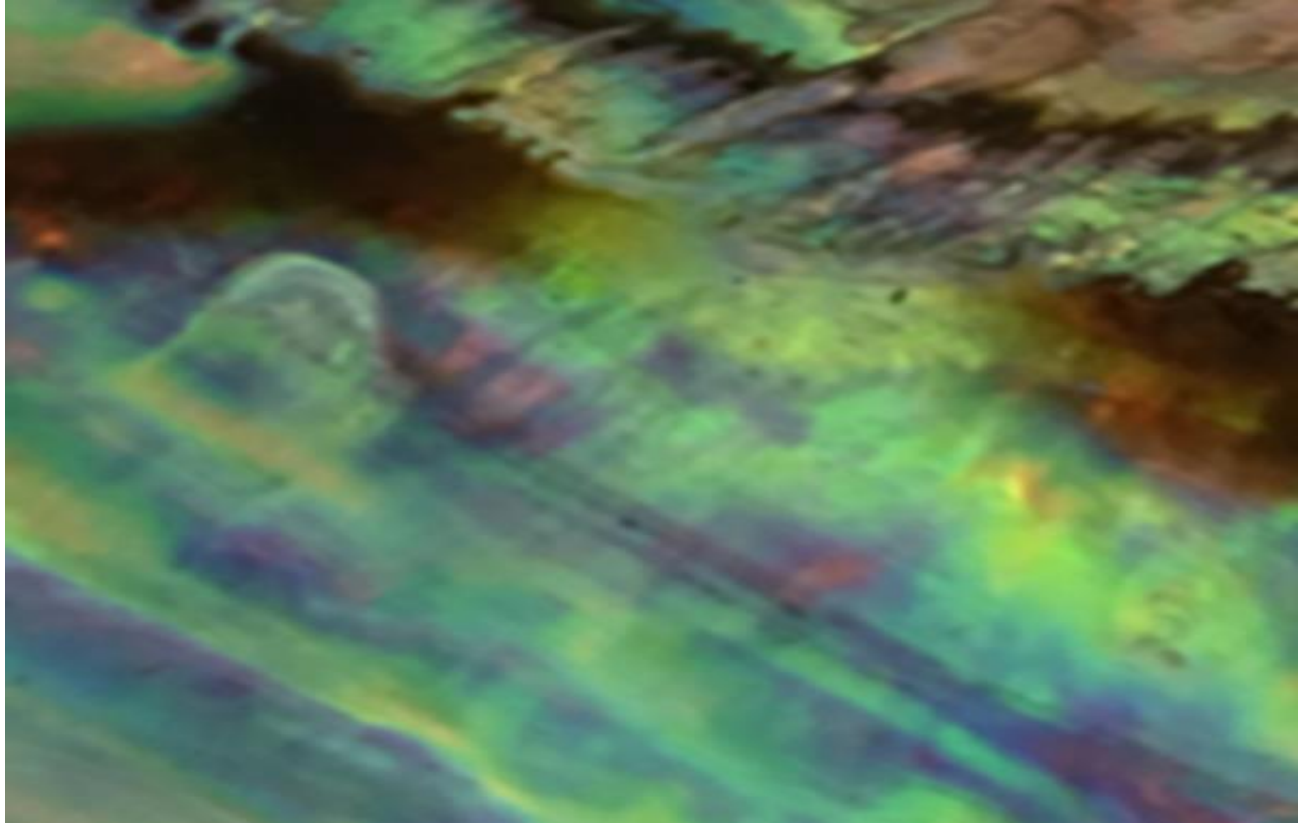
Understanding the components of evaluative rubrics and how to combine them



Developed by: Judy Oakden
Pragmatica Limited: Wellington New Zealand

March 2018

Pragmatica

A member of the Kinnect Group

Who is this e-book for?

This e-book is for practitioners and those wanting to use evaluative rubrics. Sometimes certain types of rubric are not suited for the kind of evaluation you are doing. In these instances you can get 'stuck'. This e-book will offer you some alternative approaches to rubrics and/or hopefully help you get unstuck.

In this e-book I draw on both my own experiences and the observations from Samantha Stone-Jovicich who wrote a useful reflective practice paper: *To rubrics or not to rubrics: An experience using rubrics for monitoring, evaluating and learning in a complex project* (2015) .

Note, where we talk about 'the evaluator' in this context other stakeholders as well as the client may be included.

