Next generation evaluation: Shaping better futures in Aotearoa New Zealand

David Pritchard

Keynote Opening Speech

The convergence of evaluation and impact measurement

Annual Conference of the Aotearoa New Zealand Evaluation Association
Te Papa, Wellington, 19 July 2016

Kia ora.
Although I am standing here at an evaluation conference, about to talk about evaluation, I don’t think of myself solely, and possibly not even primarily, as an evaluator. Let me give you a sense of the journey that led me to being here.

I was born and bred in suburban London although I currently live and work in the United States. I started my working life in the early 1990s as a peace activist in Northern Ireland—trying to help reduce community divisions that had erupted into violence in the 1970s but had roots that went back through centuries of religious, political, and economic conflict. I ended up there because on completing university, not ready to enter the world of salaried employment,
I was accepted onto a one-year volunteer programme at a retreat centre called Corrymeela, on the beautiful coast of County Antrim in Northern Ireland. The Corrymeela community was set up to promote reconciliation and peace-building through the healing of social, religious, and political divisions in Northern Ireland.

Intrigued by how to best to bring about peace and reconciliation to Northern Ireland and other conflicted parts of the world I applied to and was accepted onto an MA in peace studies course at the University of Notre Dame in the midwest of the United States. After that I returned to Northern Ireland to apply what I had learnt. This period was a critical time because I began to ask some awkward questions: Are we making any difference? Are we doing good—how do we know—or are we wasting time and effort? Should we be doing different things to pursue the same goal? What struck me most was that there seemed to be little appetite for questioning from other people who were putting in their time and energy into this field.

By now I was engaged to an American I had met at the University of Notre Dame, and with these questions buzzing through my head I decided my peace studies degree was of limited value. What I really needed to do was find out how best to make organisations work effectively. That took me back to the United States, but this time to business school. By the way, this may have been the best thing I did for Northern Ireland. For 18 months later the political parties came together and signed the Good Friday accord, so ending the hostilities!

I spent the next 20 years learning about and testing different approaches to promoting the public good. I was an economist for the New York City government, a consultant for government departments in Washington, DC, an evaluator of government programmes in London, an adviser to charities in the United Kingdom on how to assess their effectiveness, and most recently, back in the United States again, a teacher and a consultant to a number of organisations. And
in all of these positions the common question I was addressing was: “is there a better way to do things that serve the public good?”

This is what brings me here, and my starting point for today’s talk. In the words of the renowned American evaluator, Michael Quinn Patton, I believe the value of evaluation is to help those who are authentically engaged in making the world a better place, and to distinguish them from the resource wasters, boastful charlatans, incompetent meddlers, and corrupt self-servers. However, in my 25 years I have found making this distinction—finding the effective change-makers and helping them—is not so easy to do. The main problem is these categories are not mutually exclusive. But keep this purpose in the back of your mind through this talk, as the question of how to support effective change-makers is at the core of this topic.

When I lived just outside London, on most days I would commute to work on a train, along with several hundred thousand others. And pretty much every day I would get this experience—any of you have been in London may have experienced it too. You are sitting in the carriage in your own little world, absorbed in your own thoughts, trying not to catch anyone’s eye or look too happy in case someone mistakes you for a tourist. And you notice out of the corner of your eye another train on the track right next to you, catching up to yours and slowly pulling alongside until it is running parallel at the same speed. And you take a quick glance at the people in the next train and have the odd experience of looking in a mirror and seeing an unfamiliar face. The people in this other train are doing exactly the same as the people in your carriage, they look suitably glum(!), reading the freebie newspapers or gazing out the window, or these days checking their email. But while these other people at that moment look just like you, you also know they have different lives with different jobs, families, and so on. So at the same time you think you know what lives these people lead, you are also aware that you do not really know.
Imagine we are on such a train: an evaluation train. And when we look out the window we see another train running alongside us on a parallel track at the same speed, an impact-measurement train. And we look across and think we know what the people in this train are doing because it looks familiar, but at the same time we are not so sure. Over the next 30 minutes or so we will look more closely into the window of the impact-measurement train to see if how similar, or how different, it is from the train we are on. But before we do that, I want to first start, like a good evaluator, with a simple theory.

