

Evaluation's Contribution to the Measurement of Research, Technology and Development: From Prescription to Menu to Guidelines

Presentation at the
Canadian Evaluation Society
Victoria, BC
May 4, 2010

steve.montague@pmn.net

The Situation

- Evaluation has influenced the performance measurement of RT&D
- Traditionally – a prescription has been offered
 - Inputs
 - Outputs
 - Publication productivity
 - Citations
 - Some ‘impacts’
- Problem: Mission is sometimes neglected

Some Background

- A short 'history' of RT&D indicators..
- Paper by Fred Gault, Statistics Canada (SPRU 2006) provides an excellent review of Canadian and international R&D and innovation statistics
 - Accepted input indicators include, GERD, R&D personnel, assumed to result in innovation (input / output model)
 - Output indicators link to IP (publications, patents) and new HQP also assumed to link directly to innovation, economic growth
 - Outcomes relate to increased innovation and need to consider a wide range of factors outside the direct influence of R&D
 - Recognition that the linear innovation model is simplistic and doesn't reflect sectoral differences, influence of other policies and international economic situation
- More recently we have seen these formalized as scorecards and 'payback' categories

Evaluating the Valuation Approaches

Examples:	Assessment:
<p>Aggregate Economic Statistics</p> <ul style="list-style-type: none">- Expenditures on R&D- GERD- Market share- Trade deficit in innovative goods	<ul style="list-style-type: none">▪ Not 'outcomes'▪ Attribution problems▪ Understanding?
<p>Micro-economic Benefit-Cost</p> <ul style="list-style-type: none">- Benefit-Cost Analysis- Cost-Effectiveness Analysis	<ul style="list-style-type: none">▪ Favours easily quantifiable 'outcomes'▪ Monetization problems▪ High sensitivity to assumptions
<p>Scientific Merit Valuation</p> <ul style="list-style-type: none">- Peer review- Publications- Citations	<ul style="list-style-type: none">▪ Partial 'outcomes'▪ Bias problems – science viewpoint▪ Misses the bigger system

Recent Developments

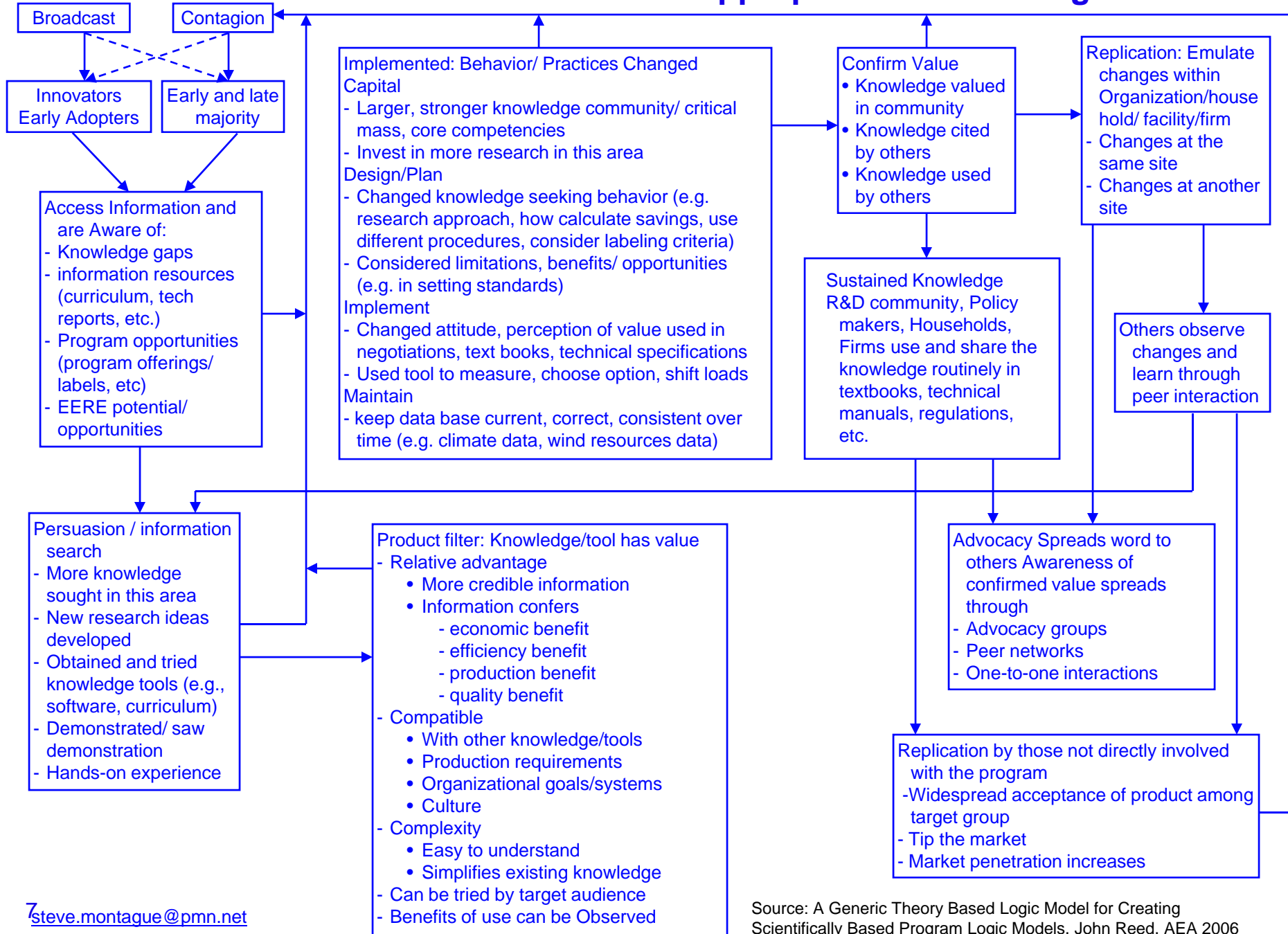
- Results logic to 'inform' indicators
- Sort the levels
- Suggest the connections

