Developing defensible evaluation criteria: A tool for practitioners

AES Conference Sydney,
17 September 2019

Mathea Roorda
mathea.roorda@unimelb.edu.au
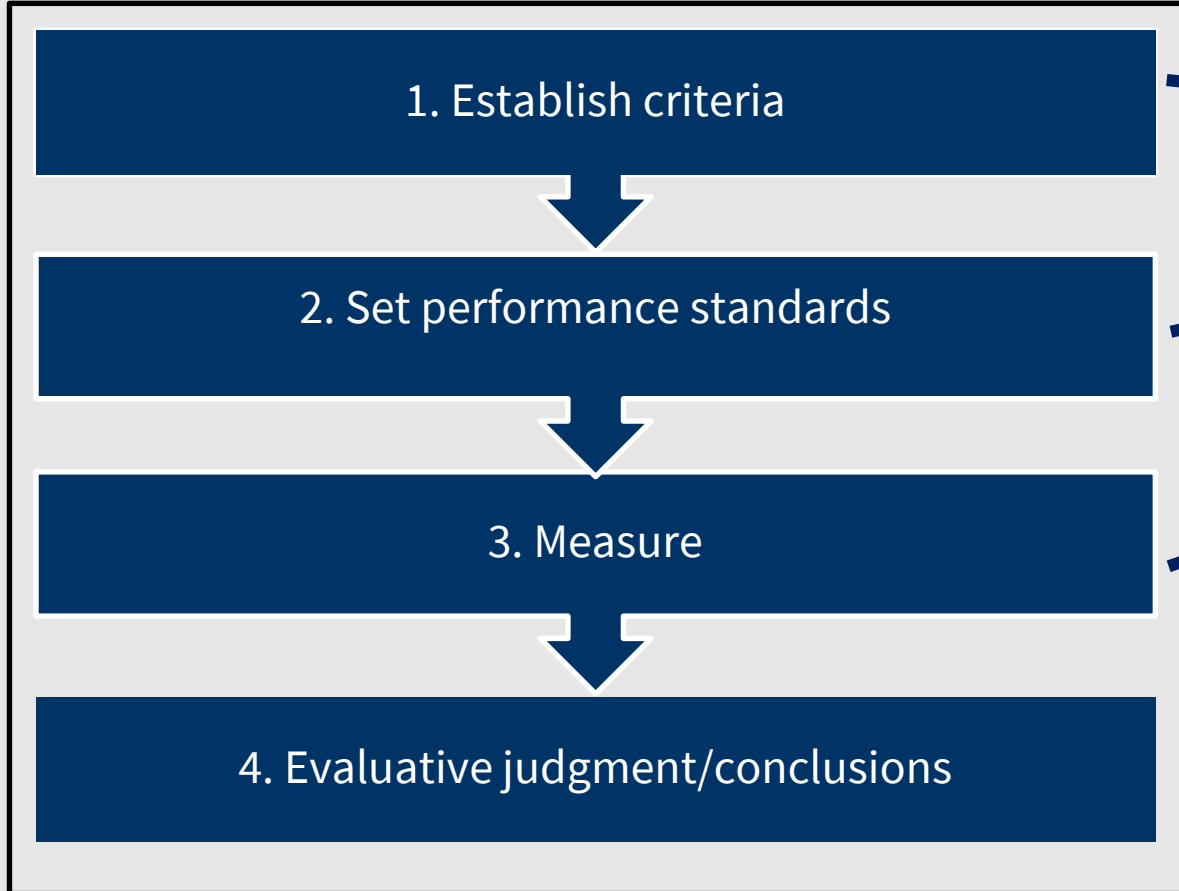mroorda@allenandclarke.co.nz

