Evaluating health services: a planning perspective

Introduction

- Like disaster planning, this is a field where people take entire courses, let alone subjects
- The purpose of this afternoon is to give you, as a planner, an overview of this really important area – in order to help you become a critical commissioner or consumer of evaluations
- To misquote a Chinese proverb (*The best time to plant a tree was 20 years ago. The second best time is now):*
    - The best time to start an evaluation is while you are developing an intervention (plan/program). The second best time is anytime after that.

Introduction

1. What is evaluation?
2. Why evaluate?
3. Types of evaluations
4. Judging a good evaluation

What is evaluation?

- WHAT IS/ISN'T EVALUATION?
- WHY EVALUATE?

Evaluation is …

- *… a systematic assessment of the design, implementation and outcomes of an intervention. It involves understanding how an intervention is being, or has been, implemented and what effects it has, for whom and why. It identifies what can be improved and estimates its overall impacts and cost-effectiveness.*

Evaluations

- *Most evaluations are conducted to answer two fundamental questions:*

  Is the program working as intended? [making a judgement/coming to a conclusion about whether a program or plan is reaching it's objectives and identifying any evidence of it having unintended or unwanted consequences]

  Why is this the case? [*helping program managers understand the reasons for program performance which can lead to the improvement or refinement of the program*]

- This involves an evaluation process which

  Defines clear and explicit criteria for success

  Collects representative evidence of program performance

  Compares this evidence to the criteria established as the outset [of the program]

Whatever the definition, evaluation always…

1. Is systematic
2. Values or in some way judges something
3. Is political (and raises ethical issues)

Evaluation isn't …

- Monitoring (descriptive rather than evaluative)
- Research

When baking a cake, I first research what kind of cake I want to make - say, a chocolate layer cake or plain vanilla. I research different recipes, and I choose one. I then set about making the cake, and I monitor what I put into it, such as how much flour or salt. I also monitor the cake to tell me how brown the top becomes, so that I know when to remove it from the oven. When evaluating my cake, I decide if the cake is good or bad, based on my preference, and sometimes I even listen to what others think of my cake

Why evaluate?

"However beautiful a strategy, you should occasionally look at results"

Winston Churchill

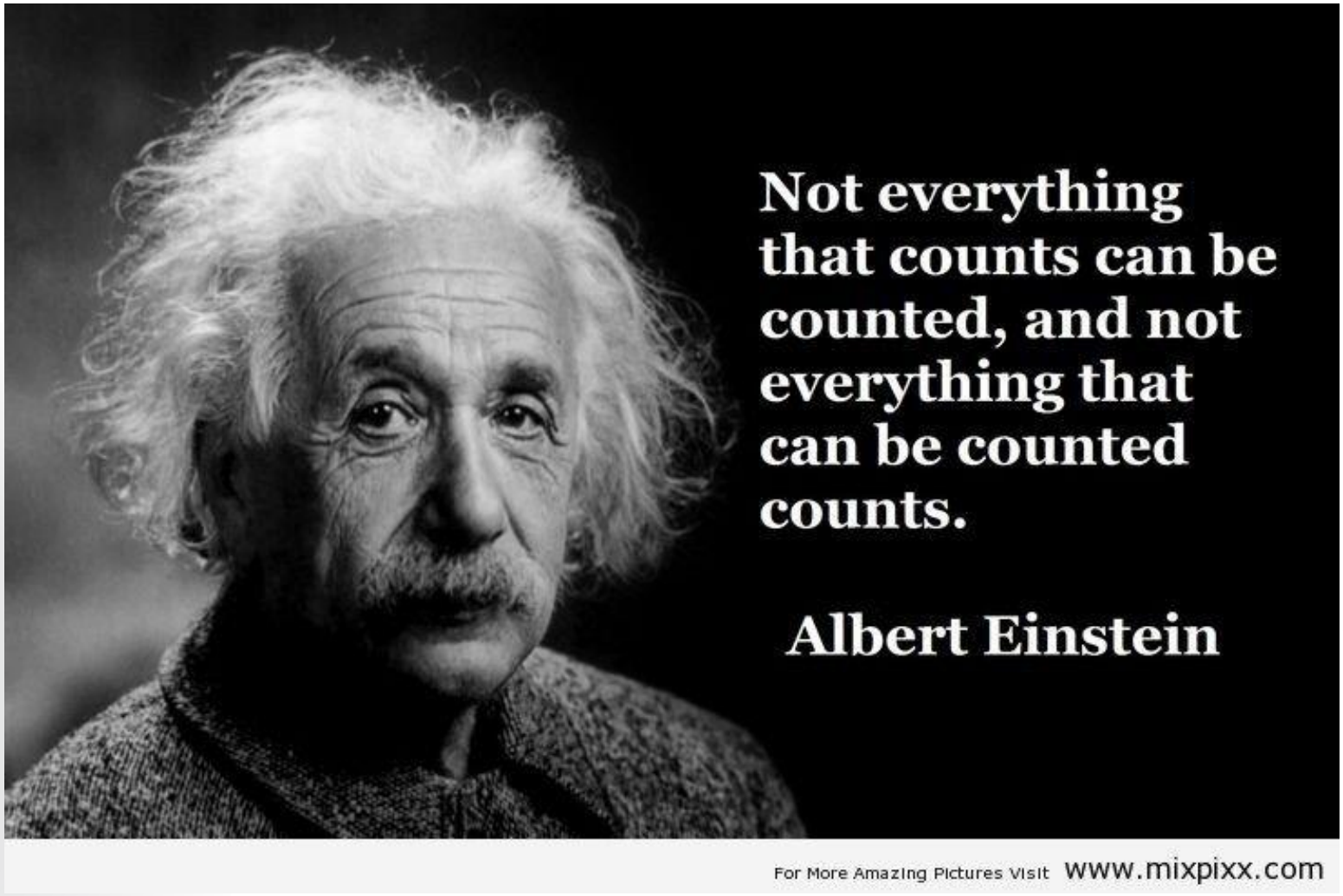
Why evaluate?

- *Contribute to broader evidence base: inform future policy and practice by others outside the organization*
- *Inform decision making aimed at improvement (formative): changing or confirming policies and practices*
- *Inform decision making aimed at selection, continuation or termination (summative): identifying best value for money*
- *Lobby and advocate: justify expenditure and demonstrate achievements*
- *Build trust and legitimacy across stakeholders: develop better understandings of each other and demonstrate that expectations are being met*
- *Ensure accountability: holding someone to account to someone for something*
- *Ensure diverse perspectives are included, especially those with little voice: make explicit the experiences and values of key stakeholders, especially intended beneficiaries*

What makes evaluation of health plans/ programs/ interventions difficult?