Case Impacts – A Structural Assessment

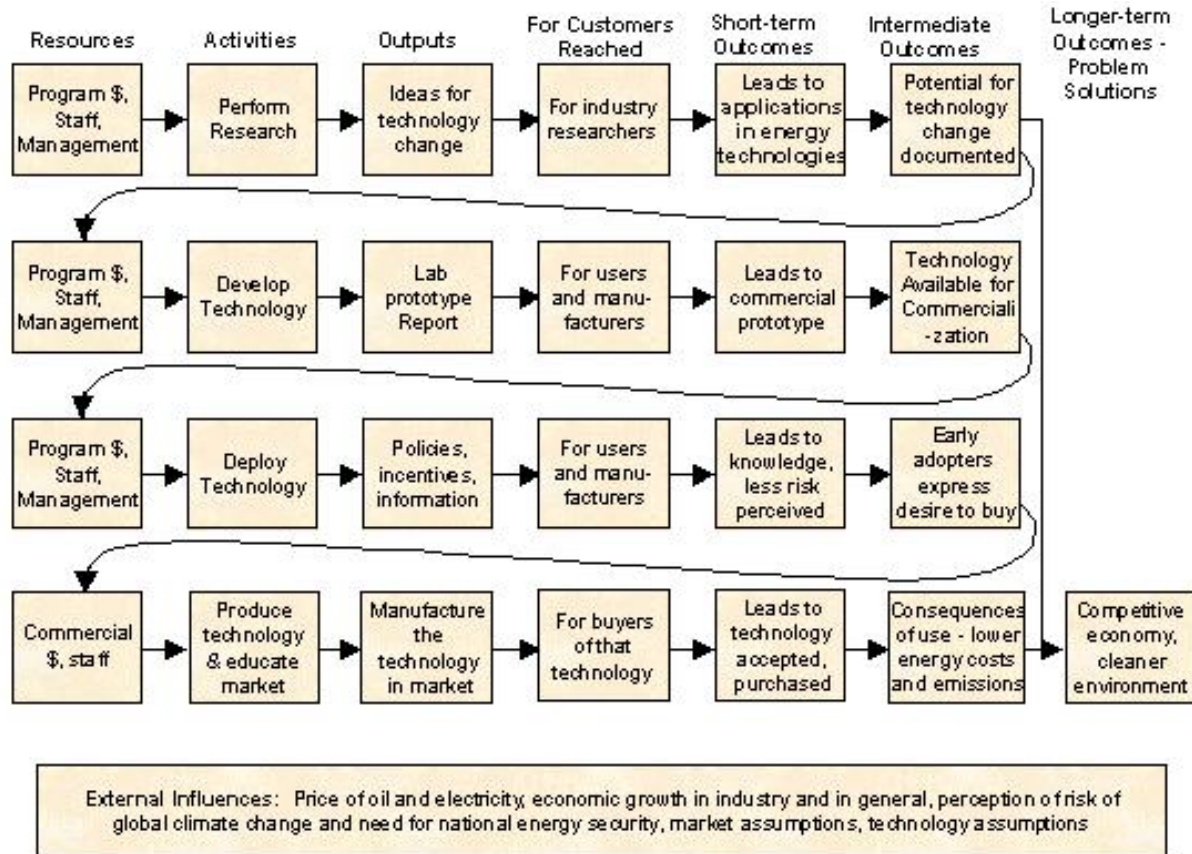
Project Incrementality (Direct Influence)	Direct Collaborator / Performer / User Impacts	Industry Sector / Supply Community Impacts	Economy / Societal Impacts
<p>Involvement</p> <ul style="list-style-type: none"> <input type="checkbox"/> project would not have been done without program assistance <input type="checkbox"/> contributed to completing the project more quickly <input type="checkbox"/> contributed to completing the project more thoroughly 	<p>Technical results</p> <ul style="list-style-type: none"> <input type="checkbox"/> new or improved product <input type="checkbox"/> new or improved process <input type="checkbox"/> advancement of knowledge <input type="checkbox"/> increased technical capabilities <input type="checkbox"/> improved quality control <input type="checkbox"/> new skills internally <input type="checkbox"/> increased efficiency / improved productivity <input type="checkbox"/> technology transfer <p>Policy / legislative results</p> <ul style="list-style-type: none"> <input type="checkbox"/> policy behavioural changes <input type="checkbox"/> agreement / accord <input type="checkbox"/> legislation / regulation <p>Infratechnology results</p> <ul style="list-style-type: none"> <input type="checkbox"/> Codes, standards, databases, protocols <input type="checkbox"/> acceptance of standards <p>Commercial results</p> <ul style="list-style-type: none"> <input type="checkbox"/> increased sales <input type="checkbox"/> increased market share <input type="checkbox"/> increased profitability <input type="checkbox"/> cost savings <p>Organizational effects</p> <ul style="list-style-type: none"> <input type="checkbox"/> increase in jobs <input type="checkbox"/> diversification <input type="checkbox"/> expansions <input type="checkbox"/> strategic alliances / partnerships <input type="checkbox"/> achievement awards / recognition 	<ul style="list-style-type: none"> <input type="checkbox"/> production process efficiencies <input type="checkbox"/> increased science and technology information <input type="checkbox"/> increased sales <input type="checkbox"/> cost savings <input type="checkbox"/> changes to industry structure (e.g., concentration, competitiveness internationally) <input type="checkbox"/> spin-off companies <input type="checkbox"/> technology infrastructure (e.g., standard scientific and engineering data, industry standards, test protocols, and instrumentation) <input type="checkbox"/> training of technological problem-solvers whose talents can be applied in many areas <input type="checkbox"/> establishment of quality, performance standards 	<p>Economic Benefits</p> <ul style="list-style-type: none"> <input type="checkbox"/> reduced consumer costs <input type="checkbox"/> increased employment <input type="checkbox"/> improved competitiveness <input type="checkbox"/> reduction in subsidies <p>Societal Benefits</p> <ul style="list-style-type: none"> <input type="checkbox"/> improved quality of life <input type="checkbox"/> protection of environment <input type="checkbox"/> improved energy efficiency <input type="checkbox"/> improved public health and safety <input type="checkbox"/> education / awareness <input type="checkbox"/> public service efficiency gains (i.e. lowered taxpayer burden)

* Note that this template attempts to integrate a wide range of impacts into a structured quantitative *and* qualitative assessment.