Introduction → Current practice → Criteria framework → Using the framework

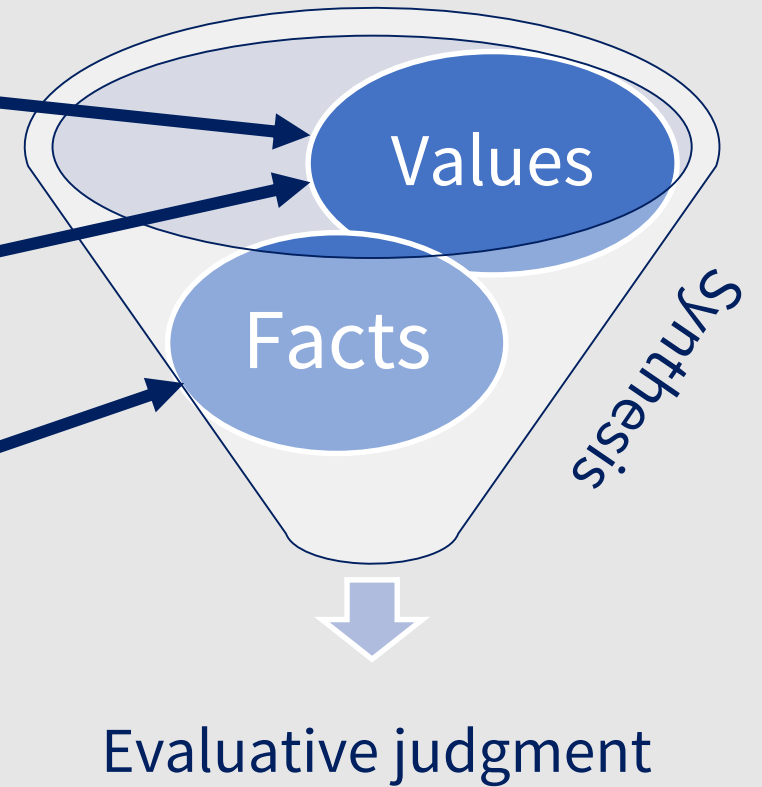**Evaluative judgements are grounded in**
– defensible empirical evidence
– argument, and
– anchored in chosen values

(Greene, 2011)

Logic of evaluation

Evaluative reasoning

1. Establish criteria

2. Set performance standards

3. Measure

4. Evaluative judgment/conclusions
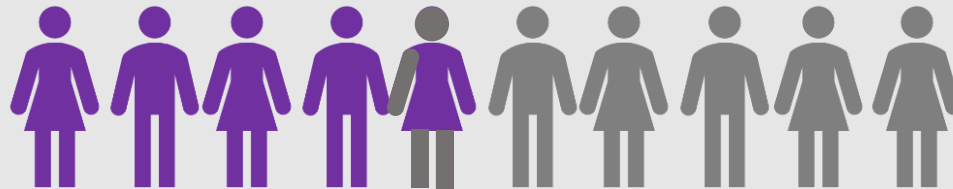
Values

Facts

Synthesis

Evaluative judgment

Evaluation persuades rather than convinces, argues rather than demonstrates, is credible rather than certain…

Ernest R. House, *Evaluating with validity* (1980)
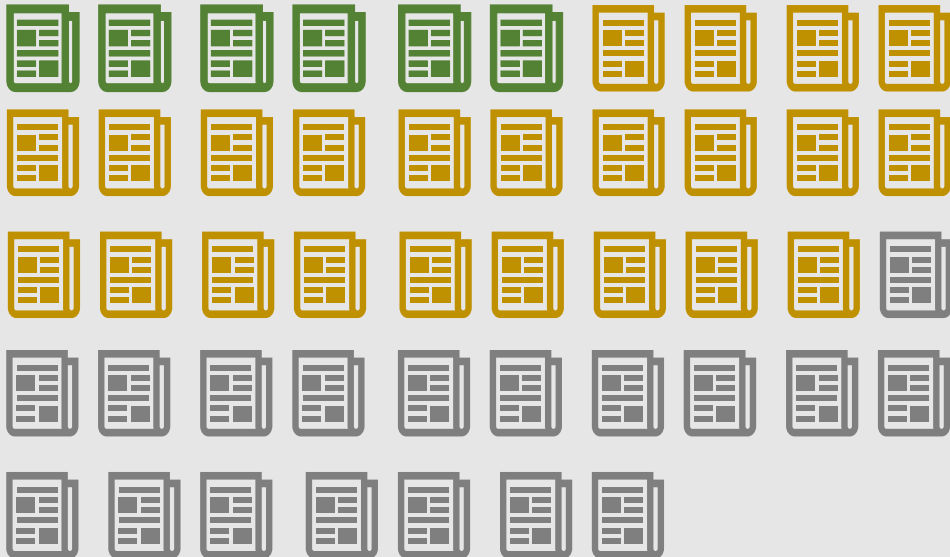
Explicit criteria are not the norm