- *The difficulties associated with a move towards an outcome approach (from output management) in public management lie principally in demonstrating impacts in complex environments.*
- *This is particularly the case when professionals in the public sector work within a multi-partnership policy landscape – such as in the public health area – whereby outcomes cannot only be reported in quantitative terms (i.e. not all outcomes are ''hard'' in that they can be measured by using official statistics on, say, changes in morbidity and mortality levels amongst population groups).*
- *Rather, public managers are also expected to be able to report on ''softer'' outcomes such as changes in skill sets, knowledge and increased organisational capacities.*
- *There are also challenges in terms of articulating the direct and indirect influence of particular programmes and initiatives in impacting on outcomes – known as the*

And …


Not everything that counts can be counted, and not everything that can be counted counts.
Albert Einstein
For More Amazing Pictures Visit www.mixpixx.com

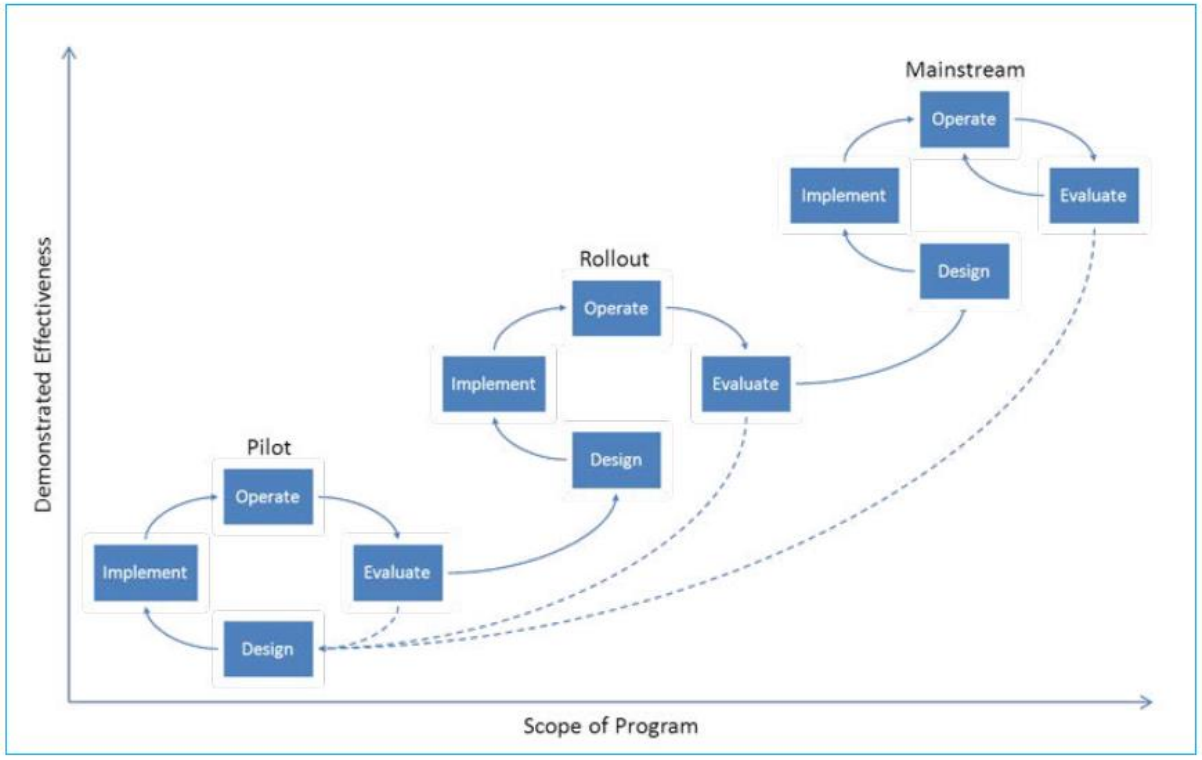
‘Evaluability’assessments

- *An evaluability assessment ‘‘is a process for clarifying program designs, exploring program reality, and - if necessary - helping redesign programs to ensure that ... program goals, objectives, important side effects, and priority information needs are well defined, program goals and objectives are plausible, relevant performance data can be obtained, and the intended users of the evaluation results have agreed on how they will use the information’’ (Wholey 1994)*

And evaluative thinking

- *Evaluative thinking is a critical part of being a thoughtful and engaged* [planner or manager]. *Thinking evaluatively means using data and information to be reflective, to learn, and to inform actions.*