A Generic Version of the Diffusion Model Appropriate to Knowledge Entities

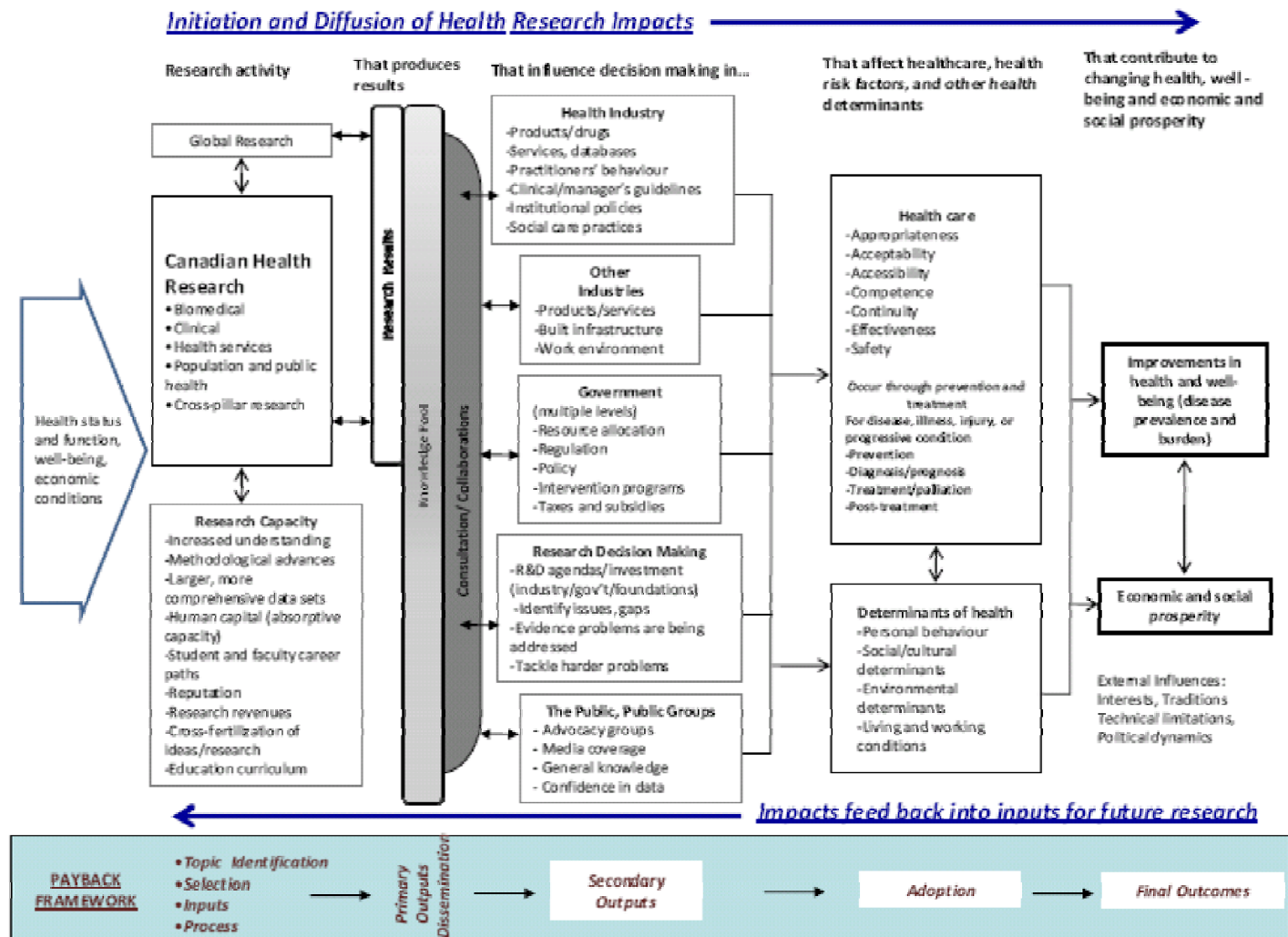


Logic Chart for a Research and Technology Development and Deployment Program



Source: Logic Models: A Tool for Telling Your Program's Performance Story . John A. McLaughlin and Getchen B. Jordan, July 1998

CAHS Framework Logic Model of Health Research Progression to Impacts



Source: Making an Impact, Canadian Academy of Health Sciences, Assessment Report, January 2009.

The Challenge: Can We Go Farther?

- Can we use results logic - or a simplified results hierarchy to sort RT&D indicators?
- Can an Innovation diffusion model help? (Rogers, Bennett)
- Can we increase the emphasis on learning.. As compared to 'justification'?

CAHS ROI Indicators

Over 60 Indicators of the following:

- *Advancing knowledge indicators and metrics include measures of research quality, activity, outreach and structure. Identified are some aspirational indicators of knowledge impacts using data that are highly desirable but currently difficult to collect and/or analyze (such as an expanded relative-citation impact that covers a greater range of publications, including book-to-book citations and relative download-rates per publication compared to a discipline benchmark).*
- *Research capacity-building indicators and metrics fall into subgroups that represent personnel (including aspirational indicators for improving receptor and absorptive capacity), additional research-activity funding and infrastructure.*
- *Informing decision-making indicators and metrics represent the pathways from research to its outcomes in health, wealth and well-being. They fall into health-related decision-making (where health is broadly defined to include health care, public health, social care, and other health-related decisions such as environmental health); research decision-making (how future health research is directed); health-products industry decision-making; and general public decision-making. Provided are two aspirational indicators for this category (media citation analysis and citation in public policy documents).*
- *Health-impact indicators and metrics include those on health status, determinants of health and health-system changes, and they include quality of life as an important component of improved health. Determinants of health indicators can be further classified into three major subcategories: modifiable risk factors, environmental determinants, and modifiable social determinants.*
- *Broad economic and social impacts are classified into activity, commercialization, health benefit (specific costs of implementing research findings in the broad health system), wellbeing, and social-benefit indicators (socio-economic benefits).*

