



Actionable Self-Assessment and Evaluation for the Real World



Keynote address for:

NZQA Self-Assessment for Quality

Otago Polytechnic, Dunedin, NZ, December 2, 2010

E. Jane Davidson, Ph.D.

<http://RealEvaluation.com>



An 'evaluative' approach?

Evaluation	Research	Monitoring	Audit
Poses questions about: <ul style="list-style-type: none"> - quality (merit) - value (worth) - importance 	Poses questions about: <ul style="list-style-type: none"> - the state of the world - relationships between variables 	Seeks to: <ul style="list-style-type: none"> - describe - track (relative to plans or targets) 	Seeks to: <ul style="list-style-type: none"> - examine - verify - inspect
Is underpinned by the collection of data	Is underpinned by the collection of data	Is underpinned by the collection of data	Is underpinned by the collection of data
Purpose(s): <ul style="list-style-type: none"> - learning - development - accountability 	Purpose(s): <ul style="list-style-type: none"> - New knowledge - Explanation - Contribution to theory 	Purpose(s): <ul style="list-style-type: none"> - To describe - To track trends and/or progress 	Purpose(s): <ul style="list-style-type: none"> - To ensure compliance with pre-set standard(s) of practice or model
Product <ul style="list-style-type: none"> - Evaluative conclusions (that say something explicit about quality, value or importance) 	Product <ul style="list-style-type: none"> - Enhanced understandings about the state of the world and relationships between variables 	Product <ul style="list-style-type: none"> - Descriptions of measurable inputs, activities, outputs and sometimes outcomes 	Product <ul style="list-style-type: none"> - Confirmation of compliance with existing standards, rules and/or obligations
Has explicit professional quality standards	Has implicit quality standards	Does not have explicit quality standards	Has explicit professional quality standards

What makes for Real, Genuine, Actionable Self-Assessment & Eval?

What's in:

1. Right questions
2. Value-based
3. Credibly evidenced
4. Direct answers
5. Well communicated
6. Usable/actionable
7. Sincere intent to use
8. Value for money

What's out:

- ❑ Value-free studies
- ❑ Designs that don't answer the questions
- ❑ Objectives-based only
- ❑ Summaries of opinions
- ❑ Average effects only
- ❑ Misleadingly reported
- ❑ Buried or censored

See Jane Davidson & Patricia Rogers' blog:
<http://GenuineEvaluation.com>

Questions drive actionability

- ❑ The questions need to "get" the purpose and cover the big picture issues
- ❑ Evaluative (value-based) questions are needed, to yield more actionable answers:
 - What's so? → So what? → Now what?
- ❑ Need transparent ways of combining descriptive evidence with 'values'
- ❑ Need reporting that gets to the point!



Descriptive questions ask “What’s so?”

1. **What** was the learning experience like for students?
2. **How** did the students rate the course/programme?
3. **Has** there been a change in retention outcomes?
4. **What** were the characteristics of those who dropped out?
5. **How** have achievement rates changed since this program was implemented?
6. **Do** programme participants have higher skills than others who did not go through the programme?
7. **Were** the graduation targets met?
8. **What** were the employment outcomes?



Evaluative questions ask “So what?”

1. How **well designed** is the programme?
2. How **valuable** is the learning experience for students?
3. How **effective** is the teaching?
4. How **well** are students guided and supported?
5. How **substantial and valuable** were any learning gains resulting from the programme?
6. How **good** are the employment outcomes?
7. How **sustainable** is the programme in terms of its enrolments?
8. Is this programme **worth** the investment (of student fees and/or taxpayer money), given its outcomes?