Figure 2. Evaluation in the (ideal) program lifecycle



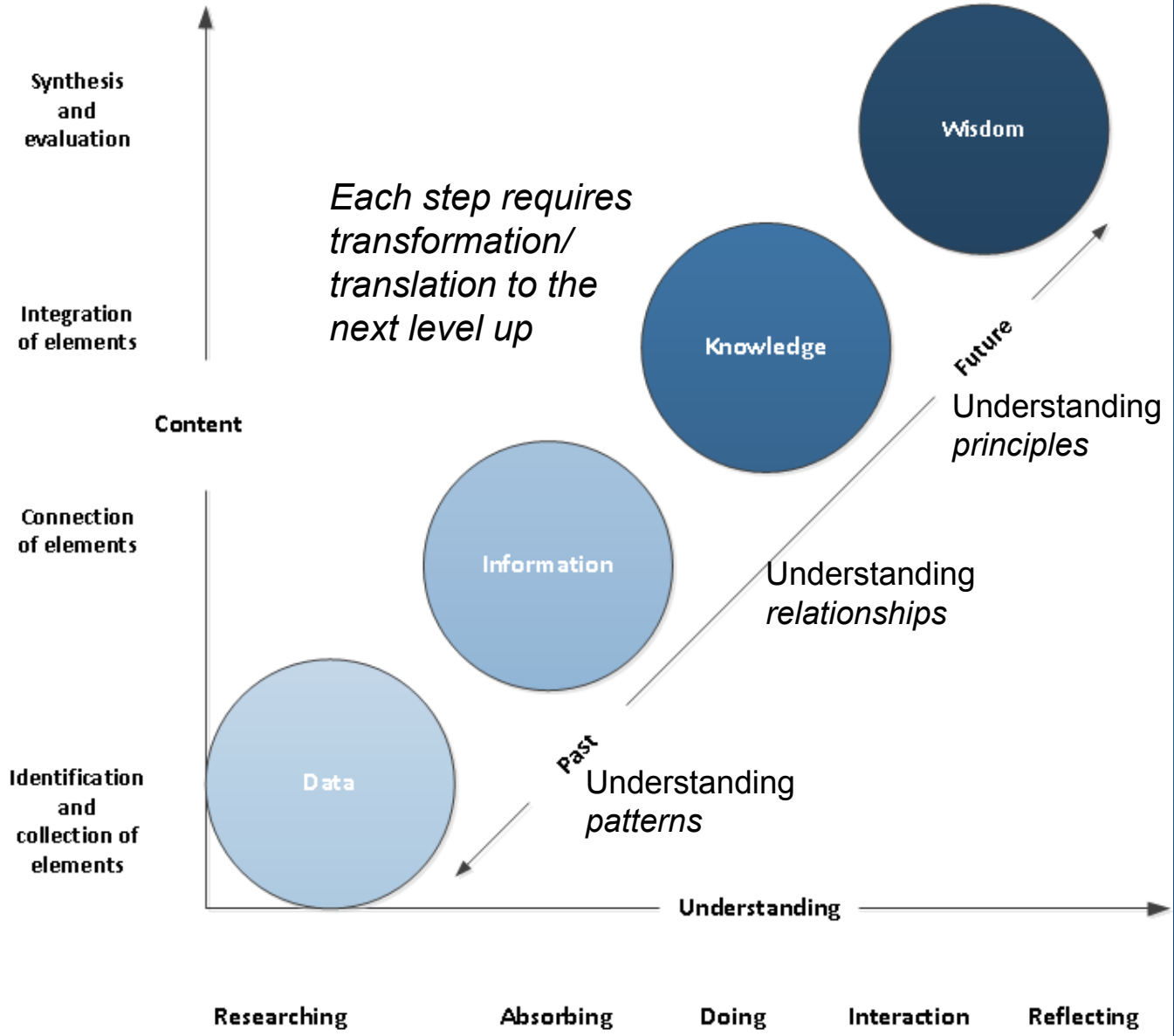
When to do an evaluation? Ideally …

Just a little side-step

WHAT ARE DATA/INFORMATION/KNOWLEDGE?

What is it we are analysing?

- **Data:** are collections of discrete facts, figures, statistics: the building blocks
- **Information:** are data with a meaning or context (organised data – the who, what, when and where)
- **Knowledge:** is applied information (contextual, relevant, actionable)
- **Wisdom:** (after that: understanding, discernment, judgement, enlightenment …)



Modified from: http://www.nwlink.com/~donclark/knowledge/context.html#join

What is knowledge (JoHARI Window)?

Luft, J., & Ingham, H. (1961). The Johari Window: a graphic model of awareness in interpersonal relations. *Human Relations Training News*, 5(9), 6-7.

	Known to you	Unknown to you
Know to others	**Arena** (public domain We all know what we know)	**Blind spot** (I don't know what I don't know)
Unknown to others	**Façade** (Avoided, hidden You don't know what I know)	**Unknown** (None of us knows!)

The conventional view serves to protect us

from the painful job of thinking

John Kenneth Galbraith

Ok, now back to evaluation …

Key evaluation questions

- *What will be evaluated? (That is, what is the program and in what context does it exist?)*
- *What aspects of the program will be considered when judging program performance?*
- *What standards (i.e., type or level of performance) must be reached for the program to be considered successful?*
- *What evidence will be used to indicate how the program has performed?*
- *What conclusions regarding program performance are justified by comparing the available evidence to the selected standards?*
- *How will the lessons learned from the inquiry be used to improve public health effectiveness?*

Key value (was it worth it?) questions

- Questions regarding values, in contrast with those regarding facts, generally involve three interrelated issues:
  - merit (i.e., quality)
  - worth (i.e., cost effectiveness)
  - significance (i.e., importance)

(One) Program evaluation framework: The CDC framework for program evaluation in public health

CDC evaluation standards

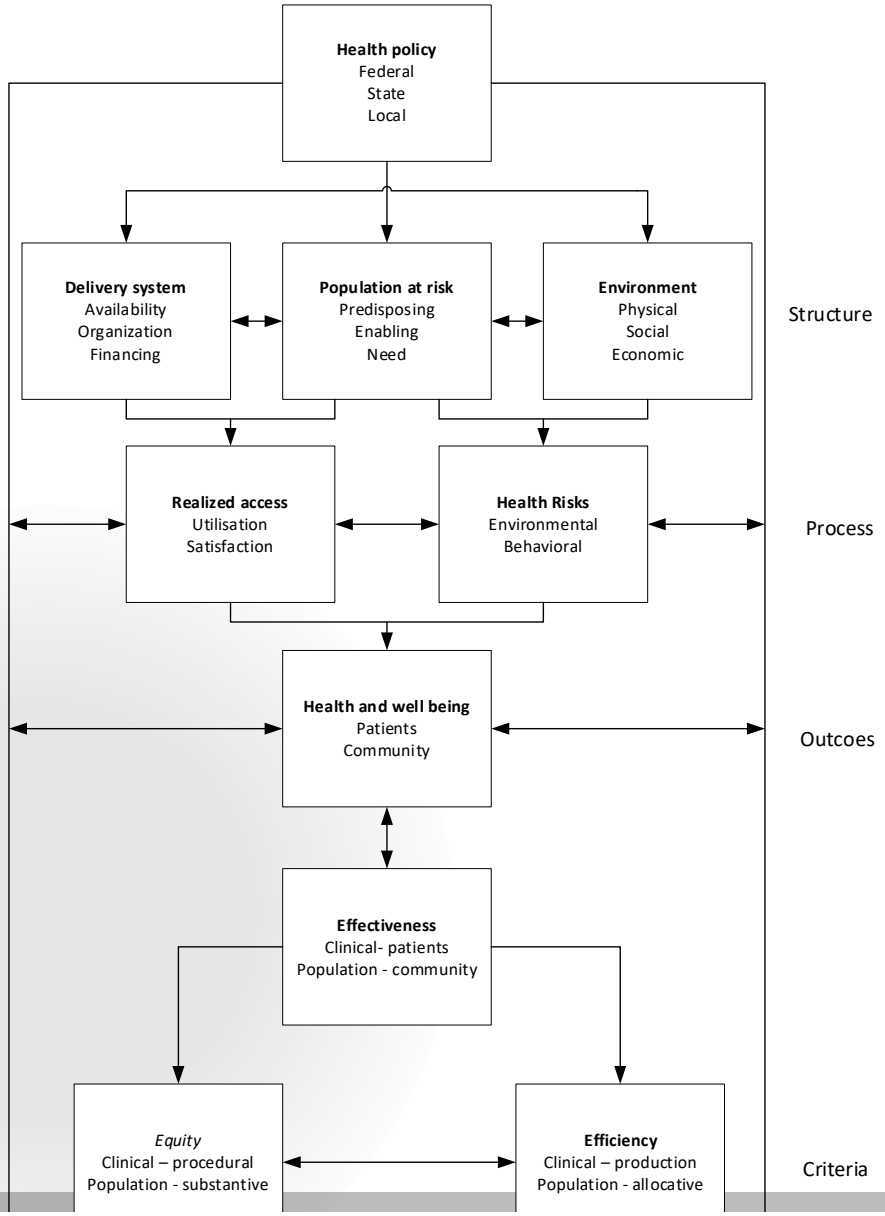
- ***Utility:*** *Who needs the evaluation results? Will the evaluation provide relevant information in a timely manner for them?*
- ***Feasibility:*** *Are the planned evaluation activities realistic given the time, resources, and expertise at hand?*
- ***Propriety:*** *Does the evaluation protect the rights of individuals and protect the welfare of those involved? Does it engage those most directly affected by the program and changes in the program, such as participants or the surrounding community?*
- ***Accuracy:*** *Will the evaluation produce findings that are valid and reliable, given the needs of those who will use the results?*

Stakeholder participation in evaluation

- That stakeholders should share influence - over what is to be studied, how, and in whose interests - signals the departure of participatory evaluation from the exclusive domain of scientific management.
- Intended as a tool for transformation, the potential of participatory evaluation stems from its democratic base; it requires a sharing of power, and stimulates a strengthening of the analytical capabilities of all participating stakeholders.
- It encourages mutual understanding and appreciation of different perspectives, and that in turn can be the precursor for both intellectual transformation and social action.