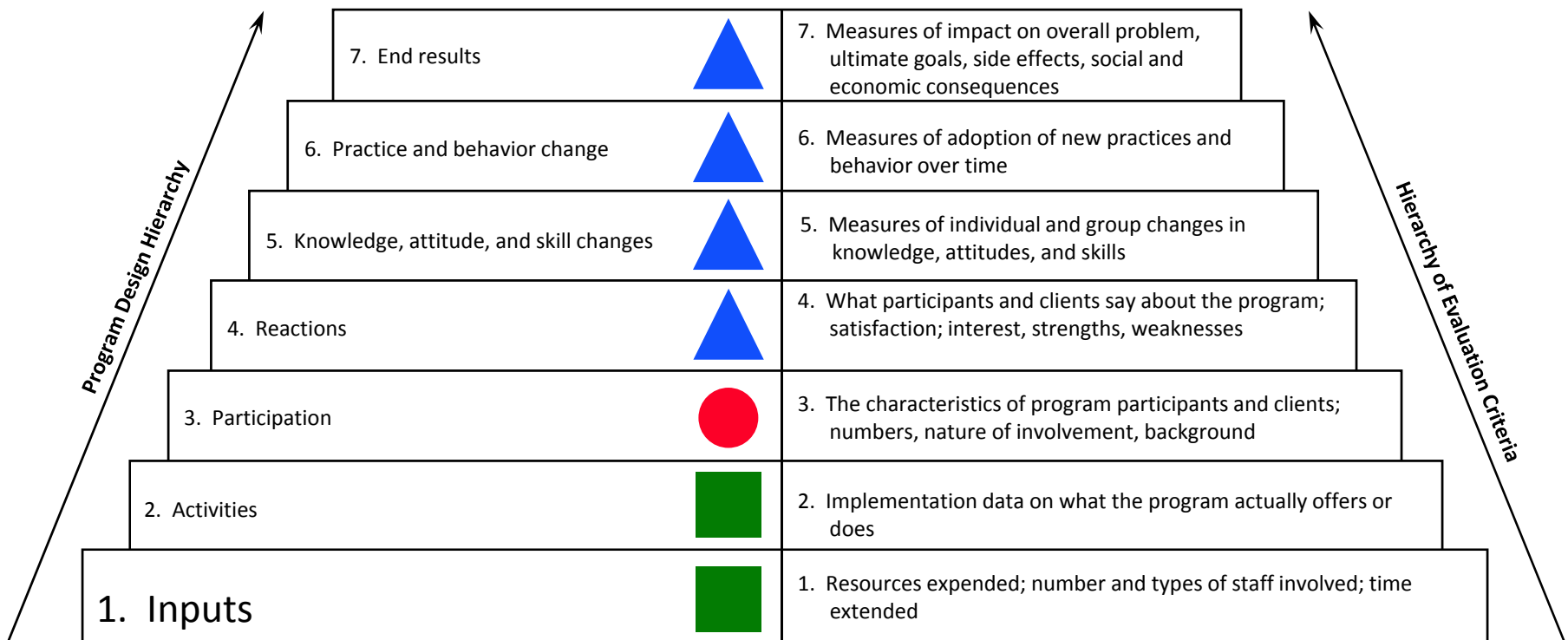
The RAND Payback Categories

Payback category	Payback	Example
Knowledge production	<ul style="list-style-type: none"> Peer-reviewed publications in the serial literature 	<ul style="list-style-type: none"> 302 papers receiving a total of 975 citations per year attributable to case studies
Research targeting and research capacity	<ul style="list-style-type: none"> Postgraduate research training Subsequent career development of PIs and research assistants The transfer of technical know-how Informing future research studies 	<ul style="list-style-type: none"> 28 PhD/MDs from work on the case studies Development of technological know-how in genetic mapping Informed >£2 million Medical Research Council (MRC) randomised controlled trial Use of biologicals as therapeutic targets
Informing policy and product development	<ul style="list-style-type: none"> Informing recommendations in clinical guidelines and other policy advice Informed development of clinical tests 	<ul style="list-style-type: none"> Recommendation in Royal College of Obstetricians and Gynaecologists (RCOG) guideline on the use of aspirin and heparin for women with antiphospholipid syndrome (APS) Recommendation in Industrial Injury Advisory Council (IIAC) assessment for hip osteoarthritis (hip OA) in farmers to be a prescribed disease Clinical test for a rare type of systemic lupus erythematosus (SLE) and chondrodysplasia type Schmidt
Health and health sector benefits	<ul style="list-style-type: none"> Improving the quality of life for people with rheumatoid arthritis (RA) Reducing the likelihood of recurrent miscarriages for women with APS 	<ul style="list-style-type: none"> Hundreds of thousands of patients treated with anti-TNF of whom 70% experience a significant improvement in health Use of aspirin and heparin for women with APS increases live birth rate by 40% compared to the use of aspirin alone and by 60% compared to no treatment at all.
Wider economic benefits	<ul style="list-style-type: none"> Unquantified economic returns resulting from a reduction in days off work and sales of licensed drugs 	

Theory of Action - Levels of Evidence and the Logic Model

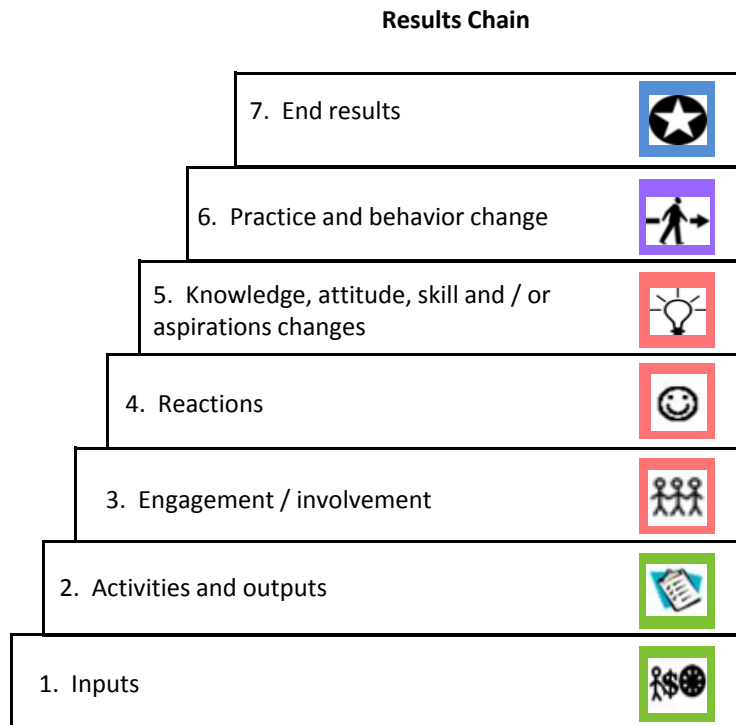
*Program Chain of Events
(Theory of Action)*

Matching Levels of Evidence



Source: Adapted from Claude Bennett 1979. Taken from Michael Quinn Patton, Utilization-Focused Evaluation: The New Century Text, Thousand Oaks, California, 1997, p 235.

Results Chains vs. RAND Case Categories










- RAND 'Payback' Scoring Areas**
- Final outcomes*
 - Wider economic sector benefits**
 - Health sector benefits**
 - Adoption by practitioners and public*
 - Informing policy and product development**
 - Knowledge production / capacity building**
 - Research targetting (*participation of groups*)**
 - Issue identification, Project selection, Research process*
 - Inputs to research*



*RAND *Returns From Arthritis Research* (2004) Section 2.1,
 ** Ibid Section 3.3.2

Note: These results categories can be tracked over time to tell a performance story.

