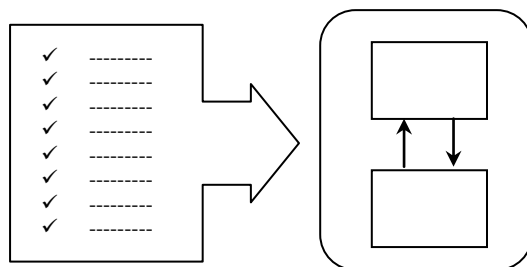


Checklists for Context, Policy and Program Delivery: Helping to Assess What Works Steve Montague, March 2013



For every complex problem, there is a solution that is simple, neat, and wrong.
- HL Mencken

As policy and program environments are increasingly recognized as complicated and complex, policy makers, program managers, planners and evaluators have increasingly recognized the important role of context. Answering the question of ‘What works’ can be daunting for both analysts and decision-makers. The Government of the UK laid out nine core competencies¹ of evidence-based policy and decision making in 1999 as follows:

- *Forward looking* – takes a long-term view, based on statistical trends and informed predictions, of the likely impact of policy.
- *Outward looking* – takes account of factors in the national, European and international situation and communicates policy effectively.
- *Innovative and creative* – questions established ways of dealing with things and encourages new ideas; open to comments and suggestions of others.
- *Using evidence* – uses best available evidence from a wide range of sources and involves key stakeholders at an early stage.
- *Inclusive* – takes account of the impact on the needs of all those directly or indirectly affected by the policy.
- *Joined-up* – looks beyond institutional boundaries to the government’s strategic objectives; establishes the ethical and legal base for policy.
- *Evaluates* – builds systematic evaluation of early outcomes into the policy process.
- *Reviews* – keeps established policy under review to ensure it continues to deal with the problems it was designed to tackle, taking account of associated effects elsewhere.
- *Learns lessons* – learns from experience of what works and what doesn’t.

The problem is that, while some reasonably elegant tools and approaches were contemplated, a decade and a half later the support of policy and public administration with credible, clear and timely evidence remains elusive. Promised investments in tools such as evaluation have not materialized and this certainly forms part of the problem. But the ‘under-funding’ of monitoring and measurement, social science research and corporate review functions like evaluation has been a chronic condition of government for years. (e.g. See Muller-Clemm 1997, Segsworth 2005, Shepherd 2012 for commentary on the evaluation function.) What if a fundamentally different approach is needed?

Early experiments in government and not for profit sectors suggest that what is needed is less traditional evaluation and analysis and more of what might be called a participative realistic approach². The following table summarizes this argument.

¹ Source: Davies, H. T.O., Nutley S. M., and Smith, P.C. (2004) What Works? Evidence-based policy and practice in public services

² See Pawson and Tilley 1997, Pawson 2006

WHAT WE NEED:

LESS:	MORE:
Aggregates, averages, summary, efficiency-effectiveness and simple minded "value for money"	Relevance and contextualized analysis and synthesis addressing how we value what works (to what extent) for whom in what conditions and why?
Linear, one-way, unexplained, context-absent box and wire diagrams	'Situating', described, systems oriented models describing theories of implementation (delivery design) as well as theories of change with key actors
'Standardized' approaches ranked by pre-determined hierarchies of 'rigor'...and approach 'worship'...in studies conducted by cloistered 'experts'.	Flexible, adapted and integrated measures and approaches fundamentally guided by issues and results logic (theories of change and implementation) and drawing on a diversity of sources and perspectives using networks and communities as active participants. Meta-accumulation and applied use of knowledge

Sources: 1. Montague, S. and Shepherd, R. (2011) *Addressing the Public Policy Evaluation Imbalance: A Realistic Approach* AEA Presentation

<https://higherlogicdownload.s3.amazonaws.com/EVAL/AEA%20Realist%20Presentation1.pdf?AWSAccessKeyId=AKIAJH5D4I4FWRALBOUA&Expires=1364230649&Signature=o42iJ2WvIQHJ7TmKTKHIX8unRdU%3D>

2. Montague, S. (2012) *Valuing Difference - The Key to Evaluation* CES 2012 Presentation http://evaluationcanada.ca/distribution/20120516_montague_steve.pdf

In order to:

- i) address key questions about what works;
- ii) apply systems oriented thinking about both implementation and broad policy and program types;
and
- iii) accumulate knowledge from a broad set of experiences,

we need to 'move-up' from a focus on 'programs' and to consider *policy instruments* (program 'types') as our key focus (evaluand). In order to do this we need to pay attention to things like Bemelmans-Vidéc, Rist & Vedung's (1998) definition of carrots, sticks, sermons etc. These tools have common characteristics and success factors (often found by consulting past research, studies and less formal experience) which can and should be referenced every time they are considered in plans, proposals, policy frameworks, evaluations and reviews. The change from a conventional program focus is a subtle but important distinction, since some programs are only part of a policy instrument and vice versa (e.g. A communications 'program' could relate to a much bigger policy instrument creating incentives. Alternatively, the 'policy instrument' of information / education could be considered a component of a bigger 'program' of incentives.) We *also* need to consider institutional arrangements and implementation considerations as part of the evaluand, or at least as part of the results logic.³

The data collection, analysis and synthesis to support a focus on policy / program type and delivery design need to move away from conventional views of evidence for policy making since policy environments are inherently complex and context dependant. Analysts, planners and managers need to accept multiple sources, the wisdom of crowds⁴ and other approaches which are structured and focussed on addressing the theories of implementation and change (e.g. realist approaches, contribution analysis). They also need to accumulate knowledge through case work.⁵

³ Chen (2005), calls the implementation component the 'action theory'.

⁴ See *The Wisdom of Crowds* Surowiecki 2005 and *The Wisdom of Crowds – It Applies to Performance Measurement Too!* at <http://www.pmn.net/wp-content/uploads/The-Wisdom-of-Crowds1.pdf>

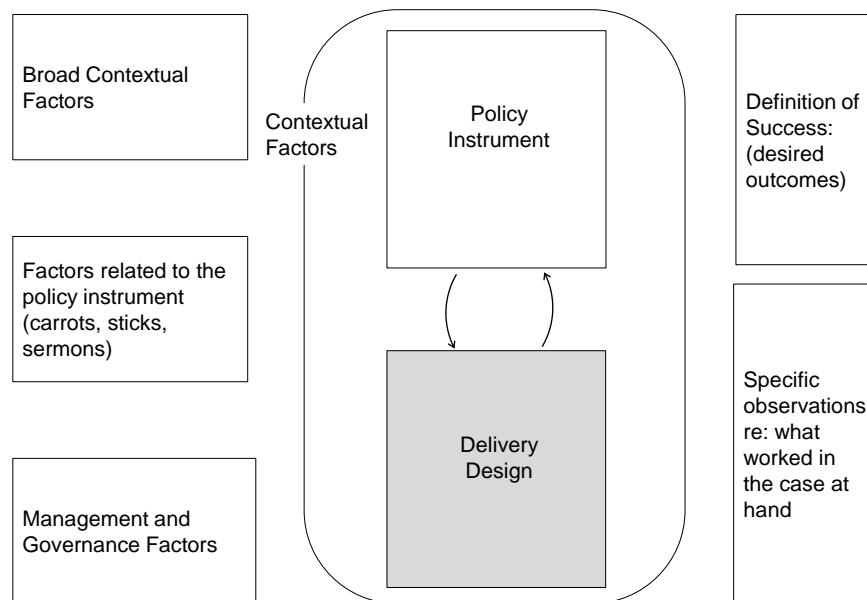
⁵ The idea is to emphasize generative approaches. See Stern et al 2012 – *Broadening The Range Of Designs And Methods For Impact Evaluations* Department of International Development (DFID) <http://www.dfid.gov.uk/R4D/Output/189575/Default.aspx>

In order to achieve this new approach, a summary method and tool is proposed as follows:

Step 1	Review the policy program / or initiative with a view to establishing its theory of need and theory of change (i.e. What is the problem? What is being done? and How will / does that work to achieve desired results?)
Step 2	Codify the theory of change into the nature of the program or policy instrument (e.g. incentive [carrots], deterrent [sticks], suasion [sermons] Bimmelmans-Vidéc , Rist and Vedung, (1998)) and the delivery design (action theory as per Chen 2005, Funnell and Rogers 2011). (e.g. Single Agency delivery, Multiple Agency, 3 P Partnership, Delegation to Sector Association etc.)
Step 3	Draw on research and collective wisdom regarding important factors influencing this type of policy instrument and this type of delivery design applied in similar or reasonably similar (or sometimes even approximately similar) circumstances.
Step 4	Construct a model (checklist) of factors related to the theory of change model, and use this to guide research, evaluation, performance measurement and even less formal discussions around the essential question of what works (to what extent) for whom in what conditions and why?
Step 5	Collect, analyse and synthesize data/information and report against, update and revise the model (checklist) based on observations and findings. (Continuous loop back to Step 1.)