Describing the program

- Need (the program is intended to meet)
- Expected effects
- Activities
- Resources
- Stage of development (how long has the program been running for at the time of the evaluation, what changes – if any – has it gone through)
- Context (including setting(s) and environmental influences)
- Logic model (we will come back to this shortly)



Aday, L. A. (2004). *Evaluating the healthcare system: effectiveness, efficiency, and equity*. Chicago: Health Administration Press.

Evaluation types

Types of evaluations

Types	Definition	Methods used
Formative	Assesses suitability of resources and program elements Needs assessment/diagnostic evaluation	Pretest surveys, focus groups
Process	Examines outputs and the processes used to create them	
Delivery	Determines whether strategies delivered as planned, using methods and materials as designed	Audit against action plan, document review, project journals
Reach	Rates of participant recruitment and participation	Attendance records, participant surveys
Exposure	Determines if participants were aware of health issue being addressed and received program components	Participant surveys, in-depth interviews, focus groups, case studies
Context	Describes the factors that influenced the quality of program implementation	In-depth interviews, focus groups, case studies
Summative		
Outcome	Determines if program has successfully achieved goals	Interviews, comparative data (with baseline)
Impact	Assesses short- and medium-term impacts of strategies	Prepost surveys, post-only surveys, interviews, focus groups, audits, partnership review, case studies, RCTs

Formative evaluations

- Investigates how the program is delivered, including efficiency, quality and customer satisfaction. May consider alternative delivery procedures. It can help to differentiate ineffective programs from failures of implementation. As an ongoing evaluative strategy, it can be used to continually improve programs by informing adjustments to delivery

Process evaluations (which can be either formative or summative)

Process evaluation examines what a program is, the activities undertaken, who receives services or other benefits, and the consistency with which it is implemented in terms of its design and across sites. Often it is undertaken for formative or program improvement processes

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). Evaluation: A systematic approach (8th ed.). Los Angeles: Sage, n.p.

Scriven differentiates between three types of process evaluation:

1. Process-Improvement evaluations;
2. Process-Judgment evaluation;
3. Outcome-Improvement evaluation

Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151-161.

What types of questions would you ask in a formative evaluation?

- *How many persons are receiving services?*
- *Are those receiving services members of the intended target population?*
- *Are they receiving the proper amount, type, and quality of services?*
- *Are there members of the target population who are not receiving services or subgroups within that population who are underrepresented among those receiving services?*
- *Are members of the target population aware of the program?*
- *Are necessary program functions being performed adequately?*
- *Is staffing sufficient in numbers and qualifications for the functions that must be performed?*
- *Is the program well organized?*
- *Do staff work well with one another?*
- *Does the program coordinate effectively with the other programs and agencies with which it must interact?*
- *Are resources, facilities, and funding adequate to support necessary program functions? Are resources used effectively and efficiently?*
- *Is the program implemented as designed?*
- *Does the program comply with requirements imposed by its governing board, funding agencies, or higher level administration?*
- *Does the program comply with applicable professional and legal standards? Do program operations or performancevary significantly between sites or locales?*
- *Are participants satisfied with their interactions with program personnel and procedures?*
- *Are participants satisfied with the services they receive?*
- *Do participants engage in appropriate follow-up behavior after service?"*

Summative evaluations (including outcome and impact evaluations)

- Determines whether the program caused demonstrable effects on specifically defined target outcomes. Identifies for whom, in what ways and in what circumstances the outcomes were achieved. Identifies unintended impacts (positive and negative). Examines the ways the program contributed to the outcomes, and the influence of other factors.

Activity: understanding the implications of different types of evaluations

IN GROUPS OF 3 TO 5 (NOT SITTING TOO CLOSE TOGETHER) IDENTIFY THE QUESTIONS ASSOCIATED WITH EACH FORM OF ASSESSMENT AND TRANSFER TO BUTCHER'S PAPER

But wait, there's more…

Economic evaluation

- Addresses questions of efficiency by standardising outcomes in terms of their dollar value to answer questions of value for money, cost-effectiveness and cost-benefit. These types of analyses can also be used in formative stages to compare different options.

Rapid evaluations

- *A set of techniques for putting trustworthy, actionable information in the hands of decision makers at critical moments*
- *A central issue in the use of rapid evaluation and assessment methods (REAM) is achieving a balance between speed and trustworthiness*
- *Real-time evaluations developed in the 1990s in response to the proliferation of humanitarian crises. The UNHCR needed a method for rapidly evaluating the effectiveness and impact of its operational responses to emergencies and ensuring that findings were "used as an immediate catalyst for organizational and operational change" (UNHCR, 2002, p. 43).*

McNall, M., & Foster-Fishman, P. G. (2007). Methods of rapid evaluation, assessment, and appraisal. American Journal of Evaluation, 28(2), 151-168, pg 152-153

Transformative evaluations

- *The transformative paradigm with its associated philosophical assumptions provides a framework for addressing inequality and injustice in society using culturally competent, mixed methods strategies.*

Developmental evaluation

- *… positions the evaluator as a part of a program's design and development process. The evaluator collects information and provides informal feedback o members of the design team and possibly organization members to help them perfect the program* [or plan] being designed before it is ready for beta or pilot testing.

Developmental evaluation

- *has the purpose of helping develop an innovation, intervention, or program. In developmental evaluation the evaluator typically becomes part of the program or innovation design team, fully participating in decisions and facilitating discussion about how to evaluate whatever happens.*
- *All team members together interpret evaluation findings, analyze implications, and apply results to the next stage of development.*
- *The evaluator becomes involved in improving the intervention and uses evaluative approaches to facilitate ongoing program, project, product, staff, and/ or organizational development.*
- *The evaluator's primary function in the team is to facilitate and elucidate team discussions by infusing evaluative questions, data, and logic, and to support data-based decision making in the developmental process. In this regard, developmental evaluation is analogous to research and development (R & D) units in which the evaluative perspective is internalized in and integrated into the operating unit.*
- *In playing the role of developmental evaluator, the evaluator helps make an intervention's development an R & D activity.*

Participatory evaluation

- *A defining feature of participatory evaluation is the active engagement of multiple stakeholders.*
- *Their various interests and needs generate the content and form of the evaluation process.*
- *That stakeholders should share influence - over what is to be studied, how, and in whose interests - signals the departure of participatory evaluation from the exclusive domain of scientific management.*
- *Intended as a tool for transformation, the potential of participatory evaluation stems from its democratic base; it requires a sharing of power, and stimulates a strengthening of the analytical capabilities of all participating stakeholders.*
- *It encourages mutual understanding and appreciation of different perspectives, and that in turn can be the precursor for both intellectual transformation and social action.*

Mathie, A., & Greene, J. C. (1997). Stakeholder participation in evaluation: how important is diversity? Evaluation and program planning, 20(3), 279-285, pg 279

Inclusive evaluation

- *… involves a systematic investigation of the merit or worth of a program or system for the purpose of reducing uncertainty in decision making and for facilitating positive social change for the least advantaged. Thus, inclusive evaluation is data based, but the data are generated from an inclusive list of stakeholders, with special efforts to include those who have been traditionally underrepresented. It does not exclude those who have been traditionally included in evaluations*

Mertens, D. M., & Wilson, A. T. (2018). *Program evaluation theory and practice*. New York: Guilford Publications.