45% of evaluators say they assess against criteria

Source: Survey of Australian and New Zealand public sector evaluators (n=137)

In a review of 47 evaluation reports, only six included explicit criteria

06 **Explicit criteria**

23 **Implicit in key evaluation questions (KEQs)**

18 **Criteria implicit elsewhere**

Source: Review of published Australian & New Zealand evaluation reports (Roorda no date)

Example of an implicit criterion

How and to what extent has each of the pilots influenced the local and wider ... sectors?

| (i) Demonstrated strategies to increase access to early childhood education | (ii) Delivery of best practice | (iii) Collaboration with research bodies | (iv) Building workforce capacity |

"Evaluators have often been unreflective, and even sloppy, in their approaches to valuing."

(Julnes, 2012, p. 4)

Prescriptive valuing
Asserts that certain values should be the criteria by which we judge the evaluand

Descriptive valuing
Describing the values held by interest groups about the evaluand, supported by trustworthy sources

	Normative perspectives				
	Consequentialist	Deontological			Ethic of care
	Maximum benefits	*Duty: Meeting obligations*	*Rights: Maximum protection*	*Equity: Being fair, equitable*	*Caring*
Interest Group 1					
Interest Group 2		Dimensions of merit			
Interest group 3					

Informed by Mepham et al. (2006) and Newman & Brown (1996)

Example: developing criteria for an evaluation of a curriculum

	Normative perspectives				
	Maximising outcomes	*Rights*	*Fairness*	*Duty*	*Ethic of care*
Students					
Teachers					
School leaders					
Program owner					

	Maximising outcomes	Rights	Fairness	Duty	Ethic of care
Students	Strong/ positive/ enhanced academic outcomes in program	Receive basic/ adequate education	Equal opportunities for all kinds of students, i.e. the program can't be a curriculum that only some students do well at / succeed in	Students attend school	Students support other students to succeed in the program
Teachers	Curriculum planning and implementation is easier and more flexible for teachers.	Teachers receive sufficient leadership support to implement the curriculum	Curriculum needs to be implementable i.e. the program can't be a curriculum that only some schools (e.g. private, well-resourced schools) can deliver	Deliver the curriculum as designed. Teachers have agreed to work at schools that deliver X, therefore they have a duty to deliver the curriculum	Teachers look after and care for students' mental and emotional well-being as well as their academic progress.

	Maximising outcomes	Rights	Fairness	Duty	Ethic of care
School leaders	Curriculum planning & implementation is easier and more flexible for coordinators and school leaders.		Cost and resourcing is not a barrier to implementing the program	Structures and resources are in place so that teachers can implement program as designed.	School leaders are concerned for and care about teachers' mental and emotional well-being as well as the quality of their teaching.
Program owner	Improved teaching and learning at schools using the program			Duty to provide schools with adequate supports / training / resources and materials so they can reasonably understand and do what they are being asked to do	

(i) Curriculum effectiveness

(ii) Equity

(iii) Implementation effectiveness

(iv) Quality of life

Relationships

Delivery of program (by program implementors) is supportive and respectful.

Teacher/school leader interactions with one another are supportive and respectful.

Balance

Curriculum respects student, teacher and school leaders' mental and emotional well-being.

Students, teachers and school leaders say they can "do [program]" while also having a positive work/life or school/life balance.