A Results Chain Typical Indicator and Measurement “Menu” for CCSRI National Programs and Networks

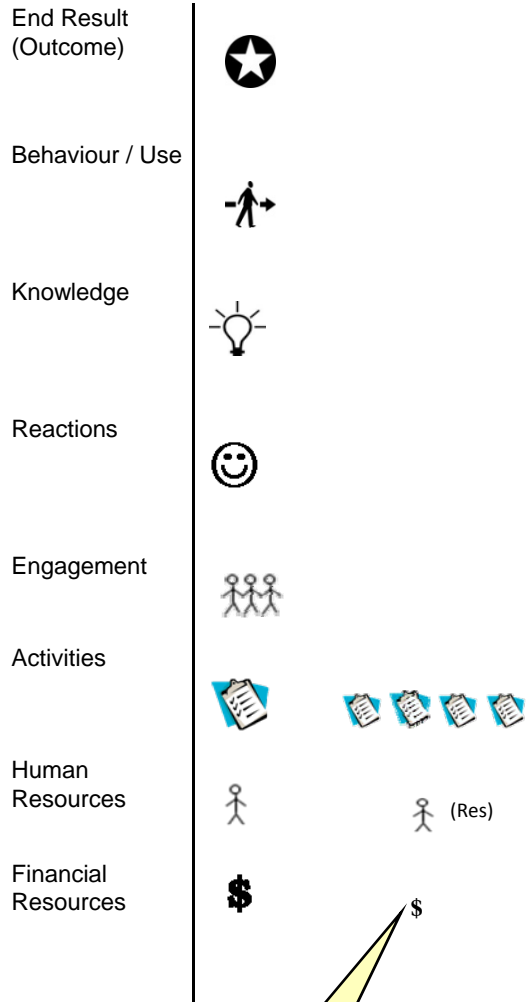
Initiative Chain of Results	Hierarchy of Evaluation Criteria / Evidence	Typical Indicators	Typical Sources / Methods
End outcomes 	Measures of impact on overall problem, ultimate goals, side effects, social and economic consequences	<ul style="list-style-type: none"> • Rate or incidence of cancer (incidence, mortality, morbidity) • Level of quality of life (Index TBD) • Level of advances in cancer science / research 	<ul style="list-style-type: none"> • Specialized analyses / evaluations • Statistical agency data • Canadian cancer statistics • Analytical and policy groups
Practice and behaviour change 	Measures of adoption of new practices and behaviour over time	<ul style="list-style-type: none"> • Level of research used (knowledge transfer, practice adoption) by scientists / policy makers / institutions / health care practitioners / consumers • Level of research used in curricula for new researchers (citation in text books and reading lists) • Level of research cited in ongoing health professional education material • Level of research cited in public policy documents • Level of research cited in advocacy publications 	<ul style="list-style-type: none"> • Physical observation • Inspections, reviews • Surveys • Evaluation studies • Content analysis
Knowledge, attitude, skill and aspiration change 	Measures of individual and group changes in knowledge, abilities, skills and aspirations	<ul style="list-style-type: none"> • Level of understanding of key related science information generated through research by scientists / policy makers / institutions / health care practitioners / consumers • Level of self-expressed commitment to specific areas of science / research or practice / protocol / policy change by scientists / policy makers / institutions / health care practitioners / consumers • Level of development of new knowledge in cancer research • Level of development of new methods in cancer research • Level of published research findings in a timely manner and in peer-reviewed journals with high “impact factors” 	<ul style="list-style-type: none"> • Independent review of target groups • Content analysis • Survey, group self-assessment • Testing / certification • Bibliometrics
Reactions 	What participants and clients say about the program; satisfaction; interest, strengths, and weaknesses	<ul style="list-style-type: none"> • Level of program recognition and support from key stakeholders / target groups / participants • Level (volume, accuracy and ‘tone’) of media coverage of research and program activities 	<ul style="list-style-type: none"> • Usage / participation tracking • Correspondence content analysis • Surveys • Media content analysis
Engagement / participation 	The characteristics of program participants and clients; number, nature of involvement, and background	<ul style="list-style-type: none"> • Level of engagement with other centres, networks, academic institutions, government agencies, etc. • Level of engagement by stakeholders / target groups / participants • Level of multidisciplinary and / or multisectorial research activities • Level of recruitment and retention of stakeholders / target groups / participants (e.g. junior investigators, researchers, review panellists, etc.) • Level of established external scientific advisory board(s) 	<ul style="list-style-type: none"> • Web use tracking • Correspondence content analysis • Observation of meetings / events • Meeting attendance records • Stakeholder relationship management / tracking (e.g. contracts and agreements) • Surveys
Activities & outputs 	Implementation data on what the program actually offers	<ul style="list-style-type: none"> • Level of research as per internal review guidelines • Extent to which plans, strategies, frameworks, etc. are delivered as per expectations (expected timelines, resource usage and quality levels) • Extent to which governance structure adheres to internal guidelines • Extent to which policy and financial decisions are made according to Board / Senior Management / Expert Advisory Committee(s) accepted guidelines and standards • Extent to which internal and external communication strategies adhere to internal standards and protocols / policies 	<ul style="list-style-type: none"> • Project / initiative tracking • Project reports • Content analysis or records • Peer-review • Operating reviews
Inputs 	Resources expended; number and types of staff involved; time expended	<ul style="list-style-type: none"> • Level of human resources (staffing) at all levels (according to norms, vacancies, expectations, benchmarks) • Level of financial resources (budgets vs. actuals) at all levels 	<ul style="list-style-type: none"> • Budget analysis • Time, reporting and budget / plan review • Activity-based costing

steve.montague@primm.net

The Vision

- Refer to a hierarchy to sort key indicators
- Focus on Mission as the key end outcome
- Tell a performance story – over time
- Emphasize behaviour change
- Employ ‘Realist Evaluation’ Questions... not simply ‘payback’ questions