Feminist evaluation

- *includes judgments of merit and worth, application of social science methods to determine effectiveness, and achievement of program goals as well as tools related to social justice for the oppressed, especially, although not exclusively, for women. Its central focus is on gender inequities that lead to social injustice. It uses a collaborative, inclusive process and captures multiple perspectives to bring about social change*

Complexity and evaluations

- Evaluations which explicitly acknowledge and try to take account of complexity acknowledge two things:
  1) Health (and social care) can be understood as complex adaptive systems
  2) Single cause single outcome interventions are rare

Complexity evaluations

- *A policy or programme is increasingly complex:*
  - *The more organisations and individuals involved*
  - *The more layers or levels of intervention involved*
  - *The more dynamic the environment*
  - *The greater diversity of opinion and views*

Complex systems

- *A good evaluation, carefully planned and managed, ca*
  greatly assist policy makers in understanding the challenge
  posed by complexity, and provide opportunities to anticipate
  and take steps to manage these challenges

Complex systems: are open systems, with many links, connections, relationships, levers and hubs

The Centre for the Evaluation of Complexity Across the Nexus (CECAN): https://www.cecan.ac.uk/

Complex system challenges	Linked to which features of complexity	Evaluation challenges	How to address in evaluation management
Multiple interactions and influences	A central feature of complex adaptive systems	Long, indirect causal chains linking inputs to impacts	Ensure appropriate evaluation approaches used
Systems may be in continual change, or may resist change	• Adaptation • Emergence and self organisation • Change over time • Domains of stability	• Objectives, design and data requirements may change over time • The programme may not be at a 'final state' when the evaluation comes to an end	• Agile management approach/regular review • 'Findings' projected forward using appropriate methods and reported with caveats
Openness to outside influences Context (and history) matters	• Open systems • Path dependency	• Hard to establish a clear boundary around the intervention • Difficult to standardise the intervention • Outcomes may vary from one context to another	• Engage stakeholders with local knowledge • Evaluation approach (and data collection) must include context and history
Multiple perspectives	Multiple actors and relationships	• Need data from multiple sources/informants	• Stakeholder involvement at all stages
The nature of the change is unpredictable Multiple causality	• Property non-linearity • Feedback loops • Levers and hubs • Tipping points • Domains of stability	• Evaluation plans may need to change to address emergence of unexpected features • New methods needed for causality and attribution	• Use experts in range of evaluation approaches • Wide range of different data sources needed to capture unpredicted features emerging
Complexity is difficult to communicate	• Features above are not widely understood • Uncertainty is difficult accept	• Difficulties in communicating methodology and findings	• Ensure alignment of stakeholder understanding • Regular feedback during evaluation

Realist evaluation

- Is a *useful method for reviewing complex social interventions involving a chain of processes (Pawson, Greenhalgh, Harvey, & Walshe, 2004). RS can help to synthesize qualitative, quantitative and/or mixed methods evidence from complex interventions (Pawson, 2002; Pawson et al., 2004*)

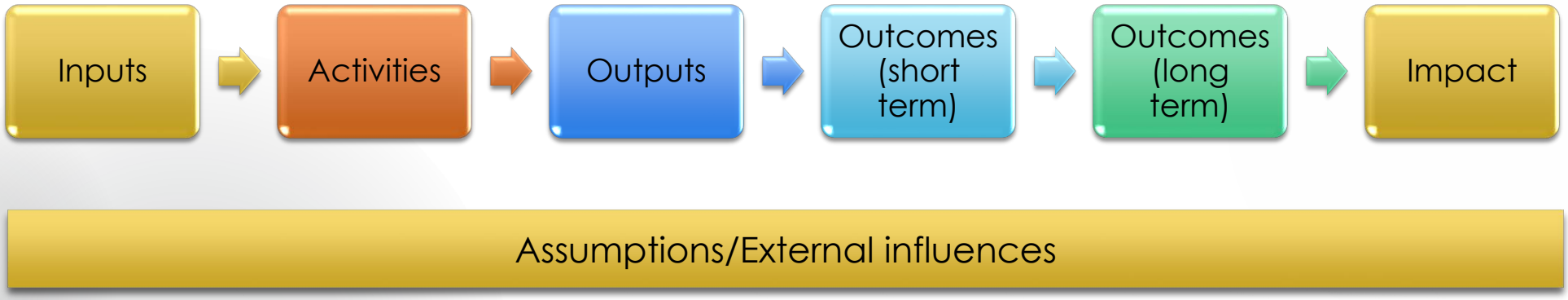
Logic models

A METHOD FOR APPROACHING BOTH PROGRAM DESIGN AND EVALUATION

Logic models

- Provide a map for the intervention/plan/activity against which the evaluation can be measures
- A logical model is "… *a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program, the activities you plan to do, and the changes or results you hope to achieve*" (WK Kellogg Foundation. (2004). WK Kellogg Foundation logic model development guide. Battle Creek, MI: WK Kellogg Foundation, pg 1)
- They are a way of considering the question "if the imitative (plan, program, activity) were successful, what would be different?"

Program logic models include



Activities

The things the program participants do with the resources to meet program objectives and interventions that are expected to produce changes or results.

Inputs

- The program inputs are all the resources that the program utilises including staff, facilities, funding, equipment and partnerships.
- Construction of the logic model allows evaluators and program staff to review the sufficiency of inputs to deliver the program outputs and achieve the intended outcomes.

Arguably the most common confusion in evaluation …

Outputs

Can include for example:

- Number of activities conducted
- Access to services
- Quantity of services
- Number and or types of services delivered to clients

Outcomes

Include those which occur

- Initial
- Intermediate (eg behaviour change as a result of training)
- Long term

Outputs

- *The logic model construction works backwards from outcomes to describe the outputs that will be necessary to achieve the stated outcomes. Outputs may be activities, products or services and describe the scope of the output and to whom it will be directed and implemented (Gugiu & Rodríguez-Campos, 2007). McLaughlin and Jordan (1999) recommend that output activities be described in no more than five groupings.*
- *Thus, outputs that achieve a common outcome can be grouped together to ensure that the logic model serves its purpose of simplifying the often complex schedule of program activities*

Outcomes

- *The first component of the logic model is the outcomes, which represents the program goals and are developed as a direct response to addressing a specific area of need as identified by the problem definition.*
- *Typically these are described as short-term and long-term outcomes with the requirement that short-term outcomes are able to generate the long-term outcomes.*

