasonably assume that the evaluators fashioned this mixed-method design primarily for practical reasons – to enable a better, more comprehensive understanding of the phenomenon of program participation. Even so, this mix of methods offers opportunities to engage more meaningfully with issues of diversity than a single-method study.

In addition to the developmental and identity characteristics likely to be invoked in the individual case studies, different dimensions of diversity may well be engaged with this design – neighbourhood characteristics such as ethnic/racial homogeneity, transience, sense of community, historical continuities, belongingness, even intensity of social capital.

Moreover, all dimensions of diversity generated in this mixed-method evaluation are viewed as potentially important and are valued as relevant to the people who named them, for this is the mixed-method way of thinking. A wider variety of methods is likely to engage a more diverse set of differences. And democratic values of equity and justice are potentially served.

Mixing theories, concepts and ways of thinking and knowing

Much of my own mixed-method work has advanced the value of mixing methods intentionally and explicitly at the level of knowledge paradigms, or how we think about the phenomena we are endeavouring to understand. That is, I believe we can do better work if we mix our methods with thoughtful attention to different ways of knowing, not just for practical reasons.

My arguments here are (1) epistemological – that we can know something better if we bring multiple ways of knowing to bear on it – and (2) political – that all ways of knowing are partial, and hence multiple, diverse ways of knowing are to be valued and respected.

Let me illustrate this admittedly abstract argument with a fanciful example that uses competing contemporary ideas about social reality and what we can know of it.

Example using competing contemporary ideas about social reality

The anthropologists from Mars and Venus

Two anthropologists, one from Mars and one from Venus, independently came to Earth to evaluate various cultural practices.

The one from Mars had prepared for his trip by studying the Scientific Realists – notably Ray Pawson and Nick Tilley from England (1997) and my American colleagues, Mel Mark, Gary Henry, and George Julnes (2001). These theorists maintain that social reality does indeed exist independent of our thinking about it, but is much more layered and contextual than previously understood. Causes of human and social phenomena are neither simple nor easily discovered. Rather, such phenomena are assumed to be multiply determined by complex interactions of causal factors that vary across context, space and time and that may well be deeply, almost impenetrably layered. Even so, our job as scientists remains one of figuring out, as best we can, these causal, contextual complexities that are social reality.

The anthropologist from Venus had prepared for her fieldwork by studying the Social Constructionists, including Bob Stake's responsive evaluation and the ideas of Egon Guba and Yvonna Lincoln (1989). In contrast to realists, these theorists maintain that much of what is important in social reality – namely, meanings and motivations for action – is socially constructed. The demographic markers of race, gender and class, for example, do not exist out in the world; rather, they are cultural constructions – albeit with real-world power and consequences (because we act as if race exists). For social constructionists, what is important to know, therefore, are the various meanings that people construct about their experiences and the processes by which these meanings are developed.

Now, it turns out that both of our alien anthropologists ended up studying at the same time a particularly unusual cultural activity, practised by some natives of the State of New South Wales on the continent of Australia. The natives called it 'cricket' – for like all anthropologists in the known universe, our two worked hard to use and understand indigenous languages in their work. Here are excerpts from each anthropologist's field notes.

The Martian anthropologist

Descriptive observational notes

The striker stood with his bat held down in front of his wicket. He tapped the bat on the pitch three times. The non-striker stood behind the other return crease. The bowler took a run-up from behind the non-striker's wicket. He passed to one side of the wicket, and when he reached the non-striker's popping crease he bowled the ball towards the striker, bouncing the ball once on the pitch. The striker hit the ball with his bat. It went 40 metres to the off side. The striker and the non-striker ran to each other's popping creases twice. Meanwhile a fieldsman ran and picked up the hit ball and threw it back to the pitch. The batsmen ceased to run. This sequence was repeated as the bowler delivered the ball again to the striker. The striker hit the ball, about 15 metres to the leg side. The batsmen did not run.