"We spent weeks writing up the descriptions for ratings categories. Soon the rubrics became bigger than Ben Hur, taking up pages and pages of tables filled with minuscule writing…

For us, the proverbial 'straw that broke the camel's back' came with the feedback from team members: no-one could agree on what constituted a 'poor', 'adequate', 'good', and 'excellent' [description]."

(Stone-Jovicich, 2015, p4).

You are welcome to use this resource, we just ask that you cite our work:

Oakden, J. (2018.) *Understanding the components of evaluative rubrics and how to combine them.* Wellington, New Zealand: Pragmatica Limited

© Pragmatica Limited 2018

What is the main idea of this e-book?

For some time it has been known that there are a number of different ways rubrics can be constructed (Davidson, 2005). My 'ah-ha' moment grappling with the different kinds, was to conceptualise rubrics as having three basic components.

This e-book explores how the components are combined for different types of rubrics.

I will show you some alternative ways the components can be combined to produce three types of evaluative rubrics for different settings. I am expecting my thinking will evolve further over time. For now, this e-book is based on the types of rubrics I have used in my evaluation practice.

Remember – there is no one 'good' type of rubric – different types of rubrics are good for different evaluation contexts.

© Pragmatica Limited 2018

Introduction – why evaluative rubrics?

"Evaluation, by definition, must answer truly evaluative questions: it must ask not only 'What were the results?' (a descriptive question) but also 'How good were the results?'. This cannot be done without using evaluative reasoning to evaluate the evidence relative to the definitions of quality and value" (Davidson, 2014, p6).

There are a number of ways of making evaluative judgements of quality, value and importance in an evaluation. I use the evaluative rubric process.

While evaluative rubrics are a useful tool in the evaluators toolkit, developing and using them can be more difficult than expected (King, et al, 2013). I think one of the reasons for this difficulty may be that there are different kinds of rubrics. Each is better suited to different contexts.

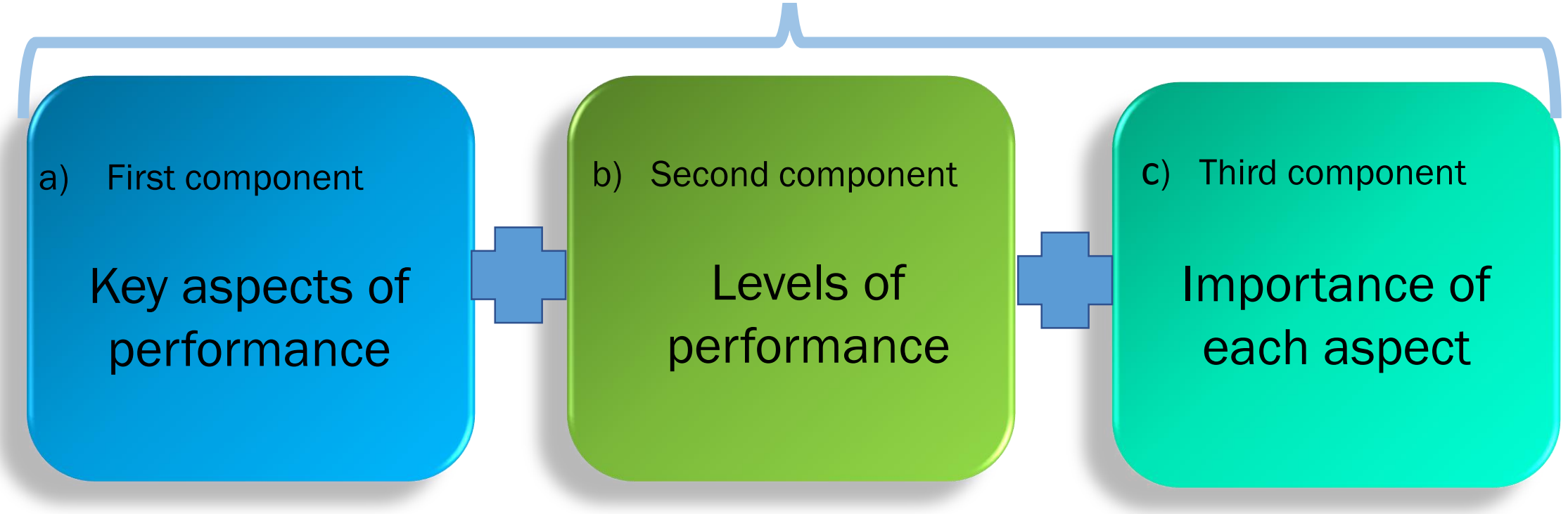
"An evaluative rubric is a table that describes what the evidence should look like at different levels of performance, on some criterion of interest or for the programme/policy overall" (Davidson, 2014, p6).