A basic framework for such a tool is contained in **Figure 1** below:

Figure 1: Basic Framework for Assessing What Works in Policy, Programs + Delivery



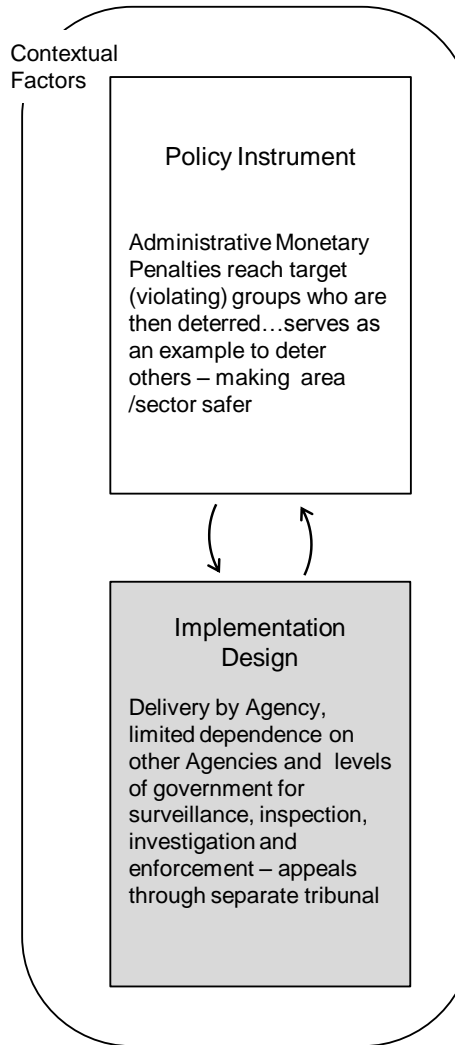
Note that the observations would be based on multiple lines of evidence related to the subject policy or program and would tie observations on the particular case (policy, program or initiative) in terms of relative success to specific contextual factors. This corresponds to the Context-Mechanism-Outcome (CMO) framework at the core of realistic evaluation. An example is provided below based on a recent evaluation of Administrative Monetary Penalties (AMPs). Note that all contents of this chart are derived from published sources including the actual evaluation itself. See **Figure 2**.

Figure 2: What Works for Administrative Monetary Penalties

What Works for Administrative Monetary Penalties? Summary of Factors and Findings for AMPs at CFIA (with reference to others)