Outcome measures

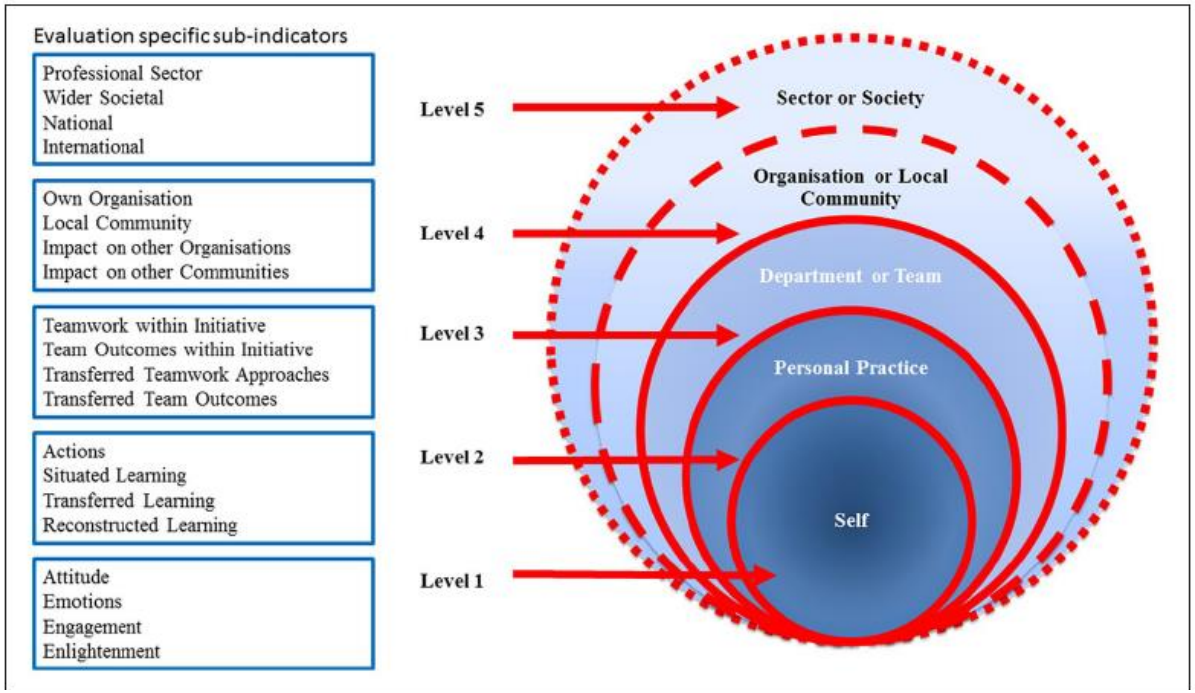
- Evaluations can identify:
  - failures of theory (where the plan or program was implemented as intended but did not produce the expected or desired results) and/or
  - failures of implementation (where the program plan or program was not implemented as intended)

Evaluating training

- Kirkpatrick's model

  Going beyond 'happy sheets'

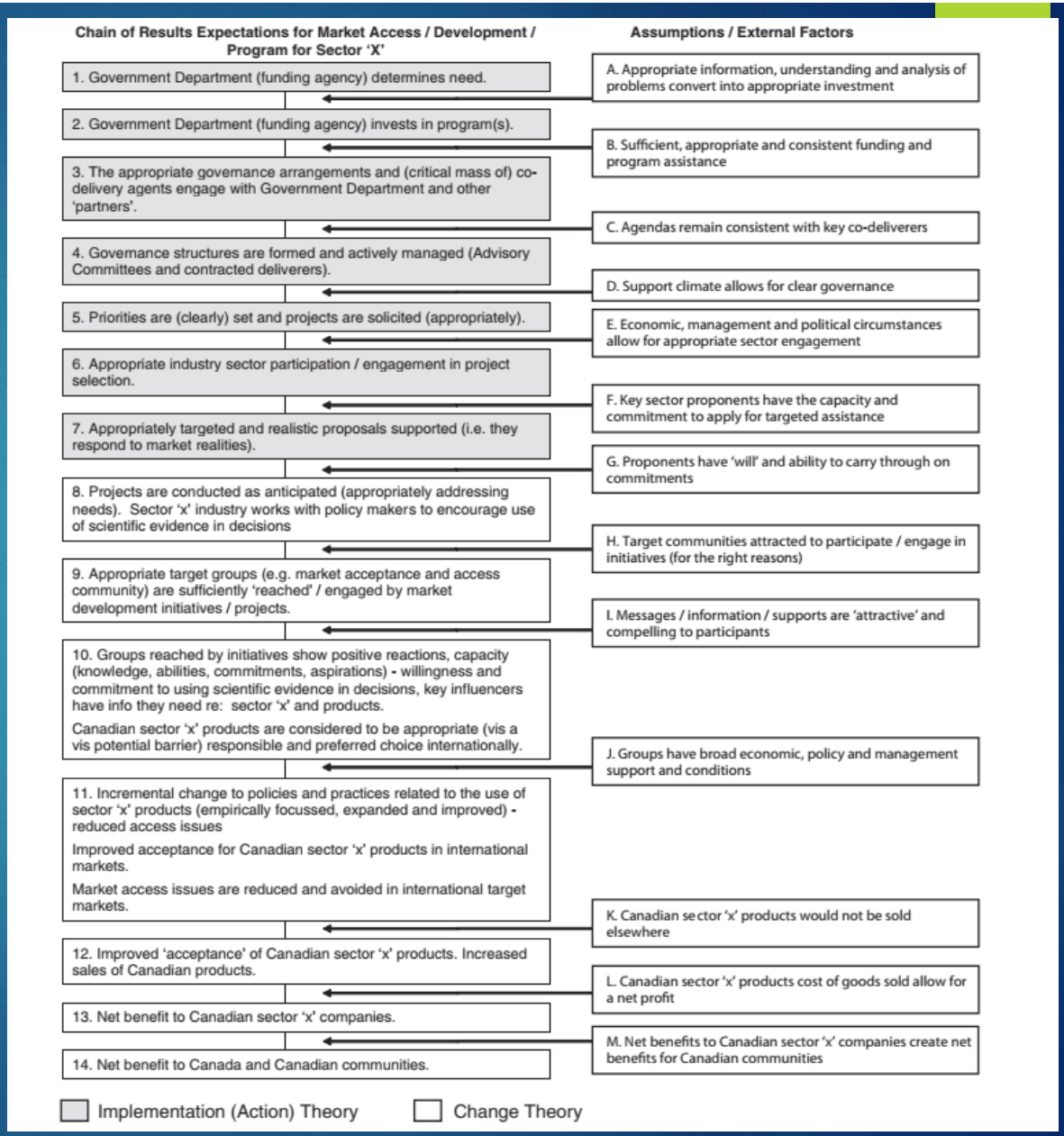
Evidence of impact



Assumptions

- *The underlying values, beliefs, and assumptions that influence the program's design, implementation, or goals.*