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

© Pragmatica Limited 2018

Basic components that make up an evaluative rubric

There are three basic components evaluative rubrics are made of: the key aspects of performance, the levels of performance and the importance of each aspect of performance (Eoyang & Oakden, 2016). **Note:** Performance can include activities, outputs or outcomes.

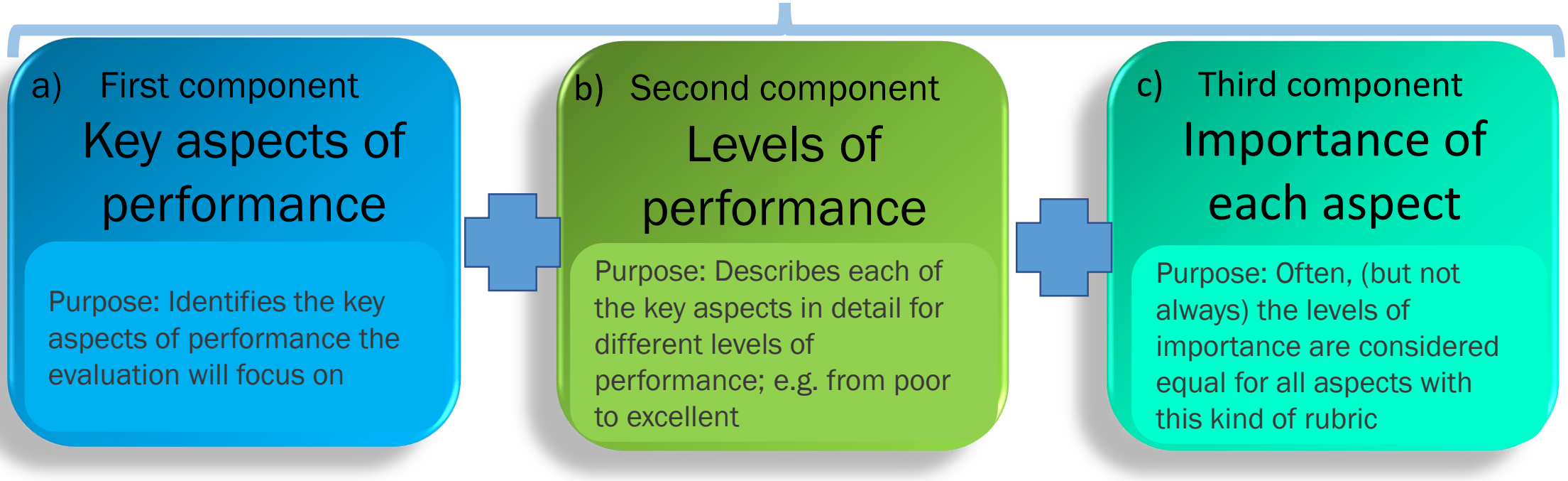


© Pragmatica Limited 2018

First example

The type of evaluative rubric most people are familiar with is an **analytic rubric**

For an **analytic rubric** the evaluator (often in conjunction with other stakeholders including the client) firstly identifies the key aspects of performance. Then they develop up detailed descriptions of what 'poor' through to 'excellent' might look like for each aspect of performance. In my experience there was not so much focus on the third component, the importance of each aspect of performance.