Broad Context (Social, Economic, Political, Technological)*
<input type="checkbox"/> Political / jurisdictional / legal (authorities over policy area) <input type="checkbox"/> Economic Factors <input type="checkbox"/> Infrastructure / value chain 'structural' readiness <input type="checkbox"/> History of policy and programming (experience in use of instruments and mechanisms) <input type="checkbox"/> Compatibility with other policies and programs (e.g., regulatory environment between and among levels and agencies of government)
Regulatory Instruments Considerations
<input type="checkbox"/> A sound understanding of community norms based on empirical evidence rather than assumptions. <input type="checkbox"/> Integration of the various phases of the deterrence process: detection, prosecution, application of sanctions. <input type="checkbox"/> Effective processes for registering and licensing the target group, especially in situations where the penalty relates to withdrawal of licenses as in the program to combat drunk driving. <input type="checkbox"/> Effective screening processes to identify potential offenders. <input type="checkbox"/> Effective inspection processes that detect non-compliance to a sufficient degree to act as a deterrent. <input type="checkbox"/> Effective prosecution processes to ensure that offenders can be successfully prosecuted following detection. <input type="checkbox"/> Penalty processes that fit the offence. <input type="checkbox"/> Effective recruitment of the community as allies in implementation (constructive engagement). <input type="checkbox"/> Effective communication processes concerning the legitimacy of the regulation and that persuade the target audience that the regulation will be effectively implemented. <input type="checkbox"/> Social legitimacy of the instrument and sanction. <i>(Source: Funnell and Rogers 2011)</i>
<input type="checkbox"/> AMPs – Specific considerations. Suggestion they should be used for conditions of: <ul style="list-style-type: none"> – High volume – Inadvertent violations (non-systematic intent) – Clear violation evidence (easy to determine) <i>(Source: Praphu 2010, Priest 2009, Tait 2007)</i>
Target Community Engagement*
<input type="checkbox"/> Level of target community engagement and type of engagement re: initiatives <input type="checkbox"/> Capacity of target area actors <i>*(Source: derived and adapted from Sager and Andereggen 2012)</i>
Partner and Sector Engagement*
<input type="checkbox"/> Level of sector engagement in priority setting, planning and governance (Public, Private, NFP, other) <input type="checkbox"/> Capacity and level of shared agenda of Sector and Partner 'co-deliverers'
Public Management*
<input type="checkbox"/> Program (Policy) priorities (planning, coordination) <input type="checkbox"/> Governance and Management Accountability components** ** Accountability requisite components include clear roles and responsibilities, clear performance expectations, balanced expectations and capacity, credible reporting and reasonable review and adjustment (see Sager and Andereggen 2012 and OAG principles for Accountability 2009)

Contextual Factors



Definition of what 'works':

1. Timely and efficient transaction
2. AMP' d group or individual paid fine
3. No appeal
4. No 'overturns' on appeal
5. No evidence of recidivism (i.e. violator went back into compliance)

Where AMPs did not work:

- Hard to interpret regulatory clauses (e.g. 'undue suffering')
- Complicated interactions (many players) difficult to know who to penalize
- Low unit value cargo (e.g. cull animals) or high total value where penalty considered cost of doing business

Example by Steve Montague, March 18, 2013

steve.montague@pmn.net

When conducted using a generative orientation and with the participation of key stakeholders such an approach both produces an important evaluation of how an initiative worked in a given context and at the same time contributes to knowledge about what works for similar policy instrument types applied with similar delivery designs (What works for whom in what conditions and why?). Experience has shown that the approach also invites stakeholder dialogue. It can furthermore and perhaps most importantly in a world increasingly suffering from attention deficits, produce a checklist take-away for people interested in pursuing related strategies, policies and programs.

Going forward it appears that this approach can be useful at several levels (e.g. In regulatory policy it can apply to broad regulatory deterrence strategies, specific deterrence methods like AMPs or specific deterrence tactics like how to deliver a successful notice of variance or how to provide appropriate redress options.) and across many areas. By featuring policy instrument and delivery design 'types' it encourages and enables far reaching comparisons which in turn can reveal important common factors (e.g. Recent reviews of outcome-based codes and standards in construction and in food safety revealed surprising similarities.) This suggests that , if not quite 'laws', at least valuable touchstones or common considerations can be developed and accumulated over time for various program (policy instrument) types and for different delivery designs. These can in turn readily be drawn upon when planning new (or reviewing 'old') policy or program initiatives. (e.g. If we are going to consider using an administrative monetary penalty as a deterrence instrument we had better review our checklist of key factors found to influence success. **See Figure 2.**)

With this article and related upcoming projects it is expected that further trial and refinement of this thinking and analytical model can be applied, knowledge accumulated and then used as a direct real-time support to decision-makers. Furthermore it is hoped that such an approach can assist evaluators to scope their work and to communicate with users and other stakeholders. In this way the approach can potentially help agencies to achieve the evaluative, inclusive, outward looking and learning oriented 'core competencies' suggested for public policy and decision-making so many years ago.

On June 10th 2013, a brief workshop was held demonstrating the ideas and models highlighted here. The areas of model exploration were:

1. Administrative Monetary Penalties (Regulatory Instrument)
2. Innovation Assistance (Incentive G&C)
3. Health Services
4. Mentoring (Case Management)
5. Performance-based Codes

For more on the conduct of this workshop and / or further developments regarding context 'checklists' and 'checkpoints', contact info@pmn.net.