The EER KEQs are evaluative questions

<i>Key Evaluation Question (KEQ)</i>	<i>What it covers</i>	
1. <i>How well do programmes and activities match the needs of learners & other stakeholders?</i>	Quality of content and design of TEO programmes and activities (including relevance to current needs & aspirations)	Process evaluation
2. <i>How effective is the teaching?</i>	Quality and effectiveness of instructional delivery	
3. <i>How well are learners guided and supported?</i>	Quality of support coverage and delivery	
4. <i>How well do learners achieve?</i>	Learning outcomes (knowledge, skills, confidence, attitudes, etc) – particularly those that occur within the bounds of the course or programme (or, by graduation)	Outcome evaluation
5. <i>What is the value of the outcomes for key stakeholders, including learners?</i>	Other outcomes (for learners, employers, communities, society, the labour market, the economy, the profession, etc.)	

7

But how do we answer evaluative questions?

There's ...

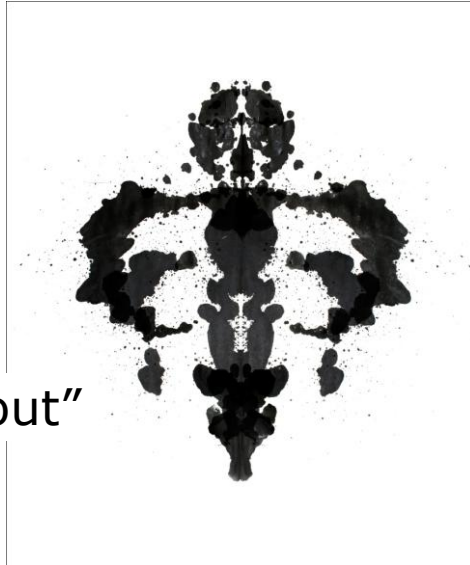
- How [many] self-assessments and evaluations "do" answer these ... (a quick review of 2 common species)
- And how we "could" or "should" answer them ...



Species 1. 'Value-free'



"You work it out"



Species 2. 'Reasoning-free'



*"I looked
upon it and
saw that it
was good"*

Leaping the critical 'values' step



Descriptive Evidence (what's so) + **VALUES** (definitions of 'quality' and 'value') = Evaluative Conclusions (so what)

11

... OR we can actually be systematic and transparent about it





7 steps to actionable answers

1. Clarify: outcome or process question?
2. Focus your question & its scope
3. Broadly define 'quality' or 'value' – and the main sources and types of evidence
4. Use/develop an evaluative rubric to define what "good" should look like in your context
5. Gather a mix of evidence and interpret relative to the rubric
6. Generate clear, direct, well-evidenced answers
7. Deliver succinct, user-friendly reporting



1. Clarify: Process or outcomes?

	Term	Definition (for TEOs)	Examples
Process	Inputs	People and things that are put into or accepted into the TEO	Staff and resources Learners
	Activities	Programmes, systems and services that the TEO delivers	Courses & programmes Support services
	Outputs	'Products' and trained learners that are <u>produced by</u> (or through) TEO programmes and activities	Graduates Course/unit credits Qualifications
Outcomes	Outcomes	Changes that <u>happen to</u> learners, employing organisations, communities and the economy that are at least partially <i>caused by</i> TEO programmes and activities	Knowledge and skills Attitudes and confidence Job placement & salaries Work performance Community contributions Impacts on the economy



2. Focus question and scope

- Start with the KEQ (broad evaluative question):
 e.g. KEQ #5 “*What is the value of the outcomes for key stakeholders, including learners?*”
- Flesh out with ‘evaluation indicator(s)’:
 e.g. “*Graduates gain employment, engage with further study and/or contribute positively to their local and wider communities.*”
- Rewrite this (or a specific part of it) as a more focused evaluation question:
 “*How good are the employment outcomes for this programme?*”



3. Define ‘quality’ or ‘value’

- Evaluation Question:
 “*How good are the employment outcomes for this programme?*”
- Key Considerations:
 - Quality of placements
 - Proportion placed soon after graduation
 - Success or impact in role
 - Demand for graduates
- Identify key sources & types of evidence:
 - nature of work
 - reputation of employer
 - starting salaries
 - gauges of success in business, voluntary community/iwi or creative enterprise
 - evidence of active recruitment or preference for graduates

4. Use rubrics to define 'good'

- Evaluative rubrics:
 - A broad-brush way of transparently defining what good, excellent (etc) quality, value or performance would look like in practice *and at different levels*
 - Allow interpretation of qualitative, quantitative and mixed method data
 - *as a set*

Evaluative rubrics may be generic ...

Rating	Performance Descriptors for Answering KEQs
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.

Source: NZQA's External Evaluation & Review framework

Or, rubrics may be very specific, e.g.



Māori learner attendance/retention, engagement

Rating	Description
Highly effective	The <u>vast majority</u> of Māori learners: <ul style="list-style-type: none"> • are motivated and positively engaged in learning and have very high attendance, retention and completion rates • have a strong sense of support for their inherent identity, language and culture in their educational institution • Actively participate in ways that are aligned with their whānau, cultural and individual values and worldviews
Minimally effective	<ul style="list-style-type: none"> • <u>Most</u> Māori learners have very few unexplained absences and want to be at their educational institution most of the time • There is clear evidence emerging that a <u>substantial number</u> of Māori learners are leading their own learning, maximising their own skills and talents, and staying at their educational institution long enough to allow them good opportunities relevant to their career aspirations
Poor or Detrimental	One or more of the following is evident – and meaningful, active steps are NOT being taken toward remedying them: <ul style="list-style-type: none"> • Unacceptable numbers of unexplained absences for Māori learners • Many Māori learners don't want to be at their educational institution

Source: MOE projects (various)

There are many wording options for ratings, performance or outcomes



Excellent	Highly Effective	Coach	Standing Tall
Very Good	Consolidating Effectiveness	Expert/ Specialist	Building on Strengths
Good	Developing Effectiveness	Competent Professional	Releasing/ Letting Go
Just Adequate	Minimally Effective	Emergent Professional	Ownership
Inadequate	Ineffective	Experienced Novice	Awareness
Poor	Detrimental	Fundamental Awareness	Crisis

5. Gather and interpret evidence

	e.g. when rating "Good"	you need to show ...
Key points from performance descriptors	Performance is <u>generally strong</u> in relation to the question.	Specifically, what evidence led you to believe performance was "generally strong" – as opposed to "clearly very strong or exemplary" (excellent) or "inconsistent" (adequate)? Include the most important examples of BOTH positive and negative evidence.
	No <u>significant</u> gaps or weaknesses,	What were the gaps or weaknesses, and why should they be considered "not significant"? Based on what?
	and less significant gaps or weaknesses are <u>mostly managed effectively</u> .	What, specifically, is the tertiary education organisation doing to manage gaps and weaknesses, and why do you consider this "effective management" in most or all instances?

6. Generate clear, well-evidenced answers

- *Question: "How good are the employment outcomes for this programme?"*
- Answer: Good, for the following reasons.
 - Generally strong quality, rates and speed of placement – based on [insert relevant comparisons, judgement criteria, reasoning – why not clearly very strong/exemplary? etc]
 - Main weaknesses are ... These are not serious because ... They are being managed in the following way ... This is an effective strategy because ...

Indicators, measures, data are not answers to evaluation questions!



Evaluation Question: Answer (NOT!):

Tell me, **how good** are the employment outcomes of our programme?

And the answer to my question is ...???

Let's see ... X% of program graduates employed within 3 months and 6 months; the mean incomes of graduates was \$Y,000, which was more than comparison Z ($p < .05$); and in interviews and focus groups with participants and employers people said ...

23

7. Deliver succinct, user-friendly reporting



- 4 "social scientist" habits for highly ineffective reporting:
 1. Pointless models
 2. Leaping to measurement
 3. Reporting results separately by data type or source
 4. Write-up like a Master's thesis
- Mix it up:
 - Small "real time" focused chunks – things you can use NOW
 - "Live" presentation & discussion of findings instead of long reports
 - Facilitate a group through evidence interpretation & action planning
 - Generate next questions together

Structure reporting so that it *demand*s actionable answers



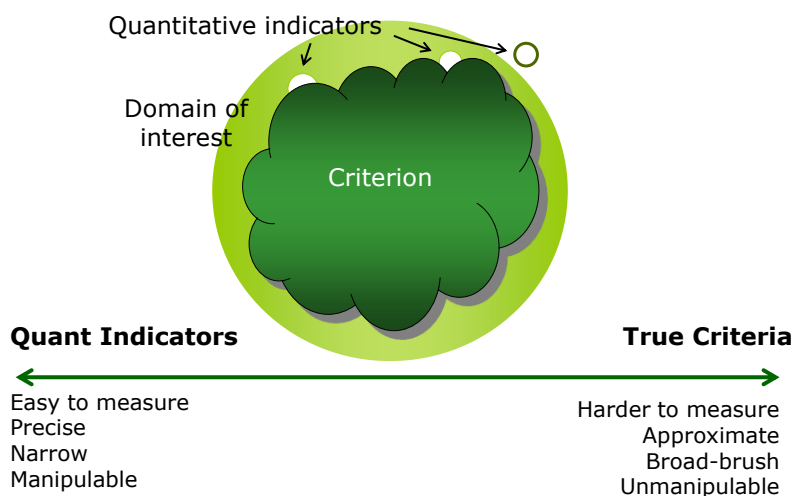
- Executive Summary
 - 2 pages
 - 7 +/- 2 headings
 - Each heading is one of your key evaluation questions
 - 1-2 paragraphs under each heading
 - These 1-2 paragraphs directly and succinctly answer the question!
- Introduction
- Methodology (& why)
- Values & Context
- Key Evaluation Questions
- Findings
 - 7 +/- 2 sections
 - Each section heading is one of your evaluation questions
 - Succinct answer, followed by the evidence

How rubrics help you do real, genuine, actionable self-assessment



- Cover the vast majority of the domain
=> largely unmanipulable
- Versatile – they work well for:
 - “fairly quick and fairly clean” or in-depth analysis
 - qualitative, quantitative or mixed methods
- Serious about values - perfect for ensuring community, cultural & organisational values/aspirations incorporated
- Build shared understanding of what constitutes ‘effectiveness’
- Identify multiple levels of performance and progress, not just one vaguely ‘acceptable’ level

Why not just stick with quantitative indicators?



Mix it up with your indicators!



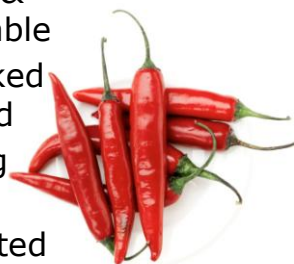
SMART Indicators

- S** Specific
- M** Measurable
- A** Achievable
- R** Realistic
- T** Time-based



SPICED Indicators

- S** Subjective (using judgement/expertise)
- P** Participatory
- I** Interpreted & Communicable
- C** Cross-checked & Compared
- E** Empowering
- D** Diverse & Disaggregated





Visible values → actionable answers

- Start with broad-brush *evaluative* questions
- Define 'quality' and 'value' & use rubrics to:
 - Spark genuine conversations about what constitute "good" course & programme design, teaching & outcomes (for students, employers, community)
 - Visualise what "success" looks like at multiple levels => more likely to realise it
 - Far better synthesis of mixed method evidence => more succinct, to-the-point findings => greater uptake and utilisation
- Use reporting structures that demand actual answers!

29



References and resources

- Davidson, E. J. (2007). **Unlearning Some of our Social Scientist Habits.** *Journal of MultiDisciplinary Evaluation*, 4(8), iii-vi. <http://tiny.cc/unlearning>
- Jane's website: **<http://RealEvaluation.com>**
[Various tips and tools available – more coming soon!
Subscribe to the site feed.]
- Genuine Evaluation blog (Jane Davidson & Patricia Rogers)
<http://GenuineEvaluation.com>
- Davidson, E. J. (2004). ***Evaluation methodology basics: The nuts and bolts of sound evaluation.*** Thousand Oaks, CA: Sage.