© Pragmatica Limited 2018

Here is an example of what an **analytic rubric** looks like with the levels of performance described in detail:

First component

Aspect of performance: Participants in professional learning are recognised as 'leading learners' in their school

Levels of performance	Rich description of the aspects of performance for each level of performance
Excellent	*Clear example of exemplary performance or best practice in this domain: no weaknesses* There are always clear and appropriate professional development goals set for FTPs Good working relationships which provide professional support and advice to FTPs are always established between FTPs and mentors FTPs always engage in reflective learning about being leading learners in their schools FTPs always understand the importance of being leaders of learning and have clear strategies to effect this FTPs always report they know how to collect, analyse and act on data to support student learning There is always evidence of the FTPs focus on equity for Māori Support for FTPs is well co-ordinated (especially where there are several support groups working with the FTP).
Good	*Reasonably good performance overall; might have a few slight weaknesses but nothing serious* There is mostly (with some exceptions) clear and appropriate professional development goals set for FTPs Good working relationships which provide professional support and advice to FTPs are mostly (with some exceptions) established between FTPs and mentors FTPs mostly (with some exceptions) engage in reflective learning about being leading learners in their schools FTPs mostly (with some exceptions) understand the importance of being leaders of learning and mostly have strategies to effect this FTPs mostly (with some exceptions) report they know how to collect, analyse and act on data to support student learning There is mostly (with some exceptions evidence of the FTPs focus on equity for Māori Support for FTPs is mostly (with some exceptions) well co-ordinated (especially where there are several support groups working with the FTP).
Poor	*Clear evidence of unsatisfactory functioning; serious weaknesses across the board on crucial aspects* There is little, or no evidence of clear and appropriate professional development goals set for FTPs Good working relationships which provide professional support and advice to FTPs are not established between FTPs and mentors FTPs do not report engaging in reflective learning about being leading learners in their schools There is no evidence that FTPs understand the importance of being leaders of learning, nor do they have strategies to effect this FTPs report they do not know how to start to collect, analyse and act on data to support student learning There is very little or no evidence of the FTPs focus on equity for Māori Support for FTPs is not at all co-ordinated (especially where there are several support groups working with the FTP).

Second component

In this case the third component is not used

Source: Adapted from: Oakden, J Oakden, J. (2013). Evaluation rubrics: how to ensure transparent and clear assessment that respects diverse lines of evidence. Melbourne: Better Evaluation. (p.6-7)

© Pragmatica Limited 2018

What do we notice about this kind of rubric?

Benefits

- Helps everyone get clear about what 'good' performance looks like
- Clearly shows progress
- Useful for non-evaluators to better understand how judgements are made.

Challenges

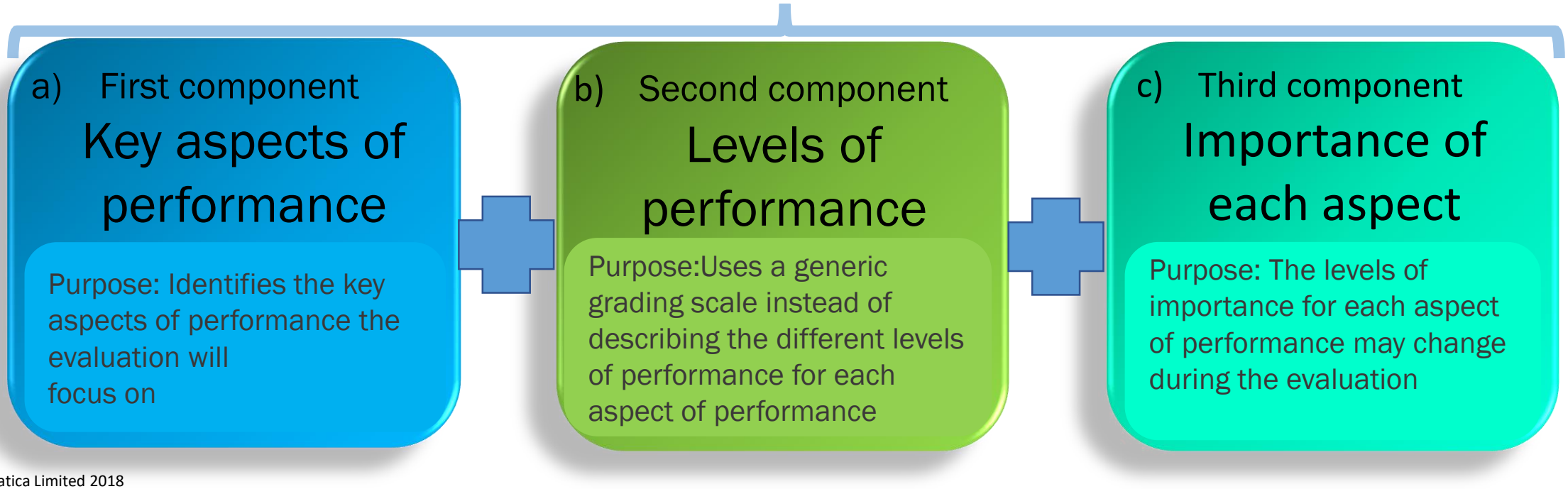
- This analytic rubric just describes one aspect of performance – if there are five or six aspects of performance the rubric gets complex and time consuming to develop
- There is a LOT of detail in the rich descriptions
- We have only shown three different performance levels; excellent, good and poor, but sometimes there are five or six different levels of performance used
- Difficult to pre-determine what might be important to capture, and can lock evaluators into aspects that are not so important if the context changes
- When I've used this type of rubric I often haven't said which aspects of performance are more important – I've treated the different aspects as being similarly important.