Looking at assumptions



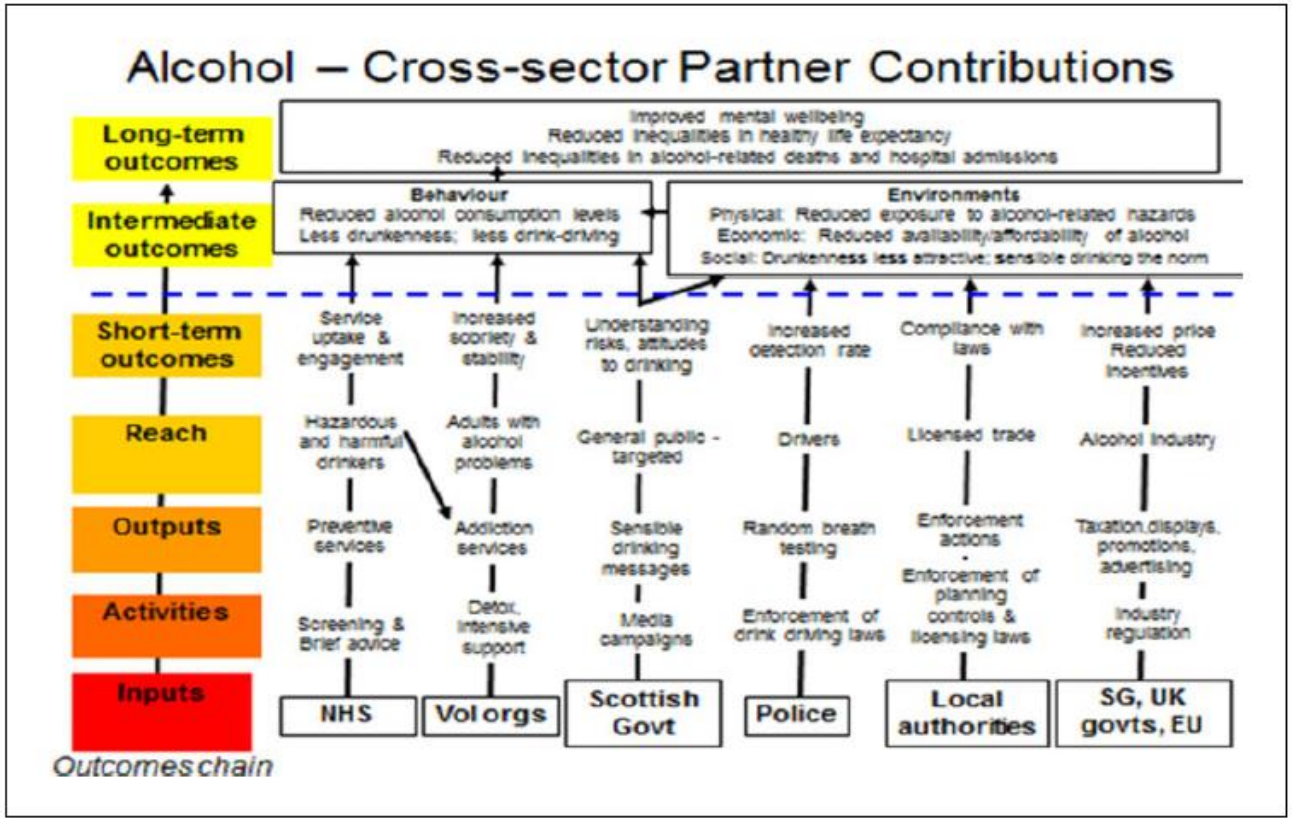
Wimbush, E., Montague, S., & Mulherin, T. (2012). Applications of contribution analysis to outcome planning and impact evaluation. *Evaluation, 18*(3), 310-329, pg 316

Example of logic model

Cullen, P., Clapham, K., Byrne, J., Hunter, K., Senserrick, T., Keay, L., & Ivers, R. (2016). The importance of context in logic model construction for a multi-site community-based Aboriginal driver licensing program. Evaluation and program planning, 57, 8-15, pg 6

Inputs		Outputs		Outcomes	
		Objectives	Implementation	Short-term outcomes	Long-term outcomes
1 Personnel 1A: Youth Workers 1B: Central Program Staff 1C: Program Steering Committee 1D: Local Steering Committees 1E: Investigative Committee 1F: Mentor driver volunteers 2 Host Organisation Facilities and infrastructure 3 Funding 4 Cars		**Organisation**			
	1 2 3	Facilitate professional development and education based upon identified training needs, career goal setting and mentoring	Training and supervising staff/volunteers, preparing funding proposals, participation in committees, presenting at conferences and to community groups	Increased staff skills in: research (data collection, ethics, interviewing, reporting, presenting), community engagement, establishing and maintaining professional networks	As for short-term
		Individuals and families			
	1A 1B 1F 2 3 4	Case management through the licensing system (assistance with ID, debt management, literacy assistance, supervised driving), financial assistance (payment of test, licensing fees, driving lessons, birth certificate fees)	Client progress through licensing system, regain licence after sanctions	Increased licence participation	Improved access to educational, employment, healthcare and social activities, better access to transport for family members
		Community			
	1A 1B 1D 1E 2 3 4	Community engagement to promote the value of licenced drivers in the community and the value of helping people through the licensing system, co-ordinate and attend community meetings to strengthen connections between existing services	Awareness and positive attitudes toward licensing, improved communication between local licensing services	Increased number of licenced and supervisor drivers, decreased licensing related sanctions and incarceration	Increased rates of licensing in participating communities, improved access to educational, employment, healthcare and social activities, decreased crash rates (unlicenced driver)
		Policy and licensing system			
	1 3 4	Promote the value of increasing licence participation, identify gaps in the licensing system, emphasise the need for culturally responsive services, liaise with and co-ordinate meetings between policy makers and licensing services	Awareness of licensing issues and barriers for Aboriginal people	Increased funding for licensing programs and better access to existing funding	Increased licence participation for Aboriginal people, decreased licence related sanctions and incarceration, decreased transport-related morbidity and mortality (unlicenced driver crashes, over-crowding due to lack of licenced drivers)

Fig. 1. Logic Model for Driving Change Licensing Support Program.



Multiple chains model

Activity 2

RETROFITTING A PROGRAM LOGIC MODEL

Program logic model

INPUTS

- Contract with NSW Rugby League
- (including one 0.5 FTE project officer)
- Sub-Contract Azion Wellness
- Participant Kits
- Develop Online Platform
- Teacher Support Resources
- Ministry of Health oversees implementation

ACTIVITIES

- Engagement of Knockout Health Challenge Town Committee, and other community engagement activities
- In service training for teachers, aides and Aboriginal Education Assistants delivered by NSW Rugby League
- Site visits to schools to support implementation
- Provision of participant kits with pedometer and logbook to all students
- Teachers and students use online platform for logging steps and 'virtual journey'
- Distribution of rewards for participation

OUTPUTS

- Children wear pedometers and accumulate steps (in school hours only)
- Steps and other activities recorded in the website for each participant
- Students participate and complete the online virtual journey
- Teachers support class participation by using teacher resources, integrating Activity Challenge into curriculum, and supporting students to log steps and participate in virtual journey
- Town Committee and broader community support Activity Challenge

OUTCOMES – SHORT TERM

- Increased engagement with communities participating in the 2014 Knockout Health Challenge
- Schools not engaged with Premiers Sporting Challenge are supported to increase physical activity of enrolled students
- An Aboriginal specific physical activity initiative is developed that can benefit all students

OUTCOMES – MEDIUM TERM

- Children decrease time spent in sedentary activity each day
- Children increase time spent being physically active each day

IMPACT – LONG TERM

- Improved health and wellbeing of Aboriginal children and families

DESIGNING EVALUATIONS

What you need to consider when designing or commissioning an evaluation …

- What specific questions do you want the evaluation answer?
- What sources of data do you have? What data collection methods do you want the evaluation to use?
- What type of relationship do you want the stakeholders have with the evaluation?