The basic neoliberal economist’s view of the world is that entrepreneurs working in free and competitive markets produce products and services that people want while creating income and wealth. This makes people happy. But economists also recognise that the same competitive nature of even corruption-free markets has negative side-effects that diminish that happiness. Common goods, such as the climate, water, and ocean fish stocks, will get overused. A free market won’t produce enough collective or public goods, such as common defence, parks, and community cohesion. There will be over-consumption of goods such as cigarettes that have harmful effects on others (known as negative externalities). And, because the market is competitive, there will be losers, people who will end up with nothing or very little. All these factors create unhappiness, which gets even worse if the stronger actors in the market collude or engage in foul play. So in steps government to ensure fair play in the markets, provide public goods, mitigate negative externalities, and engage in various forms of redistribution to provide for a level of equality that corresponds to the views, more or less, of the median voter. But as governments aren’t perfect in how they perform their duties, and because some public-spirited people don’t think governments do enough, they set up, and support, charities or non-profits to address these market and government failures.
Most social and environmental evaluation covers the effect of government services and regulation, and the impact of charities, and is mostly funded by government or philanthropists. The private sector may participate in, or be subject to, such evaluations. But there is little evaluation of the social and environmental effects of the actions of private companies that they voluntarily conduct by themselves. Indeed, if consumers are not willing to pay for this, the dynamics of the market will prevent such evaluation taking place.

However, there is a growing part of the market—made up of a diverse group of socially responsible businesses, impact investors, philanthrocapitalists, and others—who are increasingly interested in knowing about the impact on the world of their activities, products, and services. For the purpose of this talk I am calling these assessment efforts social and environmental impact measurement, or impact measurement for short, and referring specifically to practices used in the private sector. This is a simplification, because the term impact measurement is also used in the non-profit sector too. Whether the differences between these efforts and what we generally count as “evaluation” are substantial enough to warrant such a distinction is a point for discussion. But I don’t think it is controversial to say that most private companies involved in this “impact measurement” do not think of what they are doing as evaluation.

Why might we be interested in what is going on in this other train? If, like Quinn Patton, we want to help those who are “authentically engaged in making the world a better place” we should be very interested. Figure 1 shows New Zealand’s gross national expenditure from 2001 to 2015 (how much national and local governments, households, and non-profit organisations spend each year). What I described earlier as being in the domain of social and environmental impact evaluation is largely reflected in the segments in the middle of the stack chart (private non-profit organisations serving households;
central government; local government). The segments at the bottom and top of the stack chart (private households; gross fixed capital formation) largely fall in the domain of impact measurement. This is the bulk of economic activity. This is why we should care about what goes on in the world of impact measurement if we want to help the change-makers in those segments create more positive, and fewer negative, effects of this spending.


How much of the private sector, and how much of this spending, is currently or potentially open to being nudged in this way? It is difficult, if not impossible, to give a clear and meaningful answer because we do not have good gauges of either the breadth of the socially and environmentally concerned private sector, or the depth of that concern. Rather, we have pockets of interest, grouped around certain
labels. To illustrate the variability and difficulty in making a sensible estimate, here are a few of these labels.

The Fairtrade movement which originated in the 1960s guarantees a minimum price to farmers who demonstrate adherence to a set of socially and environmentally responsible practices, thus helping them gain a competitive advantage from such adherence. The most recent data from the Fairtrade Foundation’s website claims by 2011 there were 1.2 million Fairtrade producers in 63 counties, and by the end of 2010 over 900 Fairtrade producer organisations worldwide. In 2010 consumers spent $US4.6 billion on Fairtrade products. While impressive, in 2010 in New Zealand alone total household expenditure was $80 billion in 2010. If we consider only coffee, in 2010 the volume of Fairtrade coffee was just about 1.25% of the world total coffee sold by volume. So this market is significant, but relatively small.

The new kids on the socially responsible block are impact investors. The name was coined in 2007 by a group of investors who wanted to promote the idea that institutional investors, such as foundations, pension funds, and banks, should favour investments that do good while also bringing in a positive financial return. Though not an entirely new idea, it has caught the imagination of the millennial generation and in particular the new-found wealth of Silicon Valley. The Global Impact Investing Network (GIIN), the largest and most influential network of such impact investors, includes such heavy hitters as Morgan Stanley, UBS, and Deutsche Bank as well as foundations such as the Rockefeller Foundation. GIIN’s leadership council represents 60 large investors from around the world that collectively have assets with an estimated value of $US11 trillion under their management. This is equal to about one-sixth of the value of the 60 major stock exchanges in the world. But of this $11 trillion, a mere $60 billion, less than 1 percent, is classified as impact investments. (I
cannot help but wonder whether the remaining 99% of investment has no social or environmental impact or maybe is even negative!)

