Evaluative Rubrics: a Method for Surfacing Values and Improving the Credibility of Evaluation

Julian King
Kinnect Group, Aotearoa New Zealand

Kate McKegg
Kinnect Group, Aotearoa New Zealand

Judy Oakden
Kinnect Group, Aotearoa New Zealand

Nan Wehipeihana
Kinnect Group, Aotearoa New Zealand

Background: The challenges of valuing in evaluation have been the subject of much debate; on what basis do we make judgments about performance, quality, and effectiveness? And according to whom? (Julnes, 2012b).

There are many ways identified in the literature for carrying out assisted valuation (Julnes, 2012c). One way of assisting the valuation process is the use of evaluative rubrics.

This practice-based article unpacks the learnings of a group of evaluators who have used evaluative rubrics to grapple with this challenge. Compared to their previous practice, evaluative rubrics have allowed them to surface and deal with values in a more transparent way. In their experience when evaluators and evaluation stakeholders get clearer about values, evaluative judgments become more credible and warrantable.

Purpose: Share practical lessons learned from working with rubrics.

Setting: Aotearoa (New Zealand).

Intervention: Not applicable.

Research Design: Not applicable.

Data Collection and Analysis: Not applicable.

Findings: They have found that while evaluative rubrics look beguilingly simple they are hard to do well. However, when done well, evaluative rubrics can substantially increase the use and credibility of evaluation.

Keywords: Rubrics; values; valuation; stakeholder; validity; credibility; utility

Introduction

In the evaluation community and literature the challenges of valuing–determining what matters[1]– are the subject of much discussion and in part stem from a growing spotlight on evidence-based and systematic approaches to assessing performance and value (Julnes, 2012c; King, 2010). As evaluation consultants one of the challenges that we face on a day-to-day basis is to ensure that the judgments that we make are 'warranted' (Fournier, 1995) i.e., acceptable and valid to our clients and the communities we work with. Being specific about what we know and how we know it requires us to become clearer about the nature of the evaluative judgments we make, the questions we ask, the evidence we select, and the manner in which we appraise and use it (Gough, 2007). Scriven's logic of valuing (Scriven, 2012; Scriven, 1995) has several stages and steps within in it, and the challenge is how to apply this logic in practice.

It is widely acknowledged that the methodological approaches to valuing are varied and not always well developed or applied (Julnes, 2012c). As human beings we are constantly making judgments about what matters and what's important to us in different contexts. We often do this intuitively. Even when we are being deliberate we don't always do it transparently.

We agree with Gluckman (cited in Hubbard, 2012)[2], that there is a role for science to provide evidence "as values-free as possible". But even strong evidence and good science - though crucial to decision making - is often not sufficient to make evaluative judgments. There are a number of other values that are incorporated into everyday evaluative decision-making.

As evaluation consultants, we are constantly working to high stakes in complex decision making environments. Often there are multiple stakeholders and competing values at stake. In our view, therefore, a good evaluation is one in which the evaluator intentionally surfaces these values and applies them in a transparent, systematic way that results in valid and credible judgments.

Over time we've grappled with various ways of being clear about the basis on which we make evaluative judgments and were using tools and techniques from areas such as social science, market research and management consulting. All of these tools offered partial solutions to the valuing conundrum. In hindsight we'd have to admit, we found ourselves producing evaluations that weren't explicitly evaluative. Occasionally this resulted in debates about credibility of the method. We now recognise the real issue was the lack of transparency about the basis on which evaluative judgments had been made.

Te Hokinga Mai (The Return Home)[3]

Te hokinga mai literally means the return home. The initial impetus for our use of evaluative rubrics came when Jane Davidson returned home to Aotearoa New Zealand. She shared her approach to valuing with the NZ evaluation community in a very practical 'nuts and bolts' way (Davidson, 2005). Jane's approach communicates the complex ideas in Scriven's logic of valuing in simple down to earth language. She provided us with a rational and tangible approach to integrating values more effectively into our evaluation practice. This has spawned a "rubric revolution" (Davidson, Wehipeihana, & McKegg, 2011) in Aotearoa New Zealand and evaluative rubrics are rapidly working their way through the evaluation community and importantly into key government[4] and community organisations.

Evaluative rubrics (referred to hereafter as rubrics) have allowed us, four independent evaluation consultants in *Aotearoa* (New Zealand), to discuss what matters with our clients and programme and service stakeholders in a more transparent way. They have helped us clarify and set out the basis on which judgments about performance, quality, usefulness and effectiveness

---

[1] More formally, Julnes defines valuing as 'the methods by which we assist our natural abilities to judge the value of alternatives' (Julnes G. , 2012, p. 4). Scriven maintains 'the values that make evaluations more than mere descriptions can come from a variety of sources' (Scriven M. , 1991, p. 378).

[2] Professor Sir Peter Gluckman, KNZM FRSNZ FmedSci FRS. Chief Science Advisor to the Prime Minister of New Zealand. His role is to advise the Prime Minister on matters of science policy and on specific matters related to science and to promote the public understanding of, and engagment with, science particularly with young people.

[3] Te hokinga mai literally means the return home. The term is used here to acknowledge the catalytic and ripple effect of Jane's contribution to the practice of evaluation in Aotearoa New Zealand, since returning home.

[4] For example, the New Zealand Qualifications Authority use rubrics as part of its external evaluation review and quality assurance of private training establishments, institutes of technology and polytechnics, w nanga, government training establishments and industry training organisations. The Ministry of Education are using rubrics as part of the measurable gains framework to evaluate 'Ka Hikitia - Managing for Success' the Maori Education Strategy.

are made. Rubrics have also helped us identify who should be part of the judgement making.

Reflecting on the use of rubrics some of our learnings are:

- When the basis for making evaluative judgments is articulated and agreed at the start, it greatly increases the likelihood that the findings will be owned.
- When different viewpoints and values are surfaced and acknowledged early on, common ground is identified, and differences can be accommodated or acknowledged.
- It is possible to transparently have values discussions between funders and communities, so that both sets of voices are validated - a shared understanding about what matters can be reached.
- When clients and stakeholders are involved in making judgments against predetermined criteria, it increases transparency about how evaluative judgments are made.
- The sense making happens right from day one. As the evidence layers and builds, it is possible to systematically make sense of many streams and lines of evidence, in a concise and cohesive way.
- When using rubrics, we have found it is possible to deal with the richness and complexity of the real world in a transparent, appropriate, nuanced way that integrates a range of diverse data.
- When rubrics are used as a framework, reporting can be more explicitly evaluative and often reports are more focused and concise.
- Clients often find this type of evaluation easy to use, credible and they are able to more quickly to apply the learning in their organisation.

Rubrics, while useful, are not a panacea and they are not as easy as they may appear. Like any skill, you can learn the theory, but it takes time and experience to become a skilled practitioner.

This article unpacks our learnings from working with rubrics over the last five years. First we define rubrics and then we discuss how they have contributed to different aspects of evaluation and then we reflect on the value of rubrics in our evaluation practice.

What are These 'Rubric' Things Anyway?

"Rubrics offer a process for making explicit the judgments in an evaluation and are used to judge the quality, the value, or the importance of the service provided" (Oakden, 2013, p. 5).

The key feature that sets evaluation apart from descriptive research is that evaluation requires us to ask evaluative questions about *how good* something is, and whether it is *good enough* (Davidson, 2005). Deciding on what basis we determine how good something is, is the crux of the evaluation endeavour (Fournier, 1995). One of the most important parts of our job as evaluators is to facilitate a shared understanding of the basis on which we make judgments.

One way of doing this is using rubrics. Rubrics give us a structure and a methodology for addressing those '*how good is...*' evaluation questions (Davidson, 2005; Oakden, 2013). We acknowledge evaluation also includes other types of questions, such as 'what were the barriers and enablers?' or 'what were the unintended outcomes?', and these aren't necessarily addressed using rubrics. Furthermore we acknowledge that rubrics aren't the only way of addressing the '*how good*' questions.

Rubrics can take many forms (McKegg & Oakden, 2009; Ministry of Education, 2012; New Zealand Qualifications Authority, 2009; Oakden, 2013; Oakden & McKegg, 2011; Oakden & Wehipeihana, 2009; Pipi, Kennedy, Paipa, Akroyd & King, 2012) but at its most simple, a rubric often looks like a table or matrix that describes different levels of performance. The table on the following page provides a generic example. The language about rubrics is messy and can be confusing because there are no commonly agreed terms. This can make explaining rubrics to others tricky.

Suffice to say, there are two core aspects to rubrics which collectively articulate the important values in a given context:

- Criteria that *define the things that matter and what good looks like* (sometimes called evaluative criteria, quality distinctions, merit criteria, dimensions of merit or even, believe it or not, indicators)
- Descriptors that articulate *what different levels of 'goodness' or performance look like* (e.g. excellent, good, poor etc) for each of the things that matter (sometimes called merit criteria) (Davidson, 2005; Popham, 2011).

Table 1
Example of a Generic Rubric

	Criteria for Rating Answers to Key Evaluation Questions
Excellent	Performance is clearly very strong or exemplary in relation to the question. Any gaps or weaknesses are not significant and are managed effectively.
Good	Performance is generally strong in relation to the question. No significant gaps or weaknesses, and less significant gaps or weaknesses are mostly managed effectively.
Adequate	Performance is inconsistent in relation to the question. Some gaps or weaknesses. Meets minimum expectations/requirements as far as can be determined.
Poor	Performance is unacceptably weak in relation to the question. Does not meet minimum expectations/requirements.
Insufficient Evidence	Evidence unavailable or of insufficient quality to determine performance.

Note: Adapted from NZQA (2009).

In What Ways do Rubrics Contribute to Evaluation?

We have found that a strength of the rubrics approach is that it can be used to facilitate clear thinking and strengthen engagement throughout all stages of an evaluation. Rubrics offer a road map for identifying and defining: the things that matter; important criteria to guide evaluative judgments; the credible evidence needed; the selection of methods; overall evaluation design; data analysis and synthesis and finally reporting (Davidson, Wehipeihana, & McKegg, 2011; Oakden, 2013). They keep stakeholders (and ourselves) focused on the things that matter (King, 2010).

As a tool, we have found rubrics flexible and adaptable and the art of working with rubrics is in tailoring the approach to the evaluation context (Oakden & McKegg, 2011). This flexibility is a key strength of the approach but also makes it difficult to generalise about how rubrics might be used. However, the following two diagrams serve to illustrate the application of rubrics throughout the

life cycle of evaluation projects we have undertaken. Figure 1 shows the use the use of rubrics during evaluation design and figure 2 in making evaluative judgments.

Rubrics in Evaluative Design

Figure 1 illustrates the use of rubrics during evaluation design. In the early phase of an evaluation project, rubric development brings stakeholders to the table to surface the range of values and reconcile these (possibly diverse) perspectives, together with more formally documented expectations–such as those set out in policy, strategy, service specifications, intervention logic, and/or existing evidence.

Through the process, a shared understanding of what matters is reached and is articulated in the rubric. In practical terms this often involves accommodating the views of others rather than necessarily reaching consensus. This stage is iterative and we may produce several drafts of the rubric with stakeholders to reach the point where we are collectively ready to proceed to the subsequent stages of the evaluation. We have found that taking the time to craft the rubric in collaboration with stakeholders represents an early investment that pays dividends throughout the remainder of the evaluation.

By participating in the rubric development process, stakeholders necessarily become deeply engaged in grappling with the diversity of things that matter in their particular context. By the time the rubric is completed, we find that stakeholders have invested themselves in the evaluation design to the extent that they clearly understand the basis upon which evaluative judgments will be made, can see their own values represented in the rubric, and typically have a greater sense of ownership in the evaluation.

The next challenge is to determine what evidence is needed and will be credible to answer the key evaluation questions using the evaluative criteria. This involves systematic analysis of the evaluative criteria and generally involves us asking the questions, "how will we know?", "what would be credible evidence?" and "what is feasible?" for each criterion and context.

In our experience, this process helps ensure our choice of evaluation methods is clearly aligned with the context and values embedded in the rubric. Clients have told us, as a result of being involved in the rubric development process, that they are more confident that the evaluation will be able to answer the really important questions in a credible way.
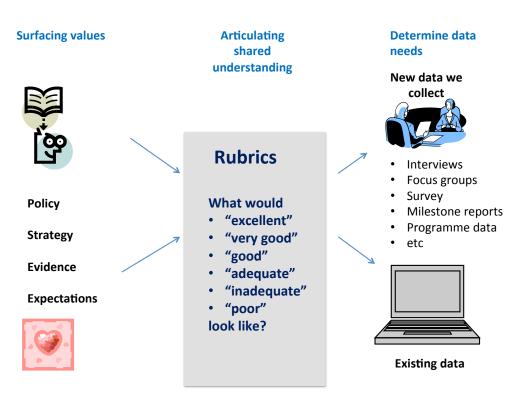
Figure 1. Using rubrics in evaluation design

Making Evaluative Judgments

It is in the synthesis of findings where our investment in rubric development really starts to pay off. As our evaluations typically involve multiple methods, there are usually several threads of analysis that may be undertaken in parallel (e.g., analysis of operational data, outcome data, interview transcripts and survey data).

The rubric has given us specific criteria, and the data collection tools were developed with reference to the rubric. Therefore the streams of analysis all map back to the rubric in a logical way. This enables us to reach sound evaluative conclusions both holistically and against individual criteria.

Importantly, we don't retreat to our 'evaluator cave' to make these judgments. Our preferred approach is to present the findings (*what's so*) and our emergent synthesis to stakeholders and facilitate a process of collective sense making. The purpose of this process is to reach a shared understanding about what the findings mean and their importance (*so what*) and what may need to happen as a result (*now what*).

Figure 2 illustrates the use of rubrics in making evaluative judgments. Throughout all the stages shown in Figures 1 and 2, the rubrics provide a focal point and a framework for the whole evaluation process. However, rubrics should not be the star of the show. What really counts to our stakeholders are the conversations that occur around the rubric, which help them to unpack the aspects of the evaluation that are their key focus. In our clients' eyes, the star of the show is a good evaluation. We like to think of rubrics as more like the best actor in a supporting role.
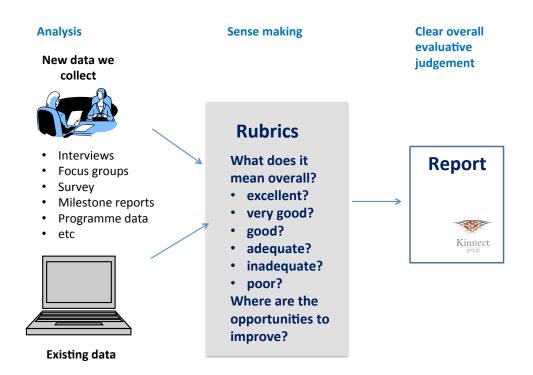
Analysis   **Sense making**   **Clear overall evaluative judgement**

New data we collect

- Interviews
- Focus groups
- Survey
- Milestone reports
- Programme data
- etc

Existing data

Rubrics

What does it mean overall?
- **excellent?**
- **very good?**
- **good?**
- **adequate?**
- **inadequate?**
- **poor?**

Where are the opportunities to improve?

Report

Kinnect
group

Figure 2. Using rubrics in making evaluative judgments

Sometimes, the rubric quite appropriately takes on its own prominence as a result of the buy-in that occurs through its shared development. For example, in a current evaluation, regional coordinators in charge of program implementation have of their own volition posted the rubric we created together on the wall next to their annual planners to keep them focused on the outcomes that matter.

What We have Learned About Working with Rubrics

Good evaluative criteria are specifically tailored to context. This means that the things that matter and therefore the evaluative criteria are different every time. At times, the evaluative criteria are used alongside a rubric which indicates generic levels of performance (McKegg & Oakden, 2009; Oakden & McKegg, 2011). At other times, fully developed rubrics are used which show different levels of performance for a project from poor to excellent (Oakden, 2013; Oakden & Wehipeihana, 2009; Wehipeihana, King, Spee, Paipa, & Smith, 2010).

There is no 'one right way'. We note that organisations in New Zealand such as NZQA use

generic rubrics, whilst the Ministry of Education has developed several specific rubrics which show performance at a range of different levels. An example of a fully developed rubric is the Measurable Gains Framework, which provides 'evidence of progress towards our objective of Māori enjoying and achieving education success as Māori' (Ministry of Education, 2012).

Just as no two rubrics are the same, nor are their development processes. However, what we think good rubrics have in common is that stakeholders can see their own values represented and validated within the rubric. Ideally, the rubric is not only resonant in terms of its detailed content, but also with regard to its overall look and feel, including the cultural cues embedded in its visual presentation.

The very endeavour and purpose of rubrics in evaluation is to assist in the reconciling and integrating of values and to provide a warrantable basis for evaluative judgments. Culture is embedded in human values, and therefore in order to do this in a valid way our practice of developing rubrics needs to be culturally responsive. In Aotearoa New Zealand, in recognition of our cultural context, our practice is nearly always collaborative and group-based rather than individual based (Julnes, 2012a).

The people who matter need to understand and buy in to the criteria. When we talk about *the people who matter*, we are not just referring to the funders and policy makers, but also those who are part of, or impacted upon, by the intervention and the evaluation (House & Howe, 1999; Mertens, 2009).

It follows that these people need to have some involvement in developing the criteria. How this is done is itself context dependent. In some cases it means getting the right people in the room and developing the criteria together. In other cases draft criteria can be developed from the literature or with experts in the field and then socialised with stakeholders. Usually reality is somewhere in between.

The criteria need to reflect those aspects of the real world that matter. They have to be consistent with evidence, policy and strategy – in ways that make sense to stakeholders.

In our experience, this range of externally and internally referenced perspectives need to be integrated. This can take some work and might culminate in reaching an accommodation rather than a consensus. We have found the time invested at the start of the evaluation, surfacing values, ironing out contradictions, managing tensions, and deciding on the boundaries for the evaluation, pays dividends throughout the rest of the evaluation because it brings a shared understanding and focus to those future conversations.

The following table summarises our key practice learnings.

Table 2
Working with Rubrics – A Summary of Our Learning

Why use rubrics?	*Rubrics sharpen evaluative practice and make evaluative judgments transparent and defensible.*
	The development of rubrics facilitates clarity and shared understanding about what's important and valuable in a given context. It keeps the evaluation focused on the things that matter and provides a road map for evaluative decision making throughout the process.
	In deciding about what to evaluate there is a fundamental need to set priorities. In the real world there are always limited budgets and timeframes. It's better to focus on the really important criteria than to skate across the surface. Don't over complicate it. Don't try to drink the ocean.
When to use rubrics	*The use of rubrics isn't an isolated or discrete step in an evaluation.*
	Rubrics provide a framework for evaluation design, data collection, analysis and sense making, synthesis and reporting.
	We suggest the development of the rubric happens early on in the evaluation and usually following confirmation of key evaluation questions and intervention logic. Rubrics are not set in stone. They can be changed, if necessary, as our collective understanding evolves – but not in order to avoid confronting findings.
Who to involve in developing rubrics	*Get the right people in the room.*
	The real magic of the rubric development process is being able to create a level of engagement and discussion about what matters. This leads to agreement or accommodation and ownership of the criteria and sets the 'tone' for the rest of the evaluation.
	It's important to capture the diversity of perspectives, motivations and values. Avoid creating a mutual admiration society. In our experience it often pays to have the fiercest critic in the room.

Table 2 Continued
Working with Rubrics – A Summary of Our Learning

How to determine the process for rubric development	*The process for developing rubrics depends on the context* The rubric development process can be highly participative and/or draw from literature and expert opinion. Our preferred approach is to combine these elements. While participation is important, we also look to draw on other sources such as past studies, expert opinion, professional standards, policy, strategies and service specifications whenever we can. Don't reinvent the wheel. The evidence about what's valuable about processes or outcomes is often out there. When working with stakeholders it's important to tailor the approach to resonate with the people in the room. Running a participative rubric development process requires skilled facilitation. The best laid plans about how to develop a rubric can go out the window in the first five minutes. Get good at taking the temperature of the room and adjusting the approach accordingly. Remember what you are doing is surfacing values about what's important. An initial activity that warms participants up to the concept of 'valuing' using our natural abilities (Julnes, 2012c) to make everyday decisions (e.g., buying a car, planning a party) can be a useful lead in to the process of developing a rubric.
What form should the rubric be presented in?	*How a rubric is presented matters a lot.* The most appropriate form of presentation depends on both the stakeholders and the context. It is important to be cognisant of the cultural. For instance, we have found the use of cultural concepts, metaphors and visual images can increase receptiveness and understanding of rubrics. The level of program maturity, and organisational maturity influences the types of rubrics used. For instance, the type of language that we use, and the way we discuss rubrics (and how we present them) when we work with government agencies, differs from our approach when working with community organisations. For community organisations we might focus on 'what you hear, what you see, what you feel' in developing up rubrics, whereas for government agencies we might focus on 'what the programme would look like if it were going really well' and 'what would constitute credible evidence of excellent performance.' Some rubrics are developed showing a range of performance levels with rich descriptions of the levels. Others are less detailed, and may even be evaluative criteria alongside a generic rubric.
Where do we use (and not use) rubrics?	*"Valuing is context dependent"* (Patton, 2012, p. 97) Although rubrics are highly effective and useful in many contexts, over-use or inappropriate use of rubrics can be damaging. There are contexts where rubrics have been embraced by whole organisations with much enthusiasm, but without an overarching coherent strategy for their use and with variable competency. Consequently, there has been an over-proliferation of what some have described as "rubric rabbits", leading to confusion over which rubric to use when, and scepticism about the value of rubrics per se. The question of where (and where not) to use rubrics is one that is ripe for further research. However, our reflections are that rubrics are best used in evaluation where: • People are willing to engage in evaluation and be reflective about performance (evaluation ready) • People are prepared to invest time and openly discuss and debate values • There are a diverse range of competing stakeholder priorities, perspectives and values. Even where these conditions do not exist, being involved in a rubric development process can help to shift mindsets and facilitate a more evaluation-ready environment.

Conclusion

In essence, rubrics:

- transparently set out the basis for making evaluative judgments,
- provide a means for reaching a shared understanding of what matters, and what 'good' looks like,
- help us to integrate expectations, policy, strategy and evidence about what good looks like,
- provide a basis for integrating multiple data sources to reach holistic evaluative judgments, and
- keep evaluation focused on the things that matter.

We first saw the potential of rubrics in a conceptual way. What attracted us to using rubrics was the prospect of a stronger base on which to make evaluative judgments. We rapidly discovered that rubrics also enabled us to work more effectively by more purposefully collecting and synthesising evidence.

We then found that rubrics didn't just benefit our evaluation practice. In fact, clients and communities embraced the use of rubrics and strongly expressed appreciation of the clarity of focus and purpose that helped them to move forward with confidence.

Rubrics don't have to be hard. But they can be done well or poorly. Done well, they are simple, cover all the bases, and are owned by the people who matter. Rubrics can evolve; they are just a guide and should not be used rigidly. In our experience rubrics can substantially increase the use and credibility of evaluation because they inform and support transparent judgments about merit, worth and significance. We believe rubrics make evaluation accessible and create demand for evaluative thinking well beyond the group of people who think of themselves as evaluators.

In hindsight, we realise that using rubrics has fundamentally changed our evaluation practice. Collaborative rubric development forms part of the glue for relationships. Rubric development opens up robust conversations about values. When we demonstrate that we understand those values – and they are reflected back in the fabric of the evaluation – relationships are cemented and the credibility of evaluation is profoundly enhanced.

References

Davidson, E. J. (2005). *Evaluation methodology basics The nuts and bolts of sound evaluation.* Thousand Oaks: Sage Publications.

Davidson, E. J., Wehipeihana, N., & McKegg, K. (2011, September 1). The rubric revolution. *Paper presenter at the Australasian Evaluation Society Conference.* Sydney, Australia.

Fournier, D. M. (1995). Establishing evaluative conclusions: A distinction between general and working logic. In D. M. Fournier (Ed.), *Reasoning in Evaluation: Inferential Links and Leaps. New Directions for Evaluation, (58)*, 15-32.

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*, 213-228.

House, E.R., & Howe, K. (1999). *Values in evaluation and social research.* Thousand Oaks: Sage.

Hubbard, A. (2012, December 1). Of a scientific persuasion. *Dominion Post*. Retrieved from http://www.stuff.co.nz/dominion-post/capital-life/8023559/Of-a-scientific-persuasion

Julnes, G. (2012a). Developing Policies to Support Valuing in the Public Interest. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, 133*, 109-129.

Julnes, G. (2012b). Editor's notes. *New Directions for Evaluation, 133*, 1-2.

Julnes, G. (2012c). Managing valuation. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, 133*, 3-15.

King, J. (2010). *We all rely on good evaluation.* Retrieved from Julian King & Associates Limited: http://www.julianking.co.nz/blog/good-evaluation/

McKegg, K., & Oakden, J. (2009). *Characteristics of good peer support.* Wellington: Wellink Trust.

Mertens, D. (2009). *Transformative research and evaluation.* New York: Guilford Press.

Ministry of Education. (2012, May 22). *Measuring and reporting progress.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/theMinistry/Poli

cyAndStrategy/KaHikitia/MeasuringandRepo
rtingProgress.aspx

New Zealand Qualifications Authority. (2009). *Policy and guidelines for the conduct of external evaluation and review.* Wellington, New Zealand: New Zealand Qualifications Authority.

Oakden, J. (2013). *Evaluation rubrics: How to ensure transparent and clear assessment that respects diverse lines of evidence.* Melbourne: Better Evaluation.

Oakden, J., & McKegg, K. (2011, August). *Waste minimisation act implementation: Evaluation of stakeholder perspectives.* Retrieved from Ministry for the Environment: http://www.mfe.govt.nz/publications/waste/ waste-act-stakeholder-perception-report-final/waste-act-stakeholder-perception-final.pdf

Oakden, J., & Wehipeihana, N. (2009). *FTP Induction programme review (unpublished report).* Wellington: Judy Oakden Consultancy.

Patton, M. Q. (2012). Contextual pragmatics of valuing. In G. Julnes (Ed.), *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, (133),* 97-108.

Pipi, K., Kennedy, V., Paipa, K., Akroyd, S. & King, J. (2012). *Evaluation of the Te Puni Kōkiri investment in the Māori Wardens Project 2007-2010.* Prepared for Te Puni Kōkiri. Auckland: FEM (2006) Ltd.

Popham, W. J. (2011). *The role of rubrics in testing and teaching mastering assessment: A self-service system for educators, Pamphlet 13 (Mastering Assessment Series).* Boston: Pearson.

Scriven, M. (1991). *Evaluation thesaurus.* Newbury Park, CA: Sage.

Scriven, M. (1995). The logic of evaluation and evaluation practice. In D. M. Fournier (Ed.), *Reasoning in evaluation: Inferential links and leaps. New Directors for Evaluation, (68),* 49-70.

Scriven, M. (2012). The logic of valuing. In G. Julnes (Ed.). *Promoting valuation in the public interest: Informing policies for judging value in evaluation. New Directions for Evaluation, (133),* 17-28.

Wehipeihana, N., King, J., Spee, K., Paipa, K., & Smith, R. (2010). *Evaluation of He Ara Tika.* Wellington, New Zealand: Ministry of Education.