  Totally independent

  Co-design

  Participatory

  Empowerment

- What resources do you have to allocate to the evaluation?

Problem statements (for the intervention/plan/[program)

- *What is the problem that this intervention aims to address?*
- *What does the intervention intend to fix, change, or improve?*
- *Why (what were the reasons for) the intervention being developed?*
- *What gap (or gaps) was this intervention intended to fill?*

Activity 3

EVALUATION PURPOSE STATEMENTS

Problem statements (for the intervention/plan/[program)

Podems also notes that an '… *often forgotten group [of stakeholders] includes those who can be hurt, damaged or negatively influences by the intervention, while others (potentially or actually) benefit.*" She argues that evaluators who consider asking one, if not all three of the following questions:

- *Who or what can be harmed, hurt, upset , or damaged by the intervention?*
- *Who or what may have diminished power, benefits, or access to resources as a result of the intervention?*
- *Who thinks the intervention is damaging in some way, regardless of whether or not it is a reality (perceptions)?*

Evaluation methods

Qualitative

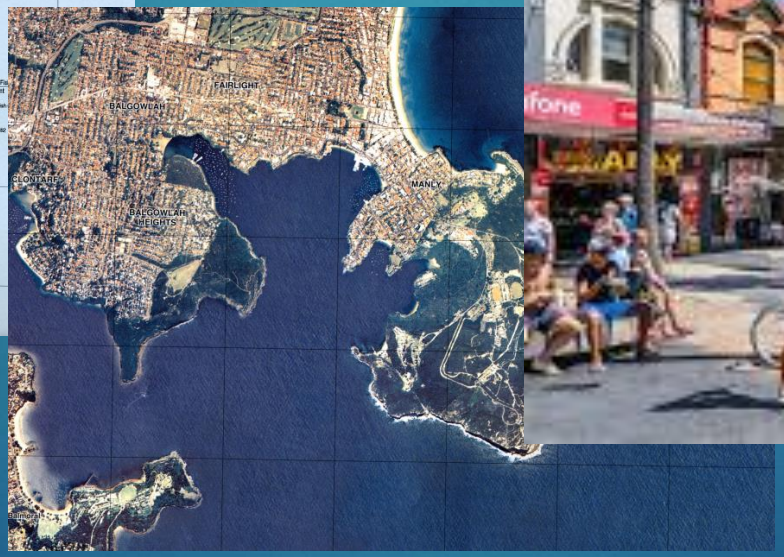
- Provides the 'flesh' on the bones of quantitative data
- Asks 'why' questions
- Provides an understanding of behaviours and motivations

Quantitative

- Provides the skeleton of an evaluation
- Asks the measurable questions (how many, how often etc)

Increasingly … Mixed Methods

Qualitative research adds



Qualitative research adds to our understanding of the context and complexity of issues and the worldview and experiences of those involved

Qualitative evaluation methods include …

- Archives
- Checklists
- Comparative Analysis
- Constant Comparative Method
- Content Analysis
- Cross-Case Analysis
- Deliberative Forums
- Delphi Technique
- Document Analysis
- Emergent Design
- Emic Perspective
- Ethnography
- Etic Perspective
- Fieldwork
- Focus Group
- Gendered Evaluation
- Grounded Theory
- Group Interview
- Key Informants
- Mixed Methods
- Narrative Analysis
- Natural Experiments
- Negative Cases
- Observation
- Participant Observation
- Phenomenography
- Portfolio
- Portrayal
- Qualitative Data
- Rapid Rural Appraisal
- Reflexivity
- Rival Interpretations
- Thick Description
- Think-Aloud Protocol
- Unique-Case Analysis
- Unobtrusive Measures

Quantitative evaluation methods include …

- Aggregate Matching
- Backward Mapping
- Benchmarking
- Concept Mapping
- Correlation
- Cross-Sectional Design
- Errors of Measurement
- Fault Tree Analysis
- Field Experiment
- Matrix Sampling
- Meta-analysis
- Multitrait-Multimethod Analysis
- Panel Studies
- Pre-Post Design
- Quantitative Data
- Quantitative Weight and Sum
- Regression Analysis
- Standardized Test
- Statistics
- Surveys
- Time Series Analysis

Judging a ‘good’ evaluation

What makes an evaluation a 'good' evaluation?

- Evaluation planning is most effective if it:
  - Is an integral element of the initial policy planning rather than an afterthought.
  - Brings together key actors and stakeholders
  - Is recognised as an important element by those involved.
  - Has a project manager who can stay with the process throughout.
  - Involves a clear and detailed brief to ensure clarity during any staff changes.
  - Ensures risk management procedures are in place to cover, for example, staff changes, sickness, or problems with data collection.
  - Takes account of the resources (including staff skills and time) available.
  - Allows sufficient time for completion and reporting.
  - Includes communication and dissemination plans.

Judging the quality of an evaluation

1. *Framing (Conceptual Framing)*
2. *Transparency (Replicability/Confirmability)*
3. *Appropriateness (Methodological Appropriateness)*
4. *Validity (Construct/Measurement Validity)*
5. *Credibility (Truth Value of Statements/Findings; Internal Validity)*
6. *Transferability (External Validity)*
7. *Robustness (Reliability/Dependability/Consistency/Stability)*
8. *Structure (Coherence, Limitations, etc.)*

The Centre for the Evaluation of Complexity Across the Nexus (CECAN): https://www.cecan.ac.uk/

A successful evaluation design will:

- *Begin with a clear aim and success criteria defined early on.*
- *Be explicit about why the evaluation is needed, who will be using it and how.*
- *Use a logic process/logic model to identify appropriate methods which are fit for purpose, rather than assume a certain method is intrinsically superior.*
- *Cover the whole process of implementing the policy – not only outcomes but how these were achieved.*
- *Be flexible enough to accommodate complexity while retaining clarity of aim.*
- *Aim to capture information at the most effective point(s) in time.*
- *Be open to different approaches to economic analysis, including qualitative as well as quantitative techniques.*
- *Draw on in depth case studies where appropriate to highlight important contextual factors.*
- *Be appropriate for integrating and building upon the work of the organisation and its wider stakeholder network*

As always the last word goes to Dilbert

Right data,
wrong conclusion

No matter how good the data (or evaluation), politics count