© Pragmatica Limited 2018

Second example

A less familiar evaluative rubric uses a **generic grading scale** as the second component

For this kind of rubric the evaluator firstly lists the key aspects of performance the evaluation will focus on. These are generally short statements. Instead of describing what 'poor to excellent' performance would look like, a generic grading description is used. There are a number of different scales that can be used for this – a common one is from 'poor to excellent', but there are also others. With this type of rubric the evaluator collects and analyses the data and then checks it against the generic grading scale to determine the level of performance later in the evaluation process.



© Pragmatica Limited 2018

Here is an example of the components of an evaluative rubric with a generic grading scale

First component

Aspect of perform-ance	General description of components of performance
Administrative efficiency	It is evident that: • administrative costs are in line with expectations (for PS, Fund applications, DFOs and TAs) • there is an appropriate balance between administrative spend and efficiency • balance of projects funded • added value of administration to applicants – find it helpful to business, has secondary benefits • low levels of complaints/challenges to processes • administrative processes are well set up, timely or robust, accurate and credible
Relationships – collaboration in the sector	Stakeholders comment that: • there is a cognisance of both political and commercial realities • they are aware of and respect each other's positions/views • they are able to work together and develop long term partnerships • media coverage reflects the positive relationships in the waste sector • relationship leverage is in broader areas than just waste
Good practice – building capability/capacity (including infrastructure) across the sector	It is evident that: • there are efficient waste data collection systems • stakeholders use WMA tools and they work effectively (including TAs, product stewardship, WMF and DFOs) • there is value for business in waste minimisation and management • TAs have good waste minimisation and management plans and implementation
Information, awareness and compliance (both in general and MfE's performance)	It is evident that : • there is awareness and knowledge of waste minimisation approaches including recycle, recover, reuse and product stewardship • data collection systems are effective and enable good measurement (both in general and MfE's systems) • data collection systems are comprehensive and well resourced (both in general and MfE's systems) • there is support for WMA at senior levels in MfE • monitoring and compliance systems support voluntary compliance • feedback loops are useful and timely • while the WMA encourages voluntary compliance, MfE can take effective enforcement action if needed • MfE has credibility with stakeholder

Source: Oakden, J., & McKegg, K. (2011). Waste Minimisation Act implementation: evaluation of stakeholder perceptions. Wellington, New Zealand: Kinnect Group

Second component

Generic grading scale

Rating level	Generic description
Excellent (Always)	• Clear example of exemplary performance or great 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.

Source: Adapted from Davidson, E. J. (2005). Evaluation methodology basics: the nuts and bolts of sound evaluation. Thousand Oaks, CA: SAGE p. 137

Third component

Importance of each aspect at different stages in the evaluation

Aspect of performance	At the start	By the end
Administrative efficiency	Medium	High
Relationships – collaboration in the sector	Medium	Medium
Good practice – building capability/ capacity (including infrastructure) across the sector	Lower	Lower
Information, awareness and compliance (both in general and MfE's performance)	HIgh	Mid

© Pragmatica Limited 2018

What do we notice about rubrics that use a **generic grading scale**?

Benefits

- Helps everyone get clear about what aspects of performance are being focussed on
- Has much less detail than the first example, quicker to develop, more flexible during the evaluation process and can accommodate contextual change

> With all of the ratings category descriptions deleted, the rubrics tables became so much more user-friendly and useful. We were finally able to use them. (Stone-Jovicich, 2015, p.5).