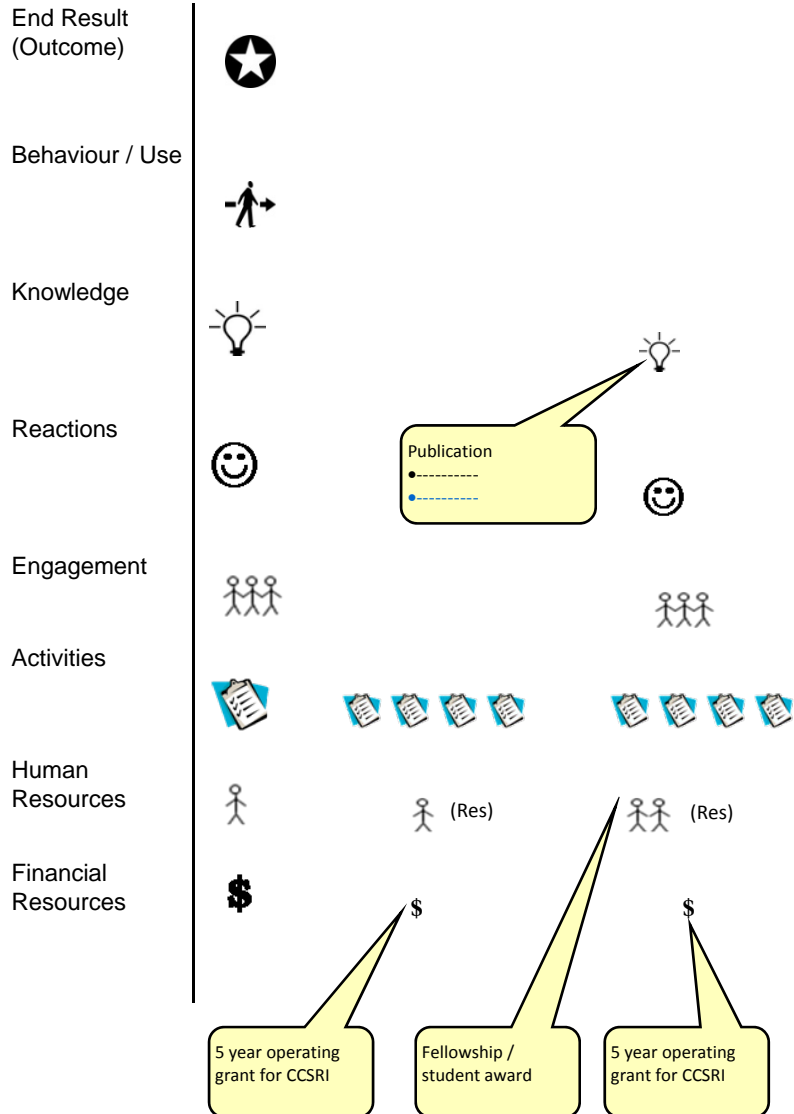
T1 (0-5)