Interpretative thoughts

This is indeed a complex reality. It remains unclear just what causes what in cricket. Clearly, the ball has to be bowled before it is hit, and it has to be hit before it is fielded, but what causes the batsmen to run? Perhaps it's the angle at which the ball is hit or the distance or the speed. More logically, it is some complex interaction of these and other factors, like the weather or the hats that are worn by some of the players. Underlying mediating causal mechanisms may well include acquired skills and confidence in running. I shall have to observe more closely.

The Venusian anthropologist

Excerpts from an interview with a cricket batsman

That was an amazing over, wasn't it? Eighteen runs – my best ever! And that ball on the full over the boundary! Wasn't that great! Especially coming off that marvellous googly. And then there was that glance right through the legs of the square leg! I almost fell over laughing at that one! No throwing my wicket away this time! Here's hoping I can do as well again before stumps.

Interpretative thoughts

Huh?

Clearly this interviewee is enthusiastic about something related to cricket. It is not exactly clear what he is enthusiastic about. I believe it has something to do with his constructed meanings of boundaries and glances. These seemed especially meaningful to him. But, his laughter – notably about the legs of the square leg – may also signal other dimensions of significant idiosyncratic constructions and understandings. I shall have to listen more carefully.

The point is that these two anthropologists can readily complement and enhance each other's work, and thereby the resultant depth of understanding reached. If Drs Venus and Mars could accept and appreciate each other's way of conceptualising and making sense of human activities, they would jointly do much better than each alone. Mars may advance his causal understanding of running in cricket by asking the runners themselves – who are likely to embed important contextualities in their causal explanations. And Venus may more successfully interpret her interviewees' enthusiasms with a more standardised and representative portrayal of the game. And in the process of sharing and engaging in dialogue, one with the other, democratising respect for diverse ways of knowing could be enacted, along with the concomitant values of equity and reciprocity – not to mention interplanetary harmony.

Being a mixed-method evaluator

Mixed-method evaluations can take many forms, and many different kinds of evaluators can ply a mixed-method trade. But there are likely to be certain practices, certain ways of being an evaluator, that underlie good mixed-method practice. Figuring out what just these practices are is an important task for future thought and study. To inaugurate this line of work, I have one fine candidate to bring forward this morning – and that is the practice of listening, of listening well.

I was struck once again by the connections between the idea of listening well and conducting good mixed-method evaluative work when I read a journal article recently featuring some of the ideas of Hannah Arendt (Coulter & Wiens 2002). Arendt was an important 20th century Jewish philosopher, who fled Germany in 1933 and went first to France and then to the US. Her post-war work can be viewed as an effort to understand how and why the Holocaust could have happened. In particular, she pursued links between thinking, judging, and acting – trying to understand how really smart people, really good thinkers (such as her colleague and lover Heidegger) could be such bad judges and do such awful things.

In this work, Arendt attended specifically to interconnections between good judgement and wise action, toward developing a portrait of the 'good judging actor'. I think evaluators could and should aspire to be 'good judging actors'. Arendt first articulated her views of what constitutes good judgement:

One judges as a member of a community ... Good judgment ... is not a matter of objective knowledge or of subjective opinion, but a result of intersubjectivity; becoming a good judge depends largely on one's capacity to consider other viewpoints of the same experience, to look upon the same world from another's standpoint, to see the same in very different and frequently opposing aspects. (Arendt 1968, p. 51, cited in Coulter & Wiens 2002, p. 17)

For Arendt, a respect for plurality is thus a necessary attribute of a good judging actor. Arendt rejects both pluralism-without-judgement, (unfettered diversity), and judgement-without-pluralism, (judgements made outside the web of human diversity) (Biesta 2001). Equally important for Arendt is a respect for the uniqueness and agency of diverse others. So, to develop an appreciation for and understanding of the viewpoints of others, to develop good judgement, Arendt advocates 'visiting', which involves carefully listening to the perspectives of others because:

... the more people's standpoints I have present in my mind while I am pondering a given issue... the better I can [judge] (Arendt 1968, p. 241).

[So] actors must go into a plural public world and engage with others ... Arendt's is not an abstract public sphere, but a world of diverse and unique individuals, all capable of agency ... Respecting diverse standpoints requires dialogue with other people, listening to their stories, and relating to their uniqueness without collapsing these divergent views into a generalised amalgam (Coulter & Wiens 2002, p. 18) without essentialising them, and without losing your own unique standpoint:

Visiting is therefore not to see through the eyes of someone else, but to see with your own eyes from a position that is not your own ... in a story very different from [your] own. (Biesta 2001, p. 398).

What a wonderful vision for the practice of mixed-method inquiry – going out into the public world, visiting diverse others, listening well to each of them – and thereby being enabled to make good judgements in the service of doing wise actions.

Evaluation as a public craft

Making good judgements in the service of doing wise actions, in the service of contributing to wise practice – that is, to social, educational, environmental and development programs that are strong and successful; that enable participants to have healthier, more productive, more meaningful lives; that help forge networks of compassion and trust; that help build caring and vital communities – doesn't this sound like a worthy vision for evaluation? Obviously, I think it does and I believe that the points I have made so far are consonant with this vision.

In lofty terms, I am talking about positioning evaluation as a public craft and evaluators as stewards of the public good. A public craft, according to Harry Boyte of the Centre for Democracy and Citizenship, is:

... work that is undertaken for public purposes and in public ways ... [it is work that] adds public judgment or wisdom to knowledge ... [work that invokes the technical canons of science and is] attentive to the local setting ... and to the civic implications of practices. (2000, p. 1).

Boyte makes the same argument as I have regarding the importance of conducting professions, like evaluation, in democratising ways and regarding the value of a mixed-method way of thinking for this purpose:

[Traditional scientific] models of knowledge ... emphasise the detached, rational, analytic observer as the highest judge of truth and the most effective problem solver. This approach is in conflict with communal common sense, folk traditions and appreciation for craft knowledge mediated through everyday life experience ... This is not necessarily to say that science is 'wrong,' and folkways are 'right' (for instance, in medicine, it is good to have a highly trained specialist to fix a broken limb). But it is to propose that there are different sorts of valuable knowledge in public life. A singular celebration of the scientifically educated expert as the actor or initiator in public affairs marginalises the amateur, while it produces mainly information and knowledge – neither wisdom nor public judgment. We need a very different and far more civic craft model of professional practice if we are to see any widespread democratic renewal in our time. (2000, p. 3)

So, the point is not to get rid of science, but rather to forge a partnership between science and citizenship, to think of evaluators as public citizens, as good judging actors, and evaluation as the advancement of wise action.

In more concrete and practical terms, let me share one last vignette, this one based on recent experience, an experience integrally wrapped up in the tangles of the New Public Management.

The actual workshop was much less dialogic and much more focused on outcomes and their measurement than imagined in this vision. But the vision remained as an inspiration, and the actual workshop was successful in assertively connecting outcomes and performance measurements to the substantial challenges and vital importance of wise practice, of engaging with diversity and thinking multiplicatively, of embracing difference, of being a good judging actor.

Evaluation in and for the community

Seated around the comfortable table were representatives from four local funders and myself. The funders included city and county governments, as well as a local chapter of the United Way (an umbrella non-profit organisation that raises local funds and then redistributes them to community social service agencies). The purpose of the meeting was to plan a workshop on outcomes and performance measurement for staff in local agencies that receive grants from these funders.

The funders were hoping that the workshop would enable local agencies to begin to move beyond their customary focus on outputs, or what activities they conducted and how many people came, to the identification of key outcomes, or what difference their services have made in the lives of participants. When I asked the funders, 'Why are you so interested in outcomes and performance measures?' they replied, 'because our funders are demanding this kind of information. We now need to report on outcomes to our funders'.

Sigh. I believe this scenario has been repeated thousands of times in meeting rooms all over the western world during this past decade. Many of you have probably attended such meetings.

My challenge at this meeting and the subsequent workshop I conducted was not to belittle or ignore these accountability pressures. While I am far from enamoured with most performance measurement systems, I agree with many that good information on the outcomes of public services is long overdue. And after all, in embracing diversity and difference, one cannot be selective. No, instead my challenge was to reconnect evaluation – in this context, in its incarnation as performance measurement – with wise practice, to reposition evaluation as not only a management tool but also a meaningful engagement with diversity and difference, a time for visiting others and listening well, a democratising advancement of the public good, a public craft conducted by good judging actors.

The workshop that I co-led aspired to the following vision. In this vision of a performance measurement workshop, participants would discuss together the following simple but profound questions:

- What is the purpose of each of our various community, social, and educational programs? Why are we doing them? What are we hoping to accomplish?
Here, we would create a vision of the kind of community we wished to live in, and we would acknowledge that program work is extremely challenging and complex.
- How would each of us know (or how do we know) if we are reaching these ambitions, if we are creating the kind of community we aspire to, if each of our programs is a good one?
Here, we would talk about the importance of reflective critique on practice and the valuable contributions that evaluative data can make to such reflections and critique. We would also talk about how difficult this is to do well.
- Let us each pick one important program outcome and talk about ways that we could come closer to understanding how well it is being fulfilled by our program and what program changes might enable better fulfilment and thus a stronger, healthier community.

Here, we would talk about holding group interviews or local forums in which community people could share their ideas about the meanings of positive youth development, substance abuse rehabilitation, social capital networks, environmental sustainability, and so forth – and we would talk about how such diverse views could be incorporated into revisions to wise practice. We would also talk about existing measures, social indicators, or even brief surveys that could chart progress on important program and community outcomes. And we would talk about how these various kinds of information could be reported to funders, yes, but also and importantly, used to get closer to our collective community vision of democratic health and vitality.

And that, I believe, is what each of us as evaluators can do to connect our public craft with stewardship of the public good. May we all become good listeners.

Note

- 1 This article is an edited version of Jennifer Greene's keynote address at the 2002 AES conference in Wollongong.

References

- Arendt, H 1968, *Between past and future: Eight exercises in political thought*, Penguin Books, New York.
- Biesta, GJJ 2001, 'How difficult should education be?', *Educational theory*, vol. 51, no. 4, pp. 385–4000.
- Boyte, HC 2000, 'Professions as public crafts', background paper prepared for the working conference on New Information Commons, Racine, Wisconsin.
- Coulter, D & Wiens, JR 2002, 'Educational judgment: Linking the actor and the spectator', *Educational Researcher*, vol. 31, no. 4, pp. 15–25.
- Greene, JC, Benjamin, L & Goodyear, L 2001, 'The merits of mixing methods in evaluation', *Evaluation*, vol. 7, no. 1, pp. 25–44.
- Greene, JC & Caracelli, VJ 2002, 'Making paradigmatic sense of mixed method practice', in Tashakkori, A & Teddlie, C (eds), *Handbook of mixed methods in social behavioural research*, Sage, Thousand Oaks, California.
- Guba, EG & Lincoln, YS 1989, *Fourth generation evaluation*, Sage, Thousand Oaks, California.
- Mark, MM, Henry, GT & Julnes, G 2000, *Evaluation: an integrated framework for understanding, guiding, and improving policies and programs*, Jossey-Bass, San Francisco.
- McNeil, C 2002, 'Evaluator as steward of citizen deliberation', *American Journal of Evaluation*, vol. 23, pp. 45–54.
- Pawson, R & Tilley, N. 1997, *Realistic evaluation*, Sage, London.
- Schwandt, TA 2002, *Evaluation practice reconsidered*, Peter Lang, New York.