And as noted in a recent presentation other benefits are:

- "Frames the evaluation differently from traditional linear thinking — captures complex dynamics and inter-linkages…
- Provides a broad-brush framework for evaluating different activities within one initiative, and captured 'messy' hard-to-measure dimensions
- Enables effective and aligned mixed methods data collection, synthesis and actionable reporting" (Source: Oakden & Bear, 2011, p.14)

> But perhaps one of the greatest (and understated) value-adds of the rubrics is not that it just integrates diverse viewpoints but that it can do so in an inductive, emergent way (i.e. as one gets more and more feedback, one starts to get a better picture of what 'excellent', 'good', and 'poor' means through the eyes of participants). This is particularly true with regards to the ratings. Rather than needing to define 'poor' upfront … it becomes defined over time as one accumulates feedback from a range of people and perspectives (e.g. different team members, external partners, participants in project sponsored activities, etc.). This is not insignificant.
>
> (Stone-Jovicich, 2015, p.8).

© Pragmatica Limited 2018

What do we notice about this kind of rubric? *(continued)*

Challenges

- At the start stakeholders and evaluators may not be clear about what 'good' aspects of performance look like, it may feel a bit 'loose'
- Has much less detail, this may be concerning to some who see it as less transparent
- Not as easy for non evaluators to understand - initially
- May make structuring the report writing challenging, if aspects of performance are not explicitly linked with the Key Evaluation Questions. While I like this approach I know of others who consider this approach can be hard to do well.

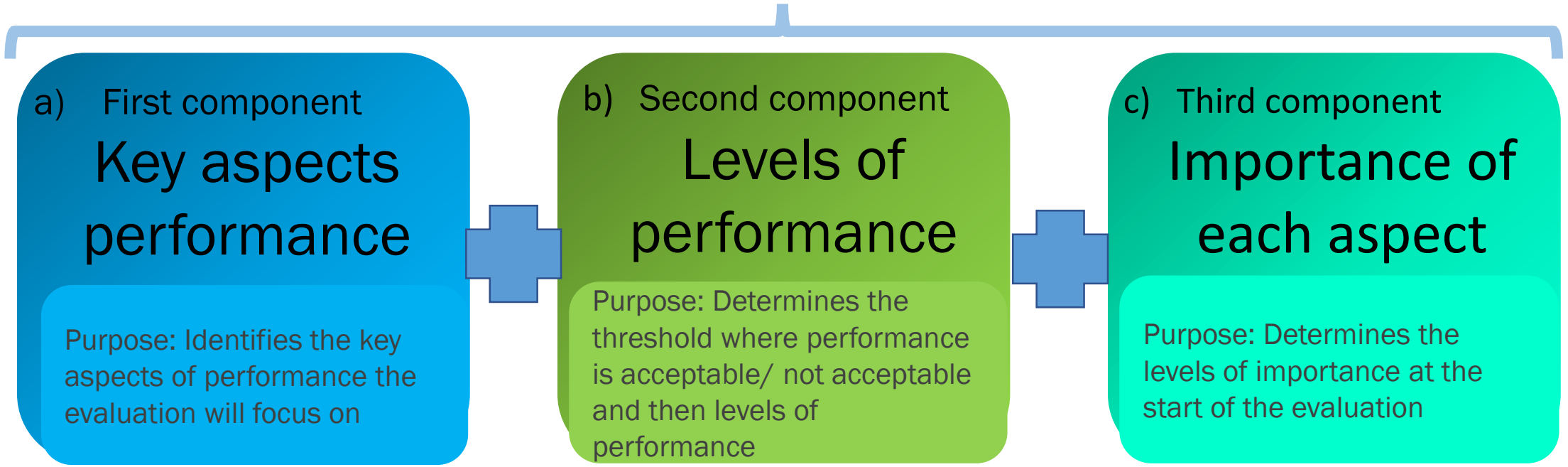
> …while rubrics are great for assessing individual or particular activities and outputs in isolation they are less effective in capturing the cumulative outcomes or impacts of multiple dimensions of a project, as it does not capture how they interact and what emerges as a result. Also the more components you have the more daunting becomes the task of pulling the data together into tables.
>
> (Stone-Jovicich, 2015, p.8).

© Pragmatica Limited 2018

Third example

The final type of evaluative rubric we will explore here is a **holistic rubric**

A holistic rubric helps assess something as a whole. For this kind of rubric firstly the evaluator identifies the key aspects of performance. Then the thresholds of performance are decided. Then the different aspects of performance are combined within the levels of performance.



© Pragmatica Limited 2018

Here is an example of the components of **holistic rubric**

First component: Aspect of performance

Second component: Levels of perform-ance

Level of performance	Value for Money
Excellent value for money	• Sufficient results from successful projects to provide clear evidence of positive return on investment from SFF investment overall AND • Credible contribution to export opportunities and improved sector productivity; and increased environmentally sustainable practice AND • Evidence of exemplary contributions to enhanced environmental, social and cultural outcomes including significant outcomes at Level 6 of the Bennett's Hierarchy and emergent outcomes at Level 7.
Very good value for money	• Sufficient results from successful projects to demonstrate we have already broken even on the SFF investment overall AND • Emerging contribution to export opportunities, improved sector productivity; and increased environmentally sustainable practice AND • Evidence of significant contribution to enhanced environmental, social, or cultural outcomes including significant outcomes at Level 6 of the Bennett's Hierarchy.
Good value for money	• Sufficient results from successful projects to credibly forecast break-even on the SFF investment overall AND • Credible contribution in encouraging primary sectors partnering, encouraging and co-investing in industry innovation and adoption, partnering innovative approaches to environmental challenges, and engaging with Māori AND • Evidence of emerging contribution to enhanced environmental, social, OR cultural outcomes including significant outcomes at Level 5 of the Bennett's Hierarchy and emergent outcomes at Level 6.
Minimally acceptable value for money	• The SFF is sufficiently well-utilised on a range of sufficiently promising projects to have a credible prospect of breaking even overall AND • Funds are being allocated and used in accordance with the intended purpose and strategic priorities of the SFF AND • Emerging contribution in encouraging primary sectors partnering, encouraging and co-investing in industry innovation and adoption, partnering innovative approaches to environmental challenges, and engaging with Māori AND • Evidence of emerging contribution to enhanced environmental, social or cultural outcomes – meets Levels 1–4 (Resourcing, activities, participation and reactions) on Bennetts Hierarchy and there are emerging examples from Level 5 (KASA – Knowledge, Attitudes, Skills and Actions).
Poor value for money	• Fund is not sufficiently well-utilised on a range of sufficiently promising projects and has no credible prospect of breaking even OR • No evidence of contribution to enhanced environmental, social, or cultural outcomes at Bennetts Hierarchy Level 5 (KASA – Knowledge, Attitudes, Skills and Actions) or higher.

Third component: Hierarchical descriptions of performance including absolute thresholds or 'hard bars' to get over (or not)

Source: Oakden, J., King, J. & Allen, W. (2014). *Evaluation of the Sustainable Farming Fund: Main Report, prepared for Ministry for Primary Industries*. Wellington, New Zealand: Kinnect Group. (p.57)

© Pragmatica Limited 2018

What do we notice about this kind of rubric?

Benefits

- Helps everyone get clear about what aspects of performance are being focussed on

> "Its most appropriate when the quality or value of the evaluand is experienced as an entire package"
>
> (Davidson (2005) p 104)

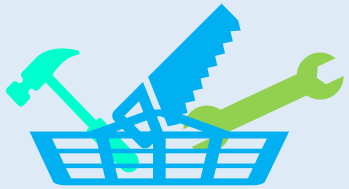
- It is clear what constitutes a 'pass' overall, and specifies the deal breakers
- It is useful where the evaluator wants to make a claim such as the programme is good value for money, as the constituent parts that make up this claim are clearly articulated and agreed at the start of the evaluation.

Challenges

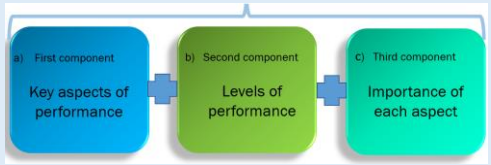
- It is difficult to pre-determine what might be important to capture, and can lock evaluators into aspects if the context changes
- Although this looks a simpler approach, I have found this type of rubric surprisingly hard to use well. In particular I find the evaluation reporting harder to develop using this kind of rubric
- BUT I know of others who find this approach very efficient and they say it works well for them.

© Pragmatica Limited 2018

Key takeaway points



- There are several different types of rubrics. If you use the right kind for a project they can be really helpful. But just like with builders' tools - if you choose a kind that is not 'fit for function' it may not work so well.



- To choose the right kind of rubric for your evaluation consider the three components of a rubric and how you combine them. There is no ONE right way, so experiment.

- Working with rubrics can be challenging – it is best to develop them with a diverse range of stakeholders, or other team members. Especially when starting out, work in a team.

- Making evaluative judgements using rubrics is a powerful way to share the emerging findings with the evaluation client and with stakeholders.

© Pragmatica Limited 2018

References

Eoyang, G., & Oakden, J. (2016). Adaptive Evaluation: synergy between complexity theory and evaluation practice. *Emergence: Complexity and Organization, 18*(3-4). doi:10.emerg/10.17357.e5389f5715a734817dfbeaf25ab335e5.

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

Davidson, E.J. (2014) *Evaluative Reasoning. Methodological Briefs: Impact Evaluation 4.* Florence: UNICEF Office of Research. www.unicef-irc.org/publications/pdf/brief_4_evaluativereasoning_eng.pdf

King, J., McKegg, K., Oakden, J., & Wehipeihana, N. (2013). Evaluative rubrics: A method for surfacing values and improving the credibility of evaluation. *Journal of MultiDisciplinary Evaluation, 9*(21) pp11–20.

Oakden, J., & Bear, C. (2011). Managing complexity in evaluation. Presented at the Aotearoa New Zealand Evaluation Association (ANZEA) Conference, 8—10 August, 2011, Wellington, New Zealand.

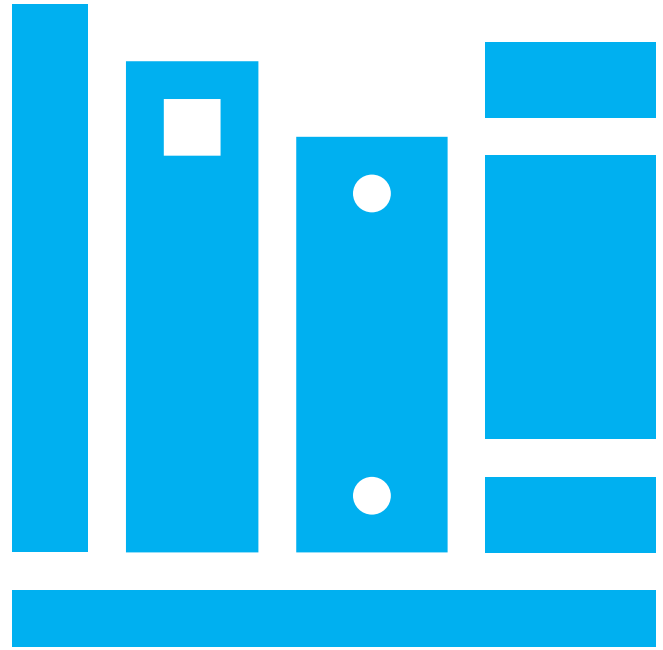
Oakden, J., & McKegg, K. (2011). *Waste Minimisation Act implementation: Evaluation of stakeholder perceptions.* Wellington. New Zealand: Kinnect Group.

Oakden, J. (2013). *Evaluation rubrics: How to ensure transparent and clear assessment that respects diverse lines of evidence*. Melbourne, Australia: Better Evaluation. Retrieved from http://betterevaluation.org/

Oakden, J., King, J. & Allen, W. (2014). *Evaluation of the Sustainable Farming Fund: Main Report, prepared for Ministry for Primary Industries*. Wellington, New Zealand: Kinnect Group.

Schwandt, T., (2015) How we think about quality in evaluation. Presented at American Evaluation Association, At Chicago, IL USA

Stone-Jovicich, S. (2015). *To rubrics or not to rubrics: An experience using rubrics for monitoring, evaluating and learning in a complex project.* Canberra, Australia: Commonwealth Scientific and Industrial Research Organisation. DOI: 10.13140/RG.2.1.3370.8565



© Pragmatica Limited 2018

Contact details

For further information or support in using evaluative rubrics

Judy Oakden
Director
Pragmatica Ltd
Kinnect Group member
Email: judy.oakden@gmail.com
Mob: +64 (0)27 479 8070
www.pragmatic.nz
www.kinnect.co.nz

To cite this resource:

Oakden, J. (2018.) *Understanding the components of evaluative rubrics and how to combine them.* Wellington, New Zealand: Pragmatica Limited

© Pragmatica Limited 2018