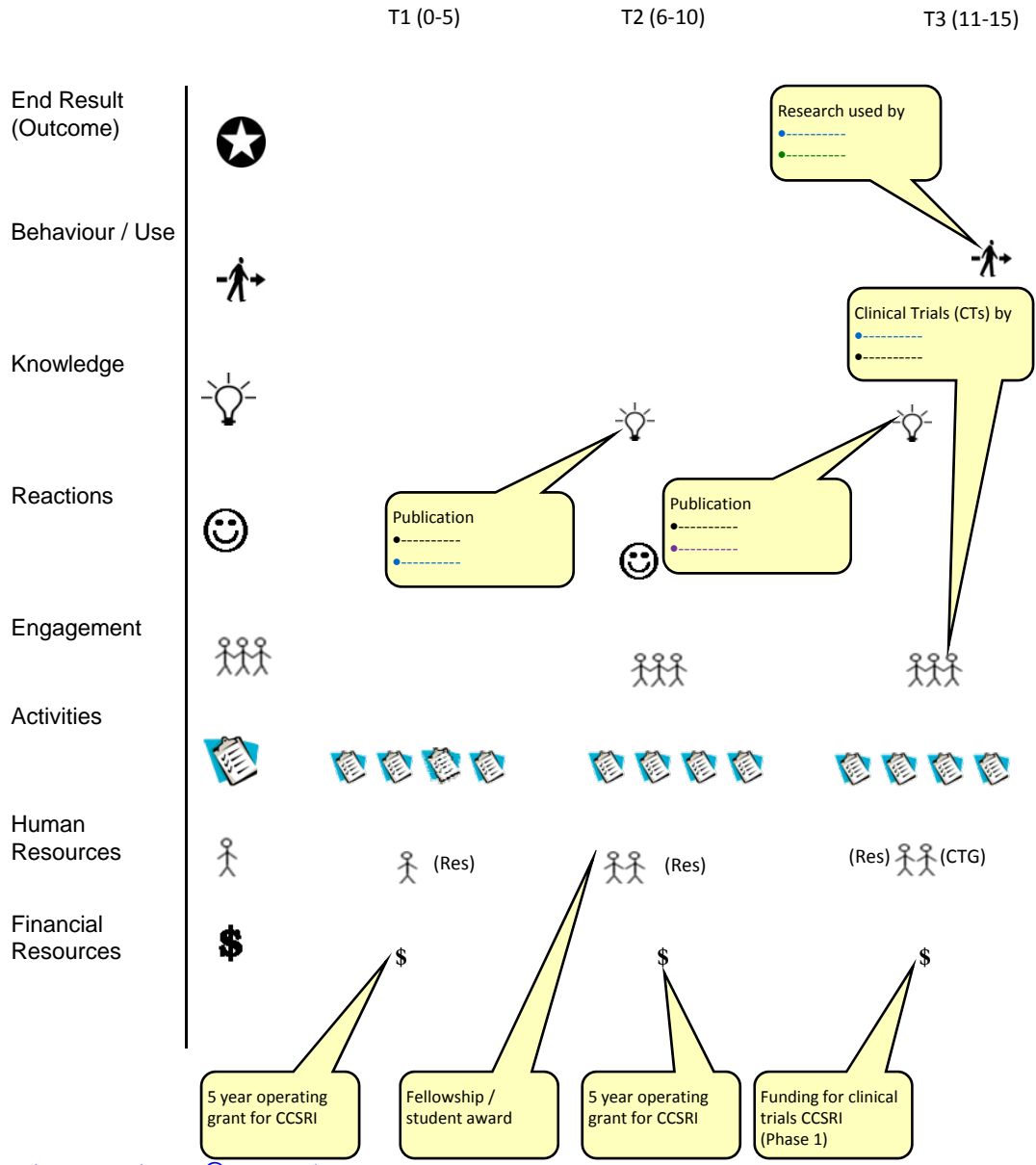


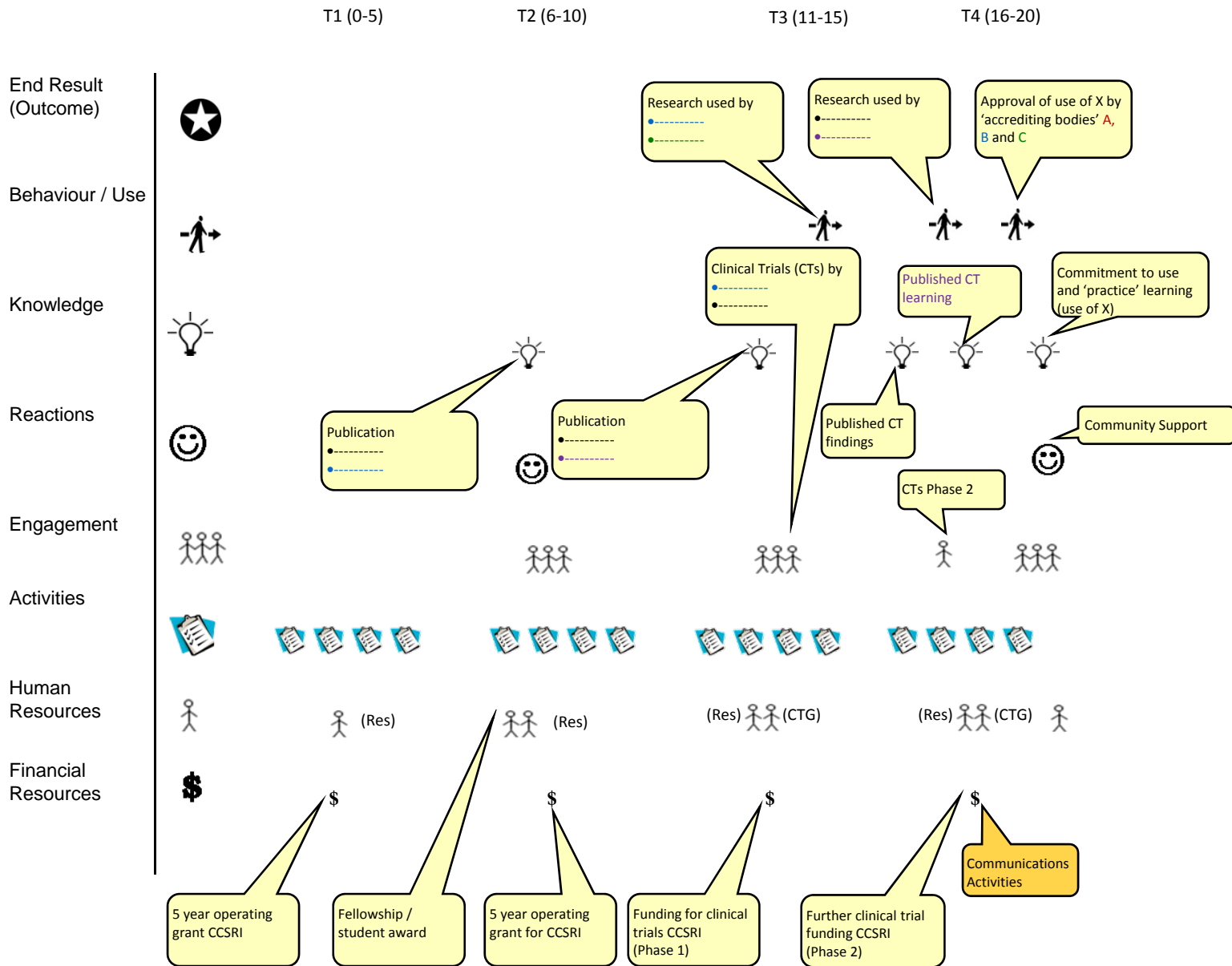
5 year operating grant for CCSRI

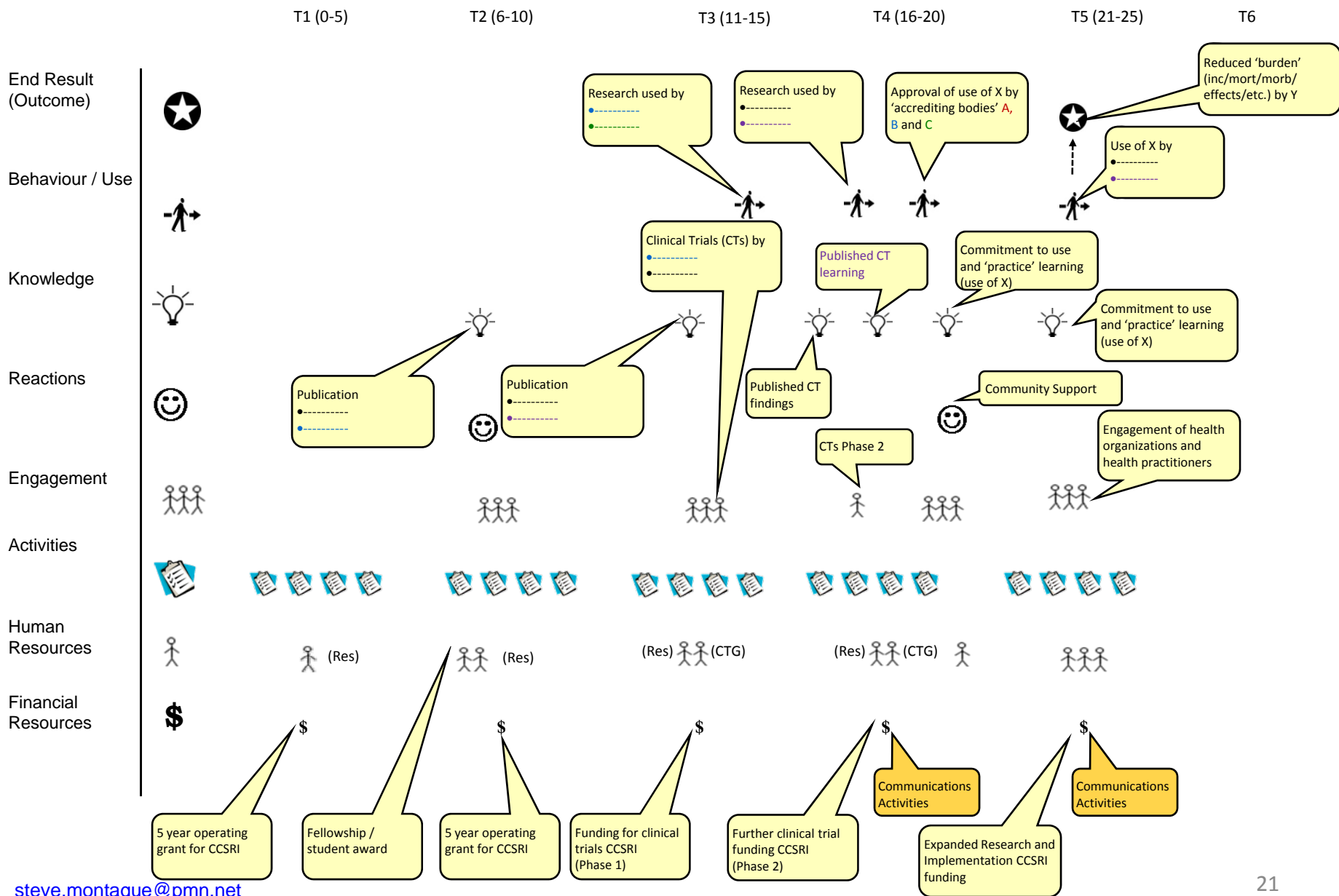
T1 (0-5)