Clearly the term impact investing is emotive rather than descriptive—and as an evaluator I find that rather frustrating—but I think it is great that there are people with wealth who are pushing the idea that investments should be tilted towards enterprises and innovations that promote the collective good.

But if we step out of these relatively small pockets of economic activity and consider what the giants of the corporate world think of social and environmental concerns we quickly come across the notion of corporate social responsibility (CSR)—the idea that companies have obligations to the communities in which they operate beyond just maximising profits. Through the lens of CSR the entire private economy—or maybe just the portion that faces end consumers—seems open to considering its social and environmental impact. The Reputation Institute’s Global CSR RepTrak is just one CSR index that illustrates the thorny definitional issues involved in trying to assess the size of the socially conscious market. So trying to get a handle on the size of this socially responsible market is like trying to nail jelly to the wall.

It is also important to note that the notion of responsible capitalism is not new. One pioneer in this field was Robert Owen, a textile entrepreneur who became owner of New Lanark Mills in Scotland in 1810. Owen promoted the radical concepts at the time of cooperatives, socialism, and the 10 hour, and then the 8 hour, working day. To show what a revolutionary he was, in 1825 Owen set up a Utopian society on the banks of the Wabash River in Indiana in the United States. Unfortunately, the scheme failed because, according to Owen’s son, the people involved in New Harmony were “a heterogeneous collection of radicals, enthusiastic devotees to principle, honest latitudinarians, and lazy theorists, with a sprinkling of unprincipled
sharpers thrown in”. That seems to describe several of the organisations that I have had the pleasure to work with over the years!

While not new, this interest in social and environmental concerns seems to me to have three important features. First, it is increasingly mainstream. Companies that you would not expect to express much interest, such as JP Morgan Chase, are getting on board. Unilever, a massive consumer multinational, is a pioneer in promoting sustainability. The last two decades have seen new social and environmental reporting standards and guidelines—the Global Reporting Initiative, the UN Global Compact on Corporate Social Responsibility, Integrated Reporting, and the Sustainability Accounting Standards Board—that mimic (or mirror in different ways) financial reporting standards. In countries such as the United States where management can get sued for taking actions that do not maximise shareholder profits, new corporate structures are being created to allow the pursuit of social and environmental goals, not just financial ones. Benefit Corps, or B Corps—f or-profit companies certified by the non-profit B Lab to meet standards of social and environmental performance, accountability, and transparency—now exist in 42 countries in the world. When I was at business school in the mid-1990s, Ben & Jerry’s ice cream was used as a rather unique and quaint case study in how a couple of hippies with some ideals, grit, and luck, could take on and beat and established food conglomerates. Today the top business schools around the world, including Massey University here in New Zealand, are setting up research centres to create and support whole cohorts of socially responsible entrepreneurs.

Secondly, not only is this interest in the social and environmental effects of business becoming mainstream, I believe it is here to stay. Why? Because the combination of forces that got us here are not changing soon. We face, as a world, existential threats that need to be addressed collectively. Through the ages different societies—including
those of indigenous people—have faced, and been hard hit by, existen-
tial threats. This is not new. But existential threats to the planet as
a whole, largely brought on by our avaricious consumption, are new,
and not going away soon. The threats of climate change, growing
economic inequalities, and depletion of the world’s resources, call for
a response on all fronts, including market economics.