Summary of steps for developing defensible criteria

Trustworthy, authoritative sources

Literature
Experts
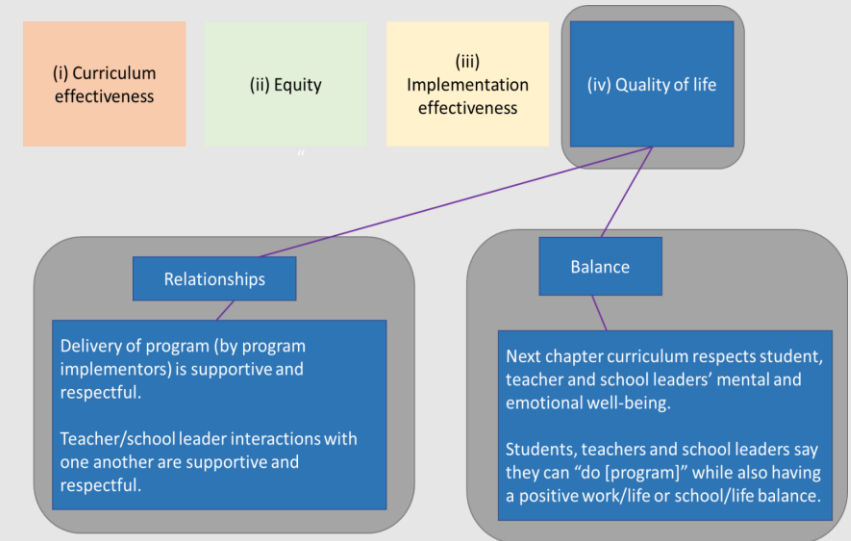Previous evaluations
Needs assessment
etc

Comprehensive description of relevant values

	Normative perspectives			
	Consequentialist	Deontological		Ethic of care
	Maximum benefits	Duty: Meeting obligations	Rights: Maximum protection	Equity: Being fair, equitable	Caring
Interest Group 1					
Interest Group 2					
Interest group 3					

Set of defensible evaluation criteria

(i) Curriculum effectiveness

(ii) Equity

(iii) Implementation effectiveness

(iv) Quality of life

Relationships

Delivery of program (by program implementors) is supportive and respectful.

Teacher/school leader interactions with one another are supportive and respectful.

Balance

Next chapter curriculum respects student, teacher and school leaders' mental and emotional well-being.

Students, teachers and school leaders say they can "do [program]" while also having a positive work/life or school/life balance.

Advantages

Includes the perspectives of all interest groups, not just the commissioner

Advantages

Includes the perspectives of all interest groups, not just the commissioner

Theory-informed, grounded in normative ethical perspectives

Advantages

Includes the perspectives of all interest groups, not just the commissioner

Theory-informed, grounded in normative ethical perspectives

Process is transparent – the evaluator can logically link values to criteria

Challenges

Prioritising criteria: How to manage competing values?

Challenges

Prioritising criteria: How to manage competing values?

Who gets to be in the 'driver's seat'?

Challenges

Prioritising criteria: How to manage competing values?

Who gets to be in the 'driver's seat'?

Potential for the framework to be used in a mechanistic way

Challenges

Prioritising criteria: How to manage competing values?

Who gets to be in the 'driver's seat'?

Potential for the framework to be used in a mechanistic way

Value theory – a gap in evaluator training?

References

Greene, J. C. 2011. *The Construct(Ion) of Validity as Argument*. New Directions for Evaluation 130:81–91.

House, E.R. (1980). *Evaluating with validity.* Sage Publications Inc., Beverley Hills.

Julnes, G. (2012). Managing valuation. *New Directions for Evaluation*, *133*, 3–15.

Mepham, B., Kaiser, M., Thorstensen, E., Tomkins, S., & Millar, K. (2006). *Ethical Matrix Manual,* (February), 45.

Newman, D. L., & Brown, R. D. (1996). *Applied ethics for program evaluation*. Thousand Oaks, CA, CA: Sage Publications, Inc.

Roorda, M. (2019). *Developing defensible evaluation criteria: A handbook for evaluators.* Wellington, NZ

Roorda, M. (no date). *Developing defensible evaluation criteria for Australian and New Zealand program evaluations* (draft PhD thesis)