

Meta-evaluation, analytic logic models and the assessment of impacts of sport policies

Chen, Shushu; Henry, Ian; Ko, Lin-Mei

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Chen, S, Henry, I & Ko, L-M 2013, Meta-evaluation, analytic logic models and the assessment of impacts of sport policies. in I Henry & L-M Ko (eds), *Routledge Handbook of Sport Policy*. Routledge, London, pp. 33.

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Eligibility for repository: Checked on 17/1/2017. This is an Accepted Manuscript of a book chapter published by Routledge in *Routledge Handbook of Sport Policy* on 23/7/2013, available online: <https://www.routledge.com/Routledge-Handbook-of-Sport-Policy/Henry-Ko/p/book/9780415666619>.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Chapter 4: Meta-evaluation, Analytic Logic Models and the Assessment of Impacts of Sport Policies

Shushu Chen, Ian Henry and Ling-Mei Ko

In Henry, I and Ko, L-M (eds.) (2013) *Handbook of Sport Policy*, Routledge: London

4.1 Introduction

The period since the 1970s has seen a considerable growth in policy evaluation research in volume, diversity, and sophistication. The major dimensions of such change may be summarised under six categories.

- Evaluation has moved from simple input-output approaches which were primarily concerned with efficiency (cost per unit of output) often based on statistical association, to explanations which look ‘inside the policy-making black box’ to develop causal explanations of how particular policy interventions can bring about the changes desired (Leeuw and Vaessen, 2009).
- There has been an increasing recognition in social policy research that classical evaluation approaches associated with experimental logic are often inappropriate for evaluating change in social contexts in which the open system nature of those contexts militates against isolation of the impact of a particular ‘treatment’. Thus in such complex social settings, the context-specific nature of outcomes needs to be understood (Pawson, 2006; Pawson and Tilley, 1997).
- Related to this recognition of the limitations of experimental method has been the development of critical or social realist ontologies (Archer, 1995; Bhaskar, 1986; Bhaskar, 1978) which have promoted ‘depth realist’ explanations of the causal influence of real social structures. Such structures, though not amenable to direct observation, exert causal influences whose effects may be observed. Pawson and Tilley (1997) codify a realist approach in their promotion of ‘realistic’ evaluation, where they contrast this with the experimental, but also pragmatic, and constructivist models of evaluation, arguing for an approach which explains policy change by reference to the operation of ‘real’ social mechanisms and how these produce outcomes in specific contexts or types of context.

- These advances at the conceptual level as to what constitutes an adequate explanation of policy impact, have been accompanied by the promotion of greater precision in terms of specifying both intended and actual outcomes of policy intervention, in particular in the case of the former through the growing prominence of the use of analytic logic models (Cooksey et al., 2001), and in the latter, developments associated with the operationalising of impact through assessing additionality (Luukkonen, 2000).
- Finally advances in understanding and application of techniques of aggregation and synthesis of data and policy explanations have taken place in the field of meta-evaluation such that not only statistical aggregation of findings (meta-analysis) but also synthesis of qualitative claims explaining outcomes (meta-synthesis) have been developed in ways which allow the strength of explanation of collections of studies to be assessed.

The growth of both academic and policy-related interest in the analysis of sport policy has also grown alongside these developments. The aim of this chapter therefore is to illustrate and evaluate three of the key developments in policy evaluation and their application to the sports field by reference to existing studies. These are analytic logic models, meta-evaluation, and the estimation of additionality.

4.2 The Application of Logic Models in Sport Policy Analysis

The use of logic models came to prominence in the 1990s and 2000s with the growing emphasis on evidence-based policy practice (Head, 2009). It is used widely across government, not-for-profit organisations, and profit-based entrepreneurial activity (Dodd-Butera and Broderick, 2011; Jordan, 2010; Lenihan, 2011) providing a means to articulate and illustrate the intended relationship between *policy content* and *inputs, activities* (throughputs), *outputs* (the immediate products of the activities), *outcomes* (longer term effects), and *impacts* (the intended and unintended consequences of the policy initiative).

As defined by Conrad and Randolph (1999), a logic model is a “*graphic representation of a program that describes the program’s essential components and expected accomplishments and conveys the logical relationship between these components and their outcomes*” (p.18). It is a picture of how a programme works and to what end, with the provision in the case of analytic logic models, of a theory or theories of change.

Logic models are used for the purpose of: a) programme (or project) design and planning, when intended linkages between inputs and activities and longer term goals are open to multi-level stakeholders and evaluators. It helps to build up a shared understanding of the programme concepts and approach; b) programme implementation, the developing of an action plan against intended goals. Using the logic model to identify and collect the needed data in order to monitor and improve programming (if necessary); and c) programme evaluation and reporting where the model provides a set of measures against which elements may be reported. By developing a logic model, with an aim of learning and programme improvement, key results (positive and negative) of a particular programme can be presented.

The structure and key elements of logic model

The elements which go to make up the logic model vary slightly in terminology and content from one author to another, but perhaps classically incorporate the six elements identified in Table 4.1:

Table 4.1: Key Elements of a Logic Model

Key elements	What are they?
1. Context/Environment	These will include contextual factors such as the size and nature of the problem; the political, economic social and organisational context relevant to the programme, the objectives of the programme which have been set out by stakeholders.
2. Inputs/Resources	The inputs refer to the financial, human, and organisational resources provided to address the policy problem.
3. Activities (Throughputs)	The throughputs refer the kinds of actions which have been taken by policy implementers. These incorporate both what is done (in terms of activities undertaken), but also how these activities are undertaken since these may be critical to success. For example, activities aimed at increasing participation in sport for young girls from conservative and traditional communities may, for a variety of reasons, be effective only when these activities are delivered by female leaders as coaches.
4. Outputs	The outputs summarise what direct and immediate results of inputs and activities, such as numbers of participants attracted, regularity of their participation etc..
5. Outcomes	The outcomes refer to subsequent changes in behaviour triggered by delivering the programme, which can be divided into <i>short-term</i> and <i>long-term</i> outcomes as it may continue for many years after a project has been completed. For example, specific changes may include changes in participants' skills, behaviours, sense of self- efficacy and propensity to act in certain ways.
6. Impact	The broader <i>intended</i> and <i>unintended</i> changes which occur in organisations or communities as a consequence of the programme/project.

As we have suggested, a number of authors employ different modifications of this framework (see e.g. Stake, 1967; Stufflebeam, 1971). For example, Taylor-Powell (1999) conflates *inputs* and *activities/throughputs*. There are good reasons however for retaining these as separate aspects of a logic model, since not only what activities are undertaken, but also how the activities are provided may be critical to the achievement of intended outcomes. Thus if we consider a logic model underpinning an intervention to promote the engagement of young girls in sports activities, in certain types of community, it will be important to have the programme delivered by female leaders. In effect female leadership

may be a necessary, though not sufficient, condition of success since it provides role-models for the girls themselves, and reassure parents from conservative communities that concerns about aspects of required 'modesty' will not be infringed. Simply providing the 'activities' themselves in such social contexts may be unlikely to result in the desired changes to girls (and as they grow older) to young women's behaviours. How the activities are delivered and by whom may be a critical success factor.

Analytic Logic models can be more than a Programming Tool

Recently, a question has been raised in relation to the extent of the effectiveness of the logic model framework. In order to answer this question, it is important to make a distinction between, on the one hand, the descriptive outline of the stages of policy development and implementation, and on the other the specification of the causes of changes, which in effect represent implicit and/or explicit assumptions underpinning the model.

In respect of the role of logic models, one can identify two main different types, i.e. the *Descriptive Logic Model*, and the *Analytic Logic Model*. A *Descriptive Logic Model* focuses on simply presenting and describing the above key elements in chronological order. One of its functions is to map out a proposed programme that helps stakeholders and evaluators visualize and understand, at a very basic level, how financial and human investments represent a precursor to achieving intended programme goals. *Analytic Logic Models*, by contrast, highlight causal relationships between inputs, activities, outcomes, and impacts which may thus be subject to evaluation. In effect theories of change (or conceptual frameworks) are built into analytic models such that the reasons for the desired change being achieved, can be tested and evaluated in ways which can contribute to future policy and practice. This function provides rich explanation of the reasons for exactly which types of inputs or resources can contribute to the change desired. It also identifies the problems or issues that are addressed by the programme, and provides a rationale for selecting certain solution strategies and providing potential activities.

An example of a descriptive logic model for a project relating to HIV/AIDS Education through sport is given below (see Figure 4.1), and its more developed equivalent, an analytic logic model is provided to illustrate the difference between the two types (see Figure 4.2). This is

not taken from a specific case but is simply used here to illustrate the mechanics (and some of the limitations) of developing and applying both types of model. Figure 4.1 merely outlines the flow of events anticipated – from provision of AIDS education through sport, to enhanced knowledge, to intentions to modify behaviour, to reduced infection rates.

Figure 4.1: An Example of a Descriptive Logic Model

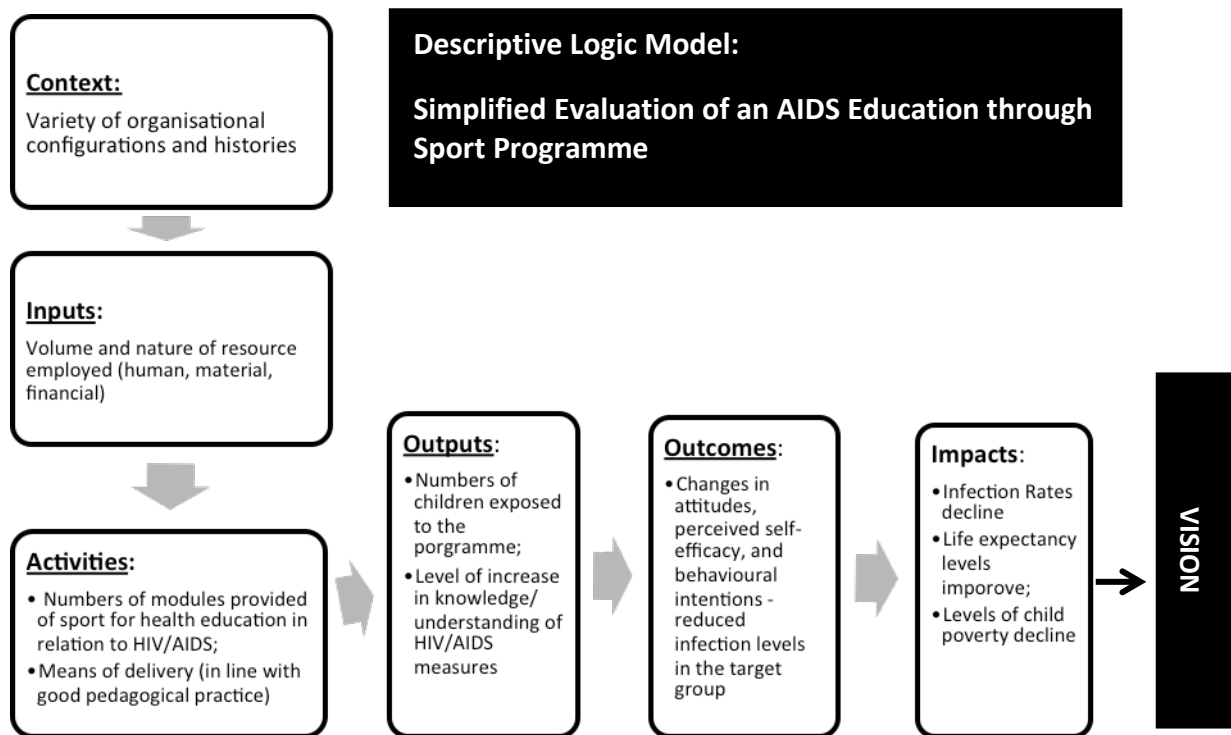


Figure 4.2 however makes explicit the assumptions on which the expectations of change are premised. The principal assumptions/theoretical premises might be defined as follows:

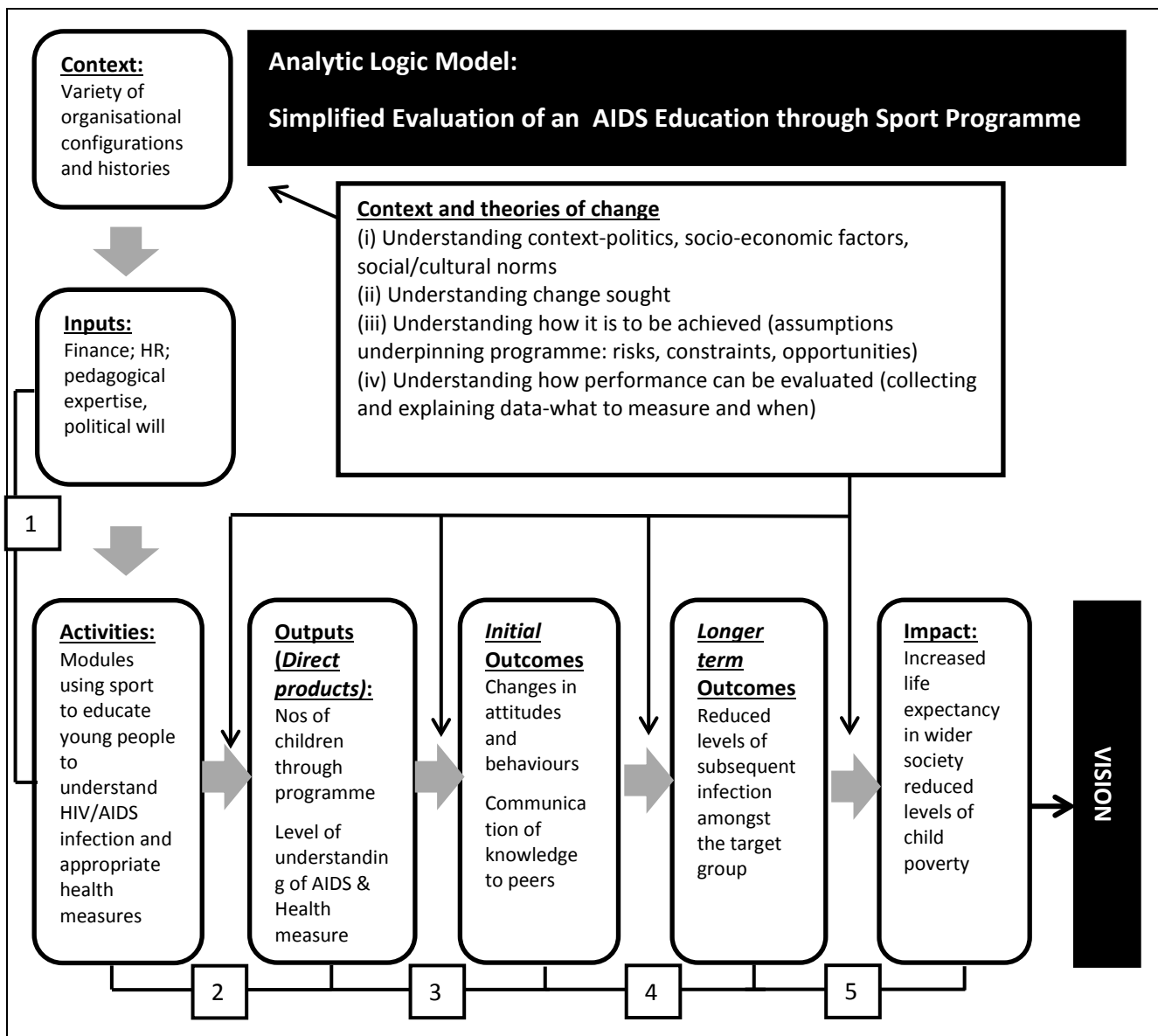
1. Young people are likely to engage with education through sports and games more enthusiastically than through traditional educational methods. Of course we know that this will not be true for all children (some children are alienated by sport), or for all sports (competitive games may be alienating for some children where they have little chance of 'success').
2. Developing self-efficacy through sport can enhance self-efficacy in wider social contexts. Perceived self-efficacy relates to peoples' beliefs about their capacities to produce effects, in short their perception that by their own efforts they may be able to affect their own lives. Teaching sporting skills can produce high levels of perceived self-efficacy (if appropriate positive experiences of sport are provided) and this is the

claimed mechanism here. However, such causal claims need to be tested or challenged in terms of internal logic or empirical evidence. In this case for example, negative experiences may produce the opposite where in competitive sport the individual perceives that no matter how hard they practice they will not be able to affect the sporting outcome. In addition it is by no means clear in the literature that enhancing self-efficacy in sport will result in enhanced self-efficacy in other domains (Biddle et al., 2007).

3. It is further assumed that improved knowledge and understanding on the part of the young people targeted will increase the likelihood of (intended) behavioural change. While one might argue that improved knowledge and understanding of sexual health matters will be a necessary condition of behavioural change, it is not necessarily a sufficient condition for such change for all individuals. The availability of resources through which to implement behavioural change (for example, the availability of condoms, or of clean needles for drug users) may well also be required.
4. Behavioural change will reduce levels of infection and thus increase life expectancy. Here the assumption is that the key behaviours in transforming HIV/AIDS infection have been identified, and knowledge about such behaviours transmitted during the education through sport programmes.
5. Higher parental life expectancy will reduce child poverty. Here there is an assumption that a major cause of child poverty is the lack of an adult/parent to provide care/shelter and some financial resource for children in societies with a high adult mortality rate through AIDS, though this may be only one factor in the production of child poverty, even in societies in which HIV/AIDS is prevalent.

Thus we can see that the underlying logic and the supporting theoretical assumptions are made explicit and can be subject to challenge in the articulation of the logic model in complex projects/programmes such as that described above. The assumptions are laid bare in the process of articulating the model, and the potential for developing measures against which reporting and evaluation might take place is clearly evident.

Figure 4.2 An Example of an Analytic Logic Model



Strengths and limitations of logic model approaches

The key advantage of the logic model approach is that it distils detailed descriptions of the assumptions underlying a programme to a ‘one-page’ format that can be easily read and followed (Cooksy et al., 2001). It explicitly depicts conceptualisation of each step of the whole chain, and helps to demonstrate a logical flow from a belief structure to related interventions, to outcomes, and then to impacts. In addition, the process of developing conceptualisation of the theory of the intervention is valuable as it helps to leverage greater insight into a programme operations and effectiveness.

Supporters of logic models believe that the process of collaboratively developing logic models together helps the often multi-level stakeholders and programme evaluators (at the beginning of the evaluation process) to reach a common understanding of the programme in order to define more clearly its vision and objectives, and to envisage whether the designed actions would accomplish the goals. It is useful in the sense that it is a mutual educational process which may help to avoid misuse of the theory by the programme operators (Patton, 1997).

Furthermore, having all the components clearly articulated in graphic form helps evaluators to link theory to practice and to accommodate changes in knowledge (Alter and Murty, 1997). The planned operation theory in the logic models may change in practice due, for instance, to the growing availability of resources, or the emergence of knowledge of unintended effects. It is therefore important for the evaluators and key stakeholders to constantly re-visit and scrutinise the defined assumptions in line with on-going activities and achieved outcomes along the delivery process, making changes to the model if necessary.

Although the logic model approaches provide benefits, as identified above, concerns in relation to its application have also been raised. For example, as Yin (1998) suggests, efforts need to be made in order to make the logic models more clearly articulated analytic strategies. He emphasises that most current logic models do not specify clearly enough the substantive processes between inputs, throughputs, outputs and outcomes. Clearly, without this articulation, logic models are simply a sequential pattern of events which do not aid understanding of how outcomes are actually produced.

When it comes to practice, some operational issues have also been identified. For example, debates around the costing of developing logic models have emerged, with some commentators suggesting that the process can take up a lot of resources (e.g. Bickman, 1989), while others have argued that the cost can be justified by benefits to programme stakeholders above and beyond their use to the evaluators (Patton, 1978), and that it is actually a way of avoiding costly evaluation in situations where evaluation efforts will not be made if there are unlikely to be observed effects as identified by the logic models (Wholey, 1994). Another concern, as Weiss (1997) has pointed out, is that logic models may depict rigid statements which limit the programme's responsiveness to new information. In

particular, programme evaluators may only concentrate on those listed outcomes and thus ignore unintended effects that are not part of the programme theory.

Furthermore, it is important to recognise that different stakeholders may have different logics. Given that a social intervention often involves different levels of stakeholders (national, regional, and sub-regional) each with their own interests and targets, there is a need to understand the way that different stakeholders view the world (and the policy problem, its intervention and so on). Adopting a single logic model may thus deny the possibility of other views. Nevertheless debates around establishing a logic model for a particular programme may well serve to bring to the surface underlying differences in assumptions, and even desired outcomes between different stakeholders.

Notwithstanding these limitations, logic models have significant potential as integrative frameworks that not only combine pattern matching and time-series analysis techniques (Yin, 1998), but also provide a unique tool in explicating underlying causal relationships in a simple picture, hence its usefulness for case study evaluation (e.g. Mulroy and Lauber, 2004; Yin, 2009).

4.3 Meta-evaluation and Policy Analysis

Meta-evaluation is an increasingly recognised approach for evaluating a number of evaluation studies. It initially emerged as a result of evaluators being required to appraise their own evaluations (Stufflebeam, 1974), particularly in the education area (e.g. in the Advanced Technological Education (ATE) Evaluation project described by Gullickson et al., 2006). Nevertheless, the number of meta-evaluation studies appearing in the literature is not substantial despite a growing recognition of the need for evaluation of evaluations. This was highlighted by Nilsson and Hogben (1983) in the early 1980s, and has been further emphasised in the past decades by a number of commentators e.g. Cooksy (1999), Scott-Little, et al (2002), Bustelo (2002b), Madzivhandila, et al (2010). In addition, some studies entitled 'meta-evaluations' are rather more in the nature of cross-case analysis rather than meta-evaluations (e.g. Ashworth et al., 2004; Russ-Eft and Preskill, 2008).

The concept of 'meta-evaluation' was first introduced by Michael Scriven writing in the late 1960s (Scriven, 1969) and was subsequently developed in his *Evaluation Thesaurus* (Scriven,

1991), in which he lays emphasis on “The evaluation of evaluations – [and] indirectly, the evaluation of evaluators” (Scriven, 1991, p. 228).

Subsequent definitions of meta-evaluation include for instance that of Patton (1997, p.193), for whom meta-evaluation is the answer to the following questions: “*Was the evaluation well done? Is it worth using? Did the evaluation meet the profession’s standards and principles?*” Bustelo (2002a) suggests that meta-evaluation is a systematic gathering, analysis and assessment of a pre-determined set of evaluation processes, while Stufflebeam based on his experience in leading the development of professional standards for evaluations in the US, (2001, p.185) defines meta-evaluation as “*the process of delineating, obtaining, and applying descriptive information and judgmental information – about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity / honesty, respectfulness, and social responsibility – to guide the evaluation and / or report its strengths and weaknesses.*”

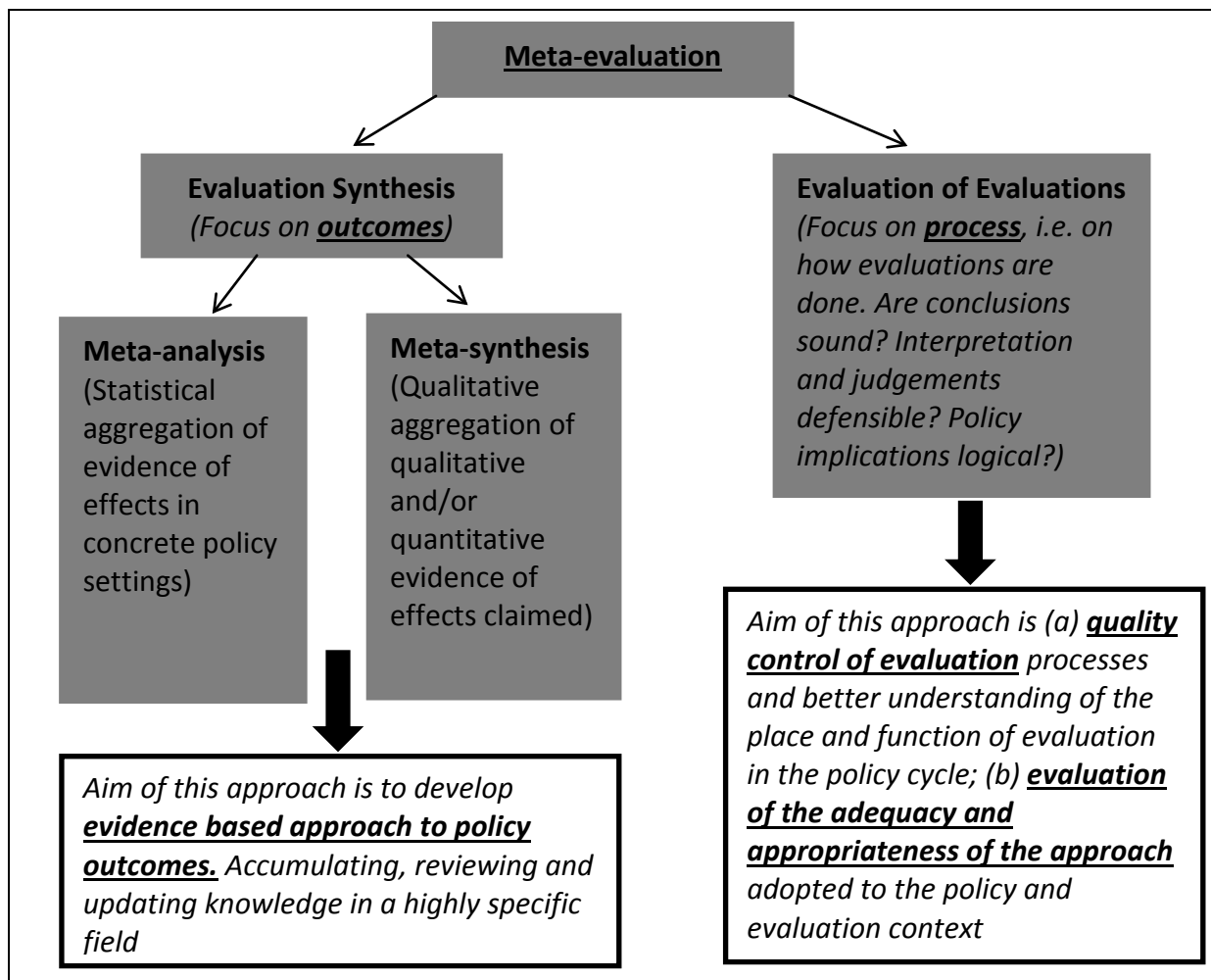
The application of meta-evaluations

The demand for evidence based policy / practice followed on from the development of evidence based medicine approach to clinical practice in the early 1990s in which the demand for clinical decisions to be based on the best available scientific evidence rather than intuition. This philosophy was quickly adopted and adapted to other fields reinforcing an emphasis on quantitative rather than qualitative analysis and positivistic methods. The evidence-based practice movement also reinforced the need for scrