

The Rubric Revolution

Presentation at the



Conference, Sydney, Australia
September 1, 2011

Jane Davidson, Nan Wehipeihana & Kate McKegg
with **Patricia Rogers (discussant)**

Evaluative Rubrics: The Basics



<http://RealEvaluation.com>

E. Jane Davidson, Ph.D.

Real Evaluation Ltd.
Auckland, Aotearoa New Zealand

Real evaluation is *one step* *harder* than descriptive research

- Descriptive research asks “What’s So?”
 - **What are** the outcomes of a particular intervention? (and why, for whom, etc)

- Evaluation also asks:
 - **How good are** the outcomes ...?
 - ... and are those outcomes ***good enough?***

Much so-called 'evaluation' skips this "how good is good" step



"You work it out"

Others do it with smoke & mirrors ...

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



Rubrics are an attempt to be systematic and transparent about it



What makes evaluation eVALUation?

What's So?

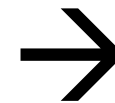
Descriptive
Facts

+

(qualitative,
quantitative
& mixed
method
evidence)

Definitions of
"quality"
& "value"

(also facts,
evidence-based
e.g. Needs,
potential,
aspirations)



So What?

Evaluative
conclusions

(saying
something
explicit about
quality, value)

What are evaluative rubrics?

- ❑ Interpretation guides for evidence (qual, quant, mixed; considered *as a set*)
- ❑ “Ladders of change”
- ❑ Definitions of “how good is good” and “how good is good enough”
- ❑ “Evidence pictures/scenarios” – which one is the “best fit” with the evidence we have?

Rubrics may be very criterion-specific, e.g. Parent & whānau engagement in education

Rating	Description
H i g h l y effective	<ul style="list-style-type: none"> ▣ Parents/whānau are extremely well-informed, confident and highly engaged in their children’s education in ways that maximise the children’s potential. ▣ Parent and whānau knowledge and perspectives are well respected, highly valued and fully integrated in ways that benefit the children’s education. ▣ Māori content and language are clearly evident and infused in ways that are appropriate for local whānau.
Minimally effective	<ul style="list-style-type: none"> ▣ Levels of parent/whānau/caregiver engagement are <u>just sufficient</u> to support children’s education, although there is significant room for improvement ▣ The school demonstrates understanding of Māori, Pasifika and other cultures, including the concepts of whānau, co-parenting and other family structures.
Poor or Detrimental	<p><u>Any one or more</u> of the following:</p> <ul style="list-style-type: none"> ▣ Levels of whānau engagement are <u>extremely low</u> or are <u>deteriorating</u> – to an extent that adversely impacts children’s education ▣ Whānau report being talked “at” or down to, made to feel unwelcome or stupid, or that their perspectives are disrespected or sidelined ▣ Information is either withheld or presented in ways that prevent meaningful whānau involvement

Source: MOE projects (various)

Rubrics can also be generic, to be applied across a range of questions/criteria

Performance Rating	Performance Descriptors for Answering 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.

Rubrics demand systematic use of evaluative inference to make ratings

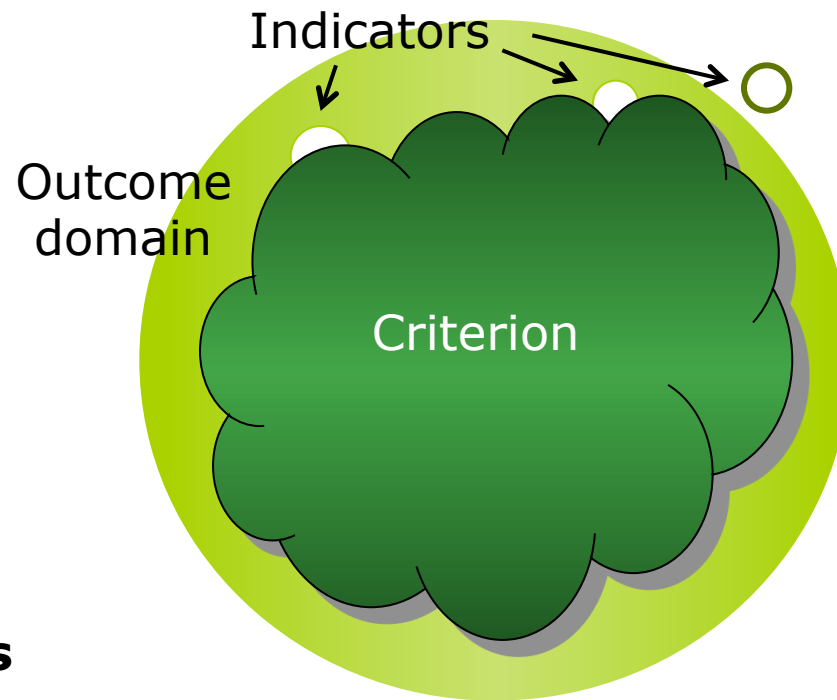
e.g. when rating “Good”		you need to show ...
Key points from performance descriptors	Performance is <u>generally strong</u> in relation to the question.	Specifically, what evidence led you to believe performance was “generally strong” – as opposed to “clearly very strong or exemplary” (excellent) or “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?

How do you create rubrics?

- ❑ Get the right people in the room!
- ❑ Clearly identify your evaluation question or criterion
- ❑ Brainstorm what distinguishes 'highly effective' from 'ineffective' (or worse) performance
- ❑ Draw boundaries around what's in and out
- ❑ Facilitate the rubric writing process
- ❑ Debate; recalibrate; field test; hone

Why rubrics?

Why not just indicators?



Indicators



Easy to measure
Precise
Narrow
Manipulable

True Criteria

Harder to measure
Approximate
Broad-brush
Unmanipulable

The three kinds of ‘subjectivity’

1. Arbitrary, idiosyncratic, unreliable, and/or highly personal (i.e., based on personal preferences and/or cultural biases)
2. Assessment or interpretation by a person, rather than a machine or measurement device, of something external to that person (e.g., expert judgment of others’ skills or performance)
3. About a person’s inner life or experiences (e.g., headaches, fears, beliefs, emotions, stress levels, aspirations), all absolutely real but not usually independently verifiable

Plus the red herring: Subjective vs. objective measures

Using rubrics to give voice to indigenous values



<http://kinnect.co.nz/>

Nan Wehipeihana



Aotearoa New Zealand

Development of rubrics in indigenous contexts

Rubrics is a difficult concept to explain in plain language.

So language matters - how we talk about rubrics, and core rubric concepts of success, levels and progression

Explicit way of defining "quality" and "value" for ourselves and for others.

Compared to

To identify what's important to us, what's important to others, what matters and come to a shared understanding...

Cultural metaphors

In Māori contexts we've found Poutama a stepped pattern seen in tukutuku as a useful metaphor

- Generally referred to as the steps to heaven, pathways
- Also symbolises various levels of learning and intellectual achievement



Cultural metaphors

Māori health provider who are using rubrics with whānau (families) talk about rubrics as 'a ladder'.

Use the rubric as a tool to frame discussion:

- Where an individual or whānau see themselves on the ladder (rubric)
- Planning to move to the next rung of the 'ladder'



Development of rubrics in indigenous contexts

So Māori get the concept (poutama) but the language of rubrics (of evaluation) can be hard to get

- An explicit way of defining “quality” and “value” for ourselves and for others:
 - Evaluative criteria (what aspects are important?)
 - Merit criteria (how good is “good”?)
 - Making a judgment, using these criteria, about how good, or excellent (or poor) something is (merit determination)

Development of rubrics in indigenous contexts

The “success” **language** can be problematic

- Use te reo Māori
- Nga moemoea (dreams) Nga wawata (aspirations, desires)

Not just the language but concerns about ‘labeling’

- Use different terms for the performance levels
 - Te Kakano (seed, latent potential)
 - Tipuranga (grow, begin to develop)
 - Oranga (well, healthy, fit)
 - Puawaitanga (bloom, come to fruition)

Text dense matrix – is a turn off for some

- Alternative presentation formats
- When they get it conceptually – the language makes sense

Giving voice to indigenous values

I learned from my Grandparents...

I learned from my Parents...

Topic

*I would like to teach my
mokopuna (grandchildren)...*

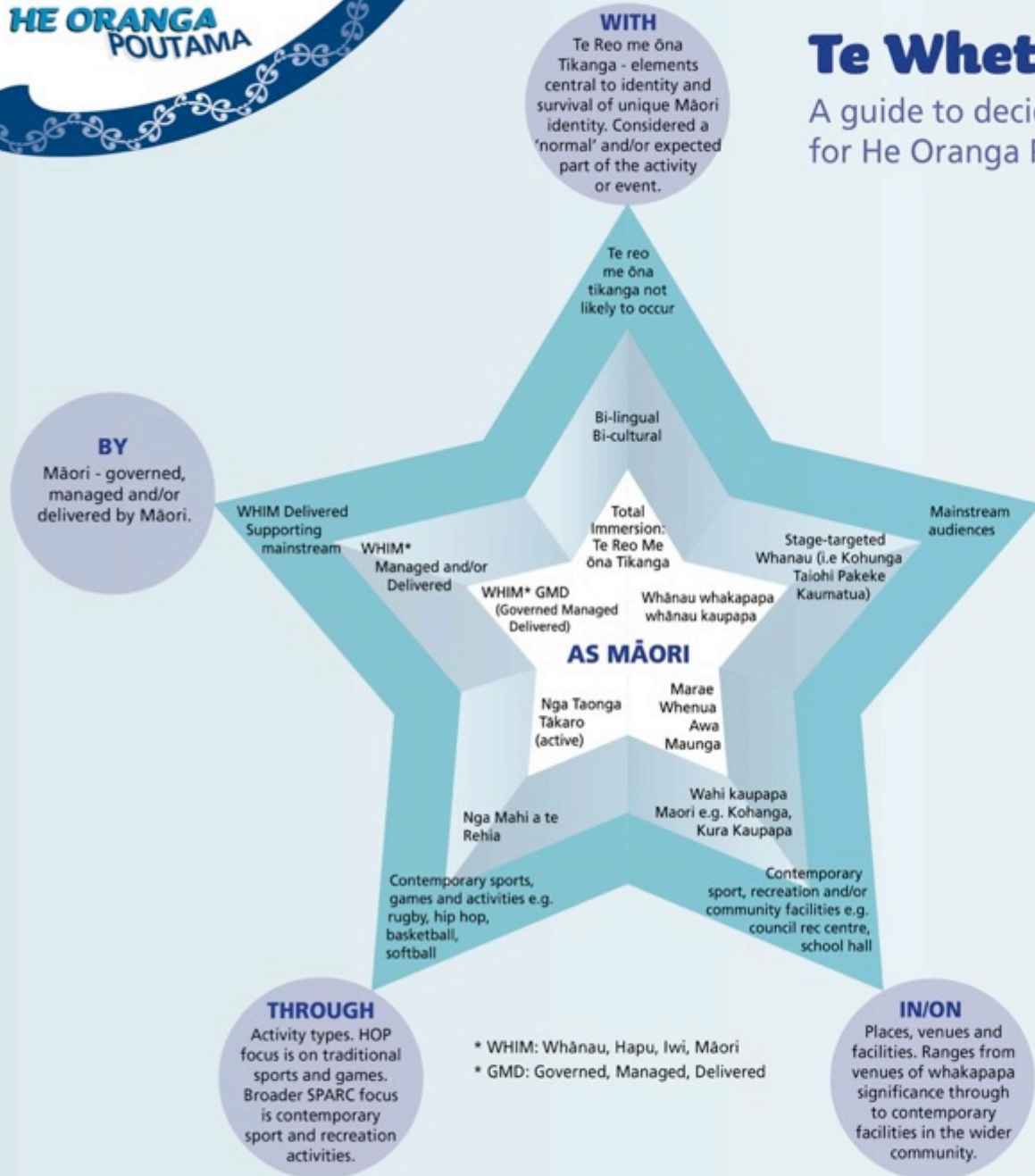
My tamariki (children) learned from me...

Example – an ‘as Māori’ developmental journey with five dimensions emerging

By Māori	Refers to the extent of control (rangatiratanga) by Māori in the governance, management and delivery of an initiative
For Māori	Emphasises the importance of family, sub-tribe, tribe (whānau, hapu, iwi), and other Māori collectives brought together for Māori purposes
With te reo me ona tikanga	Points to the importance of Māori language and customs (te reo me ona tikanga) as central to the survival and affirmation of what it means to live ‘as Māori’
In/on places of cultural significance to Māori	From ancestral places such as meeting houses, rivers, mountains (marae, awa, maunga), to more contemporary places founded on Māori principles such as immersion Māori kindergartens and schools (kōhanga reo, kura) etc.
Through	Those activities that can be distinguished as traditionally Māori (e.g., nga taonga taakaro, or ki o rahi)

Te Whetu Rehūa

A guide to deciding appropriate activities for He Oranga Poutama



* WHIM: Whānau, Hapu, Iwi, Māori
* GMD: Governed, Managed, Delivered

- This 'Whetu' outlines a five-criteria continuum to help determine 'as Māori' participation for the context of the He Oranga Poutama initiative.
- The closer an activity maps to the criteria in the inner star, closest to the 'As Māori' centre, the more likely it is to contribute to HOP's goal of participating as 'Māori' in sport and recreation.
- Generally three dimensions of the inner star are required for a strong HOP goal connection.
- Dimensions of the outer star are strongly aligned to participation in sport and recreation by Māori in mainstream initiatives or events.

Knowing about rubrics is one thing ... tips, tricks and traps

<http://kinnect.co.nz/>

Kate McKegg



Aotearoa New Zealand

Three case examples

Rubrics at:

- Individual level
- Project level
- Programme level

A mental health setting

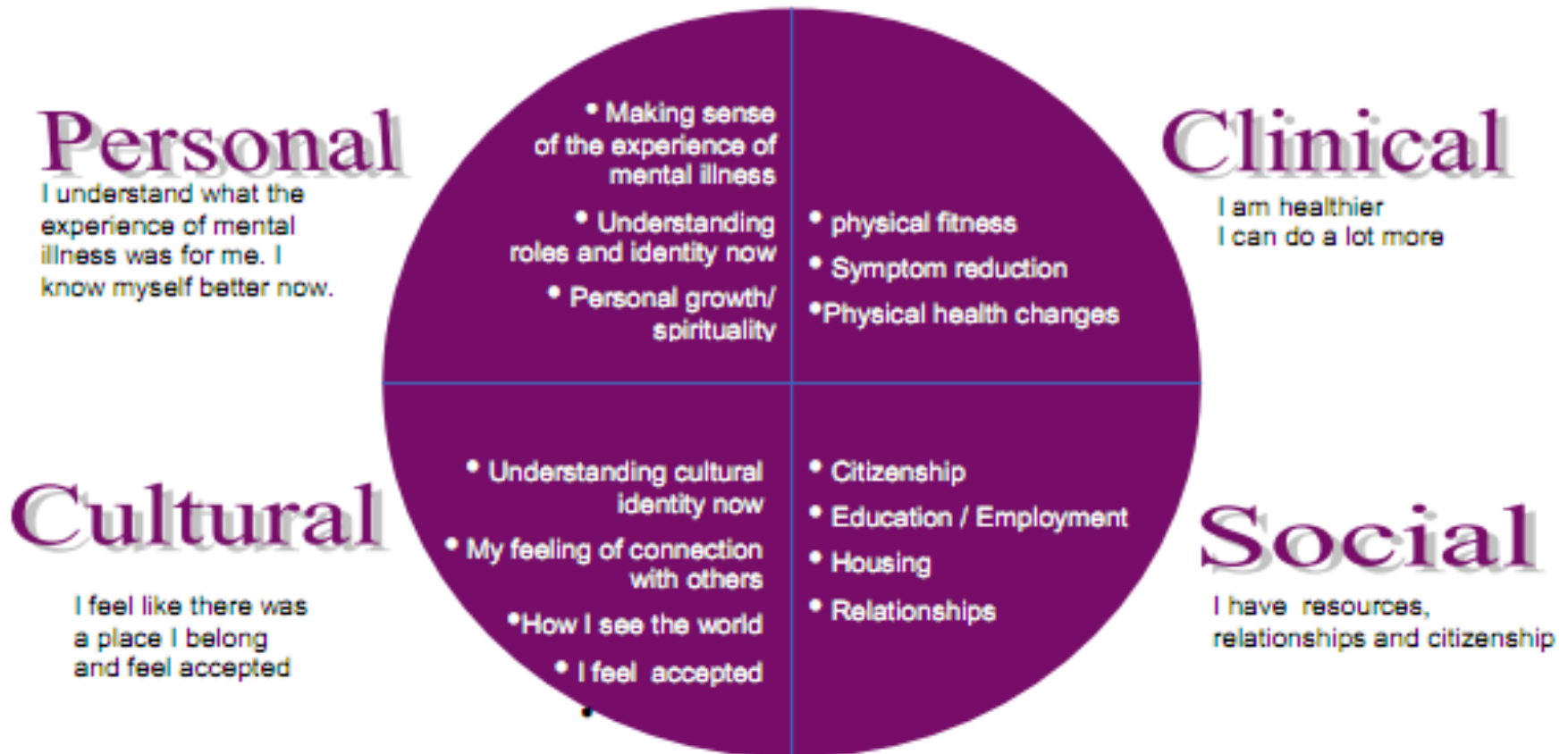
- Provider is recovery focused
- Has a strong commitment to peer led and peer informed services
- Not satisfied with many of the more 'objective' outcome measures being used in mental health, i.e., bed nights in acute care, readmission rates, etc.
- Wanted a framework that 'valued' what consumers saw as 'recovery' and 'wellness'

Wellink's outcome framework is recovery focused

RECOVERY OF WELLBEING

4 markers that indicate an environment that enhances recovery of well-being

- Hope
- Personal Power
- Self Determination
- Sense of Belonging



A suite of evaluative criteria developed

A Recovery Matrix

Personal

Hope for the future

Quality of life

Spirituality / Personal Beliefs

Cultural

Culture

Relationships (whānau)

Autonomy

Clinical

Daily living skills

Mental Health

Physical Health

Social

Housing and accommodation

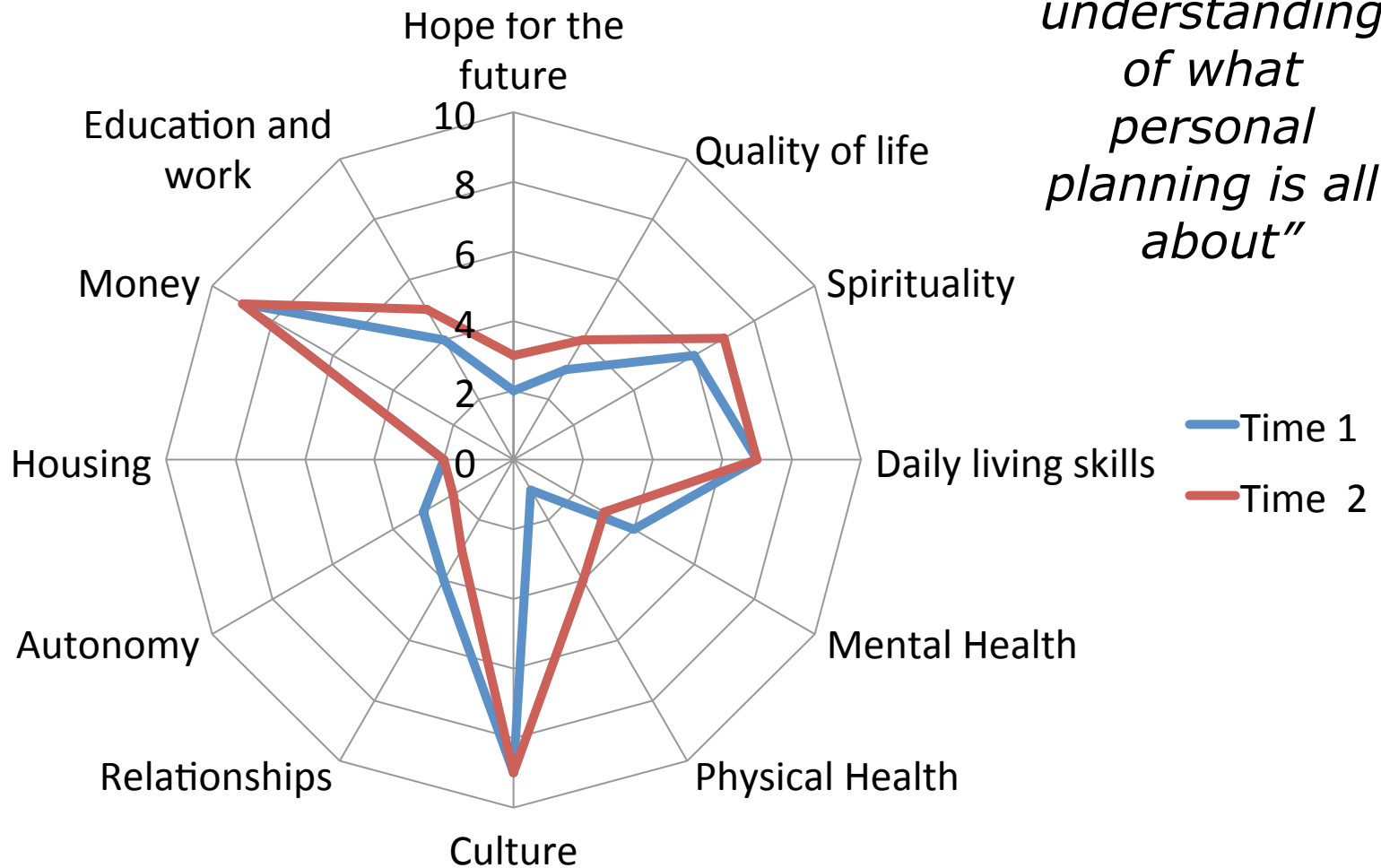
Money and finances

Education, training and work
(paid and unpaid)

Individual outcome rubric: Housing and accommodation

Crisis	At Risk	Stable / supported	Self-sufficient	Thriving
<p>I am homeless, or about to be evicted: my personal safety is at risk and I am unsafe to remain at home. I am unable to continue living either on my own or with the people I live with. The situation is intolerable</p>	<p>I am at risk of eviction, either because of my actions or because I can not afford the rent. I feel unsafe in my home. My house is in a state of disrepair. I lack some furnishings which I would like. I would like to change my living arrangements.</p>	<p>My housing Is okay – it is warm, dry and affordable. I am safe in my home and have essential furnishings. Some of the time I am ok with who I live with but would like things to be different in the future</p>	<p>I am happy with my housing. My home is comfortable, safe and affordable and furnished to my liking. I am happy with my living arrangements – I like living by myself, with family or housemates.</p>	<p>I am settled and happy in the house of my choice, at this time. It is comfortable, safe and affordable and furnished to my liking. I really enjoy living by myself, with family or with housemates.</p>

A graph of change – used for understanding the recovery journey



"Expanded our understanding of what personal planning is all about"

A program level example – acute peer support residential service

- Characteristics of peer support evaluative criteria were
 - driven by the literature
 - informed by practice
 - and fine-tuned by stakeholders

- The four key evaluative criteria developed were:
 - There is effective and appropriate client (guest) engagement
 - Peer support workers are able to provide a climate and context which assists guest recovery
 - Peer support workers have the skills and attitudes to assist guest recovery
 - There is effective and appropriate management of peer support teamwork

Then we determined levels of performance

Rating	Performance Descriptors for Answering KEQ
Excellent (Always)	Clear example of exemplary performance or best practice in this domain: no weaknesses
Very good (Almost always)	Very good to excellent performance on virtually all aspects; strong overall but not exemplary; no weaknesses of any real consequence
Good (Mostly, with some exceptions)	Reasonably good performance overall; might have a few slight weaknesses, but nothing serious
Adequate: (Sometimes, with quite a few exceptions)	Fair performance, some serious, but non fatal weaknesses on a few aspects
Poor: Never (Or occasionally with clear weakness evident)	Clear evidence of unsatisfactory functioning; serious weaknesses across the board on crucial aspects
Insufficient evidence	Evidence unavailable or of insufficient quality to determine performance

... For a combination of both quantitative and qualitative data

Rating	Quantitative data	Qualitative data
Excellent: (Always)	90% or more agree with statement	Clear example of exemplary performance or best practice in this domain: no weaknesses
Very good: (Almost Always)	80% - 90% agree with statement	Very good to excellent performance on virtually all aspects; strong overall but not exemplary; no weaknesses of any real consequence
Good: (Mostly, with some exceptions)	60% - 80% agree with statement and no more than 15% disagree	Reasonably good performance overall; might have a few slight weaknesses but nothing serious.
Adequate: (Sometimes, with quite a few exceptions)	40% - 60% agree with and no more than 15% disagree	Fair performance, some serious, but non fatal weaknesses on a few aspects
Poor: Never (or occasionally with clear weaknesses evident)	Less than 40% or more agree with statement	Clear evidence of unsatisfactory functioning; serious weaknesses across the board on crucial aspects

Reporting: can summarise evaluative criteria level overall

	Ratings				
	Poor	Adequate	Good	Very good	Excellent
Overall rating					
Effective and appropriate client (guest) engagement					
Peer workers are able to provide a climate and context which assists guest recovery					
Peer workers have the skills and attitudes to assist guest recovery					
Effective and appropriate management of peer support teamwork					

Or can summarise at a lower level

Table 4: Summary of evidence of effective and appropriate client (guest) engagement with Key We Way peer workers

Evaluative Criteria	Dimensions of merit	Ratings				
		Poor	Adequate	Good	Very good	Excellent
Effective and appropriate client (guest) engagement	Overall rating					
	Guests are relaxed in peer worker company, happy to see them, ok with physical closeness, positive language					
	Guests accept the service as something special					
	Guests are keen and willing to engage with service					
	Guests feel included in a relationship					
	Guests have a higher value of themselves – visible in body, self confidence and taking responsibility					
	Guests have a greater sense of hope, self efficacy and determination					
	Guests start to advocate for themselves					

Overall evaluative conclusion

“...The service shows great potential to affect significant and lasting change in guests for the following reasons:

- ▣ Generally strong on peer skills, relationship building and creating a climate of trust and mutual support focused on recovery
- ▣ Some management and supervision weaknesses although these were being actively addressed by the organisation”

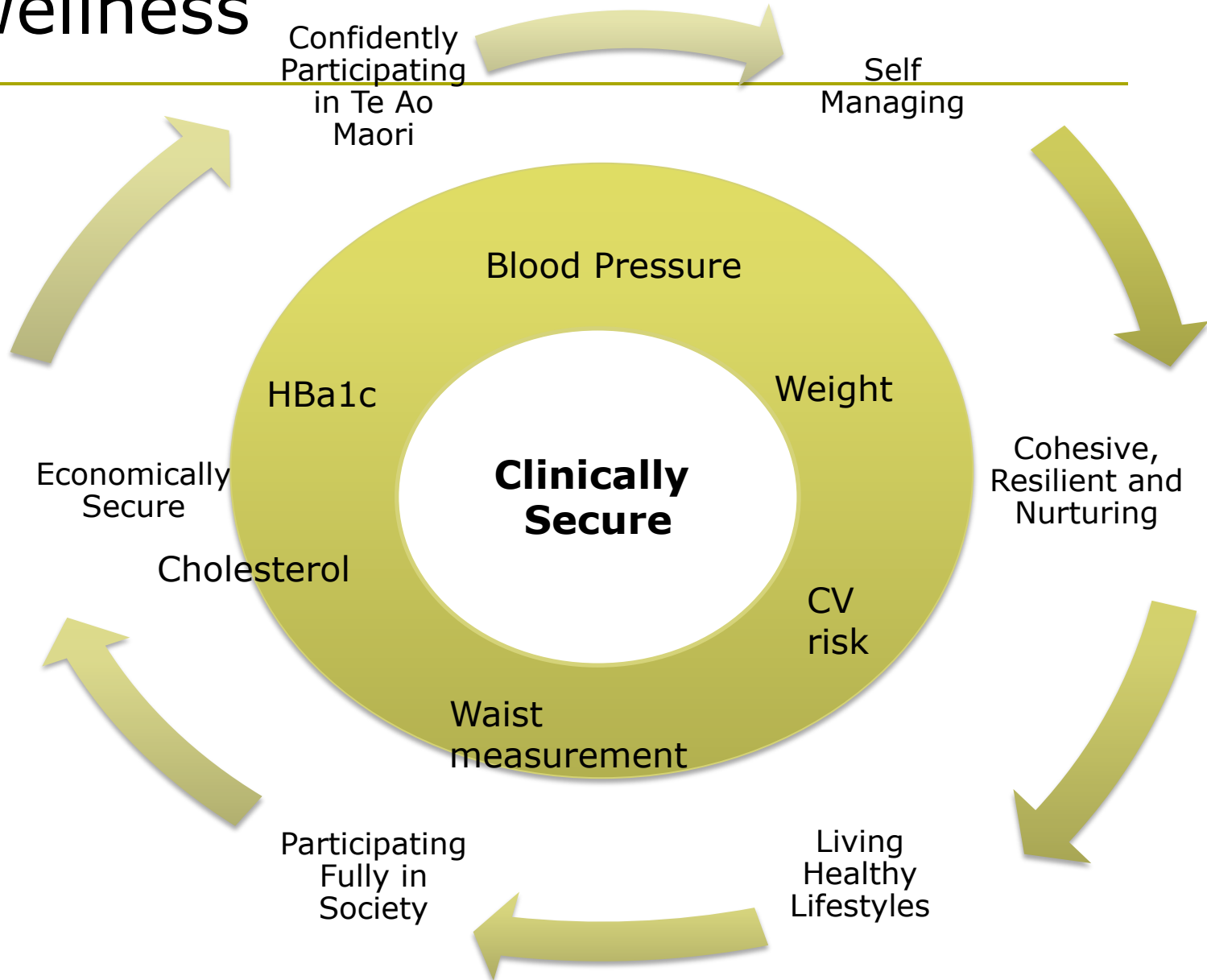
Diabetes prevention

- ❑ Complex, intractable health issue, at population level
- ❑ Among Māori in New Zealand, the profile of disease is much worse than for non-Māori
- ❑ Māori health providers – looking for innovative ways of tackling the issue

A new model of care

- The new programme model of care is intended to create a paradigm shift, transforming the more traditional 15 minute general-practice approach to care, to a care pathway and support programme involving a team working with whānau (families).

7 Outcome Domains – essential for wellness



Each outcome domain has 5 levels

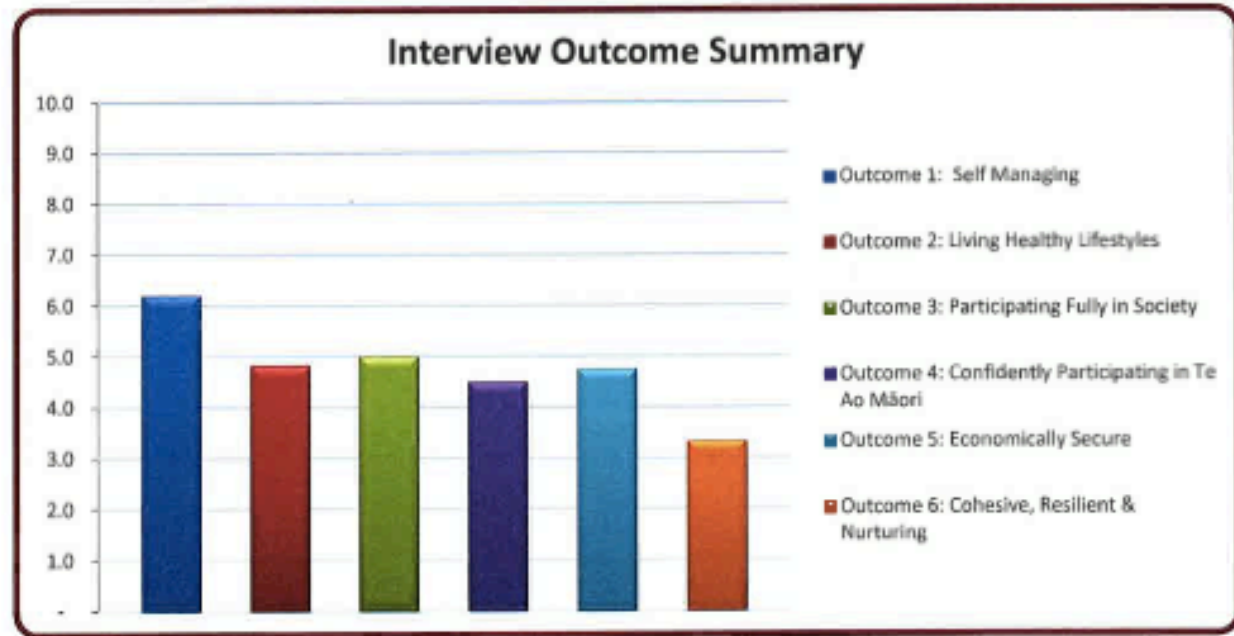
Levels	Evaluative criteria	
	Living healthy lifestyles	Participating fully in society
Self Reliance / reaching full potential	Increased participation in sport and recreational physical activity Improved choices made about consumption of healthy food and other substances	Increased recognition of whānau assets and strengths Improved educational opportunities and success among whānau
Wellbeing	Strengthened ability to access support when needed	Strengthened ability to connect and engage in and with community and whānau activities
Recognition and Belief	Improved whānau connectedness	Greater abilities and confidence to advocate for whānau needs and aspirations
Awakening / awareness	Greater ability to maintain safety (emotional, mental, spiritual and physical)	
Unwakened / untapped potential		

Whānau are self assessing

Kia Rite Outcome Summary

Outcome Assessment

- 10 Self reliance / reaching full potential - Manawa
- 9 Self reliance / reaching full potential - Hapai
- 8 Wellbeing - Manawa
- 7 Wellbeing - Hapai
- 6 Recognition and belief - Manawa
- 5 Recognition and belief - Hapai
- 4 Awakening / awareness - Manawa
- 3 Awakening / awareness - Hapai
- 2 Unwakened / untapped potential - Manawa
- 1 Unwakened / untapped potential - Hapai



Key tips, tricks and traps when developing rubrics

- Context really matters
- Getting the right people in the room is important
- Language is important for acceptance and use - Rubrics are sexy. But 'rubrics' is NOT a sexy word!!
- Directionality isn't a trivial issue
- Don't overcomplicate it

The value of rubrics

- Versatile – they work well for:
 - participatory or ‘independent’ evaluation
 - “fairly quick and fairly clean” or in-depth analysis
 - qualitative, quantitative – but especially good for mixed methods
 - needs assessment (baseline) and outcomes
 - overviews and drill down

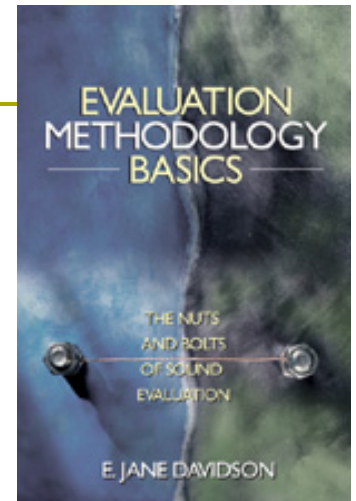
- Build shared understanding of what constitutes effectiveness – provided you get the ‘languaging’ right!

The value of rubrics

- ❑ Serious about values - perfect for ensuring community, cultural & organizational values/aspirations incorporated
- ❑ Identify multiple levels of performance and progress, not just one vaguely 'acceptable' level
- ❑ Cover the vast majority of the process or outcome domain => largely unmanipulable

References and resources

- Davidson, E. J. (2004). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage.
- Measurable Gains Framework rubrics downloadable from <http://tiny.cc/kahikitia>
- Self-review tool for schools: Focus on students achieving below curriculum expectations in literacy from <http://tiny.cc/literacytool>
- Genuine Evaluation blog (Jane Davidson & Patricia Rogers) <http://GenuineEvaluation.com>
- Jane's website: <http://RealEvaluation.com>
- Kinnect Group website <http://kinnect.co.nz/>

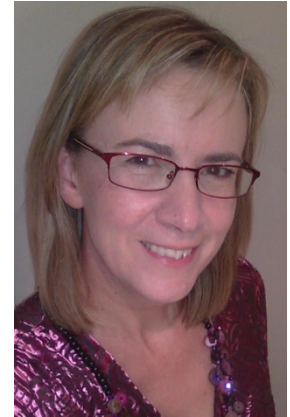


References: Rubrics

- Rubrics used in a mental health setting: downloadable from <http://www.wellink.org.nz/pdfs/CharacteristicsofGoodPeerSupport.pdf>
- Whitney, L., Bennie, N., & Wehipeihana, N. (2011) 'Māori Enjoying Education Success as Maori: Credible evidence of effectiveness for Māori learners – Ka Hikitia presented at the PPTA Conference, 20 April, Wellington.
<http://kinnect.co.nz/wp-content/uploads/2011/08/110420-Whitney-Bennie-Wehipeihana-PPTA.pdf>
- Davidson, E.J., & Wehipeihana, N. (2011) 'Credible evidence of effectiveness for Māori learners: An introduction to the Measurable Gains Framework & related rubrics' presented at the National Aspiring Principals Programme Hui, 18 April, Auckland.
<http://kinnect.co.nz/wp-content/uploads/2011/08/110418-Davidson-Wehipeihana-NAPP.pdf>
- Davidson, E.J., & Wehipeihana, N. (2011) 'Credible evidence of effectiveness for Māori learners: An introduction to the Measurable Gains Framework & related rubrics' presented at the National PLD Facilitator Training, 8 February 2011, Auckland.
<http://realevaluation.com/pres/Credible-evidence-PLD-conf-Waipuna-8-2-11.pdf>

Presenters

Jane Davidson
jane@RealEvaluation.com



Nan Wehipeihana
nanw@clear.net.nz

Kate McKegg
kate.mckegg@xtra.co.nz