T2 (6-10)

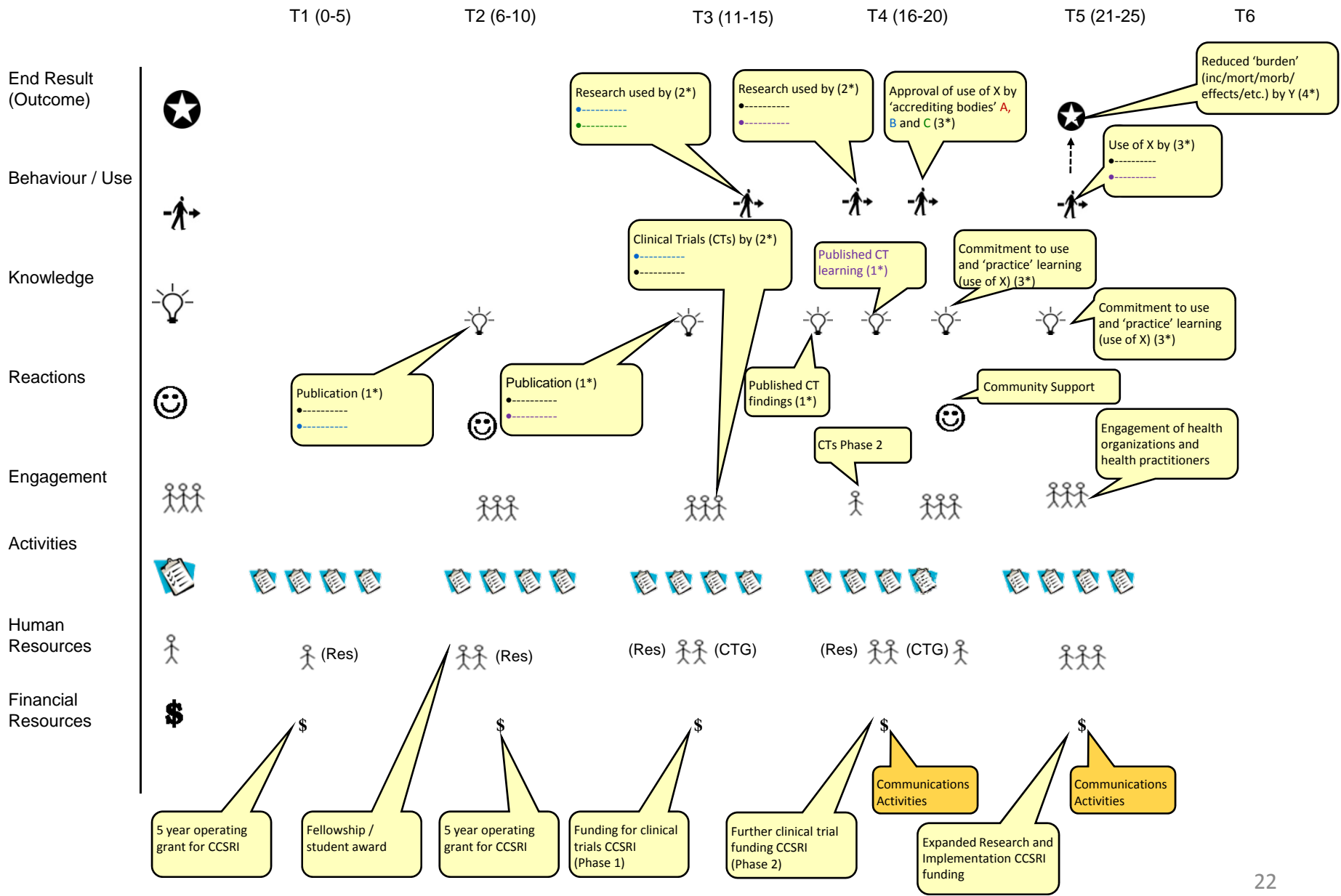








Mock-up Results Chain Applied Over Decades to a Research Initiative



*Linkage to RAND Payback categories respectively: 1. Knowledge Creation; 2. Research Targeting, Capacity Building; 3. Informing Policy or Product Development; 4. Health Benefits; 5. Broader Economic Benefits
 Source: S Montague – ACOR Presentation May 27, 2009

Realist Evaluation Questions

An explanatory compendium for complex RDT initiatives
(Adapted from Evidence-based Policy, Ray Pawson)

1. Program theories – how is the RDT initiative supposed to work?
2. Reasoning and reactions of stakeholders – are there differences in the understanding of the RDT initiative theory?
3. Integrity of the implementation chain – is the RDT initiative theory applied consistently and cumulatively?
4. Negotiation and feedback in implementation – does the RDT initiative theory tend to bend in actual usage?
5. Contextual influences – does the RDT initiative theory fare better with particular individuals, interpersonal relations, institutions and infrastructures?
6. History of the RDT initiative and relationships with other policies – does the policy apparatus surrounding the theory advance or impede it?
7. Multiple, unintended, long-term effects – is the theory self-affirming or self-defeating or self-neutralizing?

Challenges

- Research culture vs. management culture...
Investigator vs. mission driven
- The criticism that this might drive out explorative innovation...
- ‘Getting’ the point that the real evaluation story is all about human behaviours

Conclusions

- A hierarchy of results can help to sort RT&D indicators
- A storyline is key to provide indicator context – evaluation logic models offer this.. but they need to be adapted for RT+D.. Stay tuned..
- Realist evaluation questions may prove to be more promising as a support to management than simple ‘payback’ enquiries... at least if we are serious about learning – as well as ‘justification’
- Principle – based guidelines work better than prescriptions