This is because the market economy is the main game in town.
Despite its weaknesses, most people believe a free market is the best
economic system. Figure 2 shows the percentage of respondents in
selected countries who believe that “most people are better off in a
free market economy, even though some people are rich and some
are poor.” Out of the 42 countries surveyed, only in seven did the
majority disagree. Most of those have suffered from corruption or
recent economic collapses, namely Greece and Spain. Countries that
are newcomers to the free market, such as China and Russia, are big
supporters. But the underlying driver of this interest in social and
environmental concerns, at least in the United States, is the opinions
of millennials, those born in the 1980s or 1990s. In recent months
there has been some hand-wringing among certain commentators
in the United States about the apparent dismissal of capitalism by
young people. A few opinion polls suggested that young people were
losing faith in capitalism. In this Harvard Public Opinion Project
poll, only 19% of young people polled self-identified as capitalists,
compared with 16% of young people who self-identified as social-
ists. The percentage who support capitalism (42%) is still larger than
the percentage who support socialism (33%), but this still suggests a
generational disillusionment with the benefits of outright capitalism.
As noted above, the free market seems to be the only game in town. But the existential threats continue to grow despite the role of government and philanthropy in trying to stem the tide. So maybe the way to solve these threats is at the source—changing the way that the market operates. This combination of factors led Bill Gates, in 2008, to call for a kinder capitalism. Whether or not we are evaluators, this is something to be interested in.

But it is the third feature that is the one most relevant to everyone here. The private market’s interest in social and environmental concerns is mainstream and here to stay. But, of most relevance to everyone here, it is backed, at least in theory if not altogether in practice,
by a commitment to measure the concerns. Measurement is a much-talked about topic in this space, and in a world where so much data we need is at our fingertips, aspirations and expectations are high. So, when the United Kingdom held the presidency of the G8 in 2013, David Cameron promoted the concept of impact investing—or social impact investing as it is known in the United Kingdom—by setting up a Social Impact Investing Taskforce, headed up by the father of British venture capitalism Sir Ronnie Cohen. The task force included a working group tasked to address the knotty problem of measurement.

For the remaining minutes I want to take a closer look into the windows of this parallel train so you can get a sense of what some of these practices are, and how they are similar to, or differ from, the train we are all on. Keep in mind that I don’t claim to be showing you the full picture. Rather, these are my observations, as well as those of others, as we look across the tracks. I am keen to hear the observations of others of you who work in this space.

An initial observation is that people who self-identify as evaluators are not so visible in the world of impact measurement. I know a number, but not all, of the members of the Impact Measurement Working Group I referred to earlier. To the best of my knowledge, though I cannot be sure, none of these members primarily think of themselves as evaluators. That is not to say they are not qualified or experienced. The people I know in this working group are all good people with lots of relevant skills and experience to bring. But it is not an evaluator’s perspective.

Instead we find that people with expertise and experience in assessing financial returns are applying the tools and knowledge they have to the analysis of social returns. This observation is not a complaint, as there is much value in this perspective, but the voice of evaluation is small in this space by comparison. Indeed, a 2015 survey of
members of the GIIN showed that two-thirds of the investors who responded said they rely on members of their investment team for tracking progress against social and environmental goals (see Figure 3). Maybe these teams include evaluators and social researchers, but I suspect not.

Figure 3. Responsibility for social/environment performance management (n=146) (Source: JP Morgan and GIIN, 2015)

This emphasis on financial expertise means that impact measurement is a hybrid of accounting and finance, and evaluation. The former is made up of a well-developed infrastructure of IT systems, rules, processes, regulations, and so on. The latter, while also mature, is less structured, focused on methodologies and customisation, and largely unregulated. Impact measurement sits in the middle, trying to mirror the standardisation and systematisation of accounting and finance while recognising the need for customised approaches and flexibility practised in evaluation.

When considering approaches to measuring impact, it is important to differentiate between the perspectives and practices of investors—who are analogous to funders—and the investees, who are similar to grantees or providers. For example, Bridges Ventures, an established and well-regarded social-impact investor in the United Kingdom, has developed a system that is based on the risk and return model for managing a financial portfolio—where an investor will
tolerate a risky investment if it is expected to generate an above-market financial return—and applied this system to how it selects and tracks investments in its social-impact portfolio. The four points of the radar in Figure 6 are: 1) the degree to which the investment satisfies a set of environmental, social, and governance criteria; 2) the degree to which the investment is meeting target social and environmental outcomes; 3) the level of confidence that the investment provides additionality, or is actually making a difference; and 4) the degree of alignment between the investee’s goals and its own strategic goals. Each investment is given a score for each of these, and a score for the risk that the first score might be wrong.

![Figure 4. Social-impact portfolio radar (Source: Bridges Ventures, 2013)](image-url)
It is clear that the type of model in Figure 4 borrows more from portfolio-management theory than programme-management theory. It also relies heavily on judgement and experience rather than objective data. Indeed, I think the greater reliance on judgement compared with data is an important difference between evaluation and impact measurement.

But when we look at practices at the level of the enterprise we see things that are more familiar. For example, FRC is a social enterprise in Liverpool, England, that started out recycling used furniture to people in need, but has since grown. They use a handful of surveys and output data to collect a mix of subjective and objective data at various points along their logic model, in ways that might be similar to non-profit organisations or some government programmes. But it is important to note that there is a lot of diversity in the measurement practices that take place at the enterprise level and the levels of effort and rigour that are applied.

That may be why investors do seem concerned about the capacity of their investees to track the impact of their work. In a 2015 survey of impact investors conducted by GIIN and JP Morgan, nearly 40% of 134 investors surveyed believe that the lack of capacity to track impact at the investee level was a serious challenge, though they are not so worried about the lack of qualified staff or understanding of best practices (see Figure 5). One wonders whether investors could similarly invest more in the capacity of their investees to help them up their game, similar to how government and philanthropic funders support capacity-building efforts to improve the evaluation capabilities of their grantees. It may be that while investors see this as a significant challenge, it is not so significant to warrant putting money towards solving. Several experienced colleagues tell me that this is the view they find.
This also hints at the possible relaxed or resigned attitude towards rigour. Seventy percent of the respondents thought that the ability—or rather inability—to attribute impact was not a challenge, or only a moderate challenge. It is not that these investors have cracked this nut. No one would suggest that statistically robust approaches to attributing impact are widespread. The opposite is widely acknowledged. Instead I think this response reflects a view that attribution is a “nice to have”, not a “must have”. Maybe that is the right approach? That is not to say there is no concern about rigour and objectivity. As part of the effort to mirror the financial world’s ability to compare and manage different investment options, impact investors are keen on standardised tools and frameworks. There are many such frameworks, and new ones are being developed all the time.

There is a similar interest in having standardised metrics. GIIN responded to the demand for such metrics by developing the Impact Investing Reporting Standards, better known as IRIS. But this is a bit of a misnomer as these are not really standards. Rather, IRIS is a set of defined measures created for specific subsectors, such as microfinance, education, and health. So IRIS is better thought of as a catalogue of potential common outcome measures for different sectors.

Anyone who has been involved in trying to develop shared outcome measures across different organisations knows that while it sounds great in theory, in practice differences between types of
interventions, context, beneficiary groups, and so on, place severe constraints on the use of common measures. So it is with impact investing. The 2015 JP Morgan / GIIN study found that while many investors use IRIS metrics, or frameworks that are aligned with IRIS, many do not (see Figure 6).

This question also illustrates the frequency with which qualitative anecdotes are used to assess social and environmental performance.

**Convergence**

One evaluation tool that has been widely adopted by the private sector is the logic model. But it has been adopted with some subtle differences. Whereas evaluations will often focus on the outcome and impact end of the chain, private companies often focus on what they have most immediate control over, namely inputs and activities, or their supply chain and business processes, to make them better. They might look for measures of these activities and their outputs to use as “leading indicators” of positive social and environmental effects they expect will occur downstream. For example, increased school attendance might be a leading indicator for higher educational outcomes and a reduction in the risk of a student ending up in poverty in the future.

In fact, private enterprises will often think of their impact not as what happens at the *end* of the logic model or theory of change,
but often the effect of changes they make at any point in the logic model, such as reducing their carbon footprint in manufacturing, introducing social and environmental criteria to their supply chain, or donating to charitable causes. So one issue we all face is that the term impact has so many meanings that it is almost always necessary to start a conversation about impact measurement by first discussing which definition is being used.

One reason why people involved in measuring impact are reluctant to think of themselves as engaging in evaluation is that they think of the latter as retrospective rather than prospective. The business mindset is about the future, not the past, and accordingly practices and models and tools lean to timely, actionable indicators rather than measures of the past.

Part of managing impact is ensuring that any organisation that has a positive impact is financially sustainable. In the evaluation world, economic evaluation, and assessments of cost-effectiveness may be add-ons to impact evaluations or may not happen at all. Evaluators are rarely asked to look at the financial sustainability of the evaluand. By comparison, in the world of impact investing and social enterprises, financial viability either sits alongside assessment of social and environmental impact or takes centre stage. This is the idea behind integrated reporting, one of the new reporting standards I referred to earlier. The point is to show the links between the financial, social, and environmental returns, and impact of a company’s operations.

Looking through the last window of this train, my final observation is that, like the world of evaluation, the different players in the impact-measurement world recognise they are part of a diverse population that has different goals, contexts, and practices. I have discussed today general features of impact-measurement approaches and practices but I want to emphasise there is much variation among these.
This diversity is recognised and practices reflect it. Last year some of my former colleagues at New Philanthropy Capital started classifying the different models and approaches used by impact investors at different stages in their organisational life-cycle, differences in the weight investors give to achieving social and environmental goals as opposed to financial returns, the nature of the business model they are supporting, and the specific sector they operate in. My guess is that, as impact measurement matures, we will see common practices emerging in pockets of specific sectors and types of investors, rather than convergence to an overarching framework or set of practices.

Now the train has passed let me summarise what we have just seen. Currently, the impact-measurement ecosystem is fragmented and fluid, reflecting not just the underlying methodological and practical complexities of assessing the impact of a heterogeneous group of organisations, but also the fact that it is a rule-free zone where almost anything goes to fit the different motivations and priorities of those who are interested in, and paying for, the data. This set of actors often have a different mindset and background to evaluators, are more sensitive to the cost of rigorous information, and tolerate the risk of not always being right.

The question facing us, individually and collectively, is how to respond as we watch this train going forward. I see three possibilities: 1) worry about what they might be doing in case they make matters worse, rather than better; 2) ignore, as the private sector has different goals that makes engagement too difficult; or 3) embrace this as a new opportunity to spread the influence of evaluative thinking and, in Michael Quinn Patton’s words, support those who are authentically engaged in making the world a better place (Patton, 2008, p. xviii).

In case you are in any doubt, my preference is for the third of these options, so I will finish with some suggestions.
First, “lean in” and get engaged with the social enterprises, social-impact investors, and any company interested in improving the social and environmental effects of what they do. The assessment practices are made by the people working in the sector. I think it would be good for evaluators to have a greater voice.

You can do this here in New Zealand which has a number of social-innovation and social-enterprise organisations and initiatives.

Secondly, you can also engage in the discussion internationally by becoming part of IMPCON2016, a discussion on this very topic that the American Evaluation Association (AEA) and Social Value International, supported by the Rockefeller Foundation, are hosting a few days before the AEA conference. The aim is to have representatives from the evaluation community and the impact-measurement community meet and discuss points of common interest and identify how best to move the agenda forward. So if you were thinking of attending the AEA conference, come a couple of days early. And at this point I have to give a shout out to John Gargani, President of the AEA, who is providing leadership in this area and has done more than anyone to make this conference happen.

A third option is to join Social Value International, a body of professionals a variety of disciplines including evaluation, accounting, and finance, who are interested in improving how we incorporate social and environmental value into decision making at all levels of society.

There is a national affiliate of SVI here, Social Value Aotearoa, and I believe, from talking to Julian King and others here, that ANZEA may begin to link up with them in the future.

Here is a final quote from two people very involved in the process of assessing social and environmental impact.

Many impact assessment methodologies seek to create a single estimate that is both precise and accurate. We have found that most of these methodologies are too expensive or complex to scale across
a large number of projects, while lighter touch methodologies are often imprecise, inaccurate, or both. (Michael McCreless, director of impact and strategy at Root Capital and Brian Trelstad, Partner at Bridges Ventures, in McCreless and Trelstad, 2012, p. 21)

Are we in two trains running parallel, two trains heading in different directions? Or in one train heading in the same direction? I am not entirely sure. But I do know that there is still work to be done, and you have the skills and experience that are needed so I encourage all of you in ANZEA to reach out and get involved.

Thank you.

References


The author
David Pritchard’s career has centred on helping effective change makers increase their positive social impact. He has more than 25 years of experience working with organisations in the non-profit, public, and private sectors on assessing the social value they create and finding ways to improve it.

Email: dpritchard@david-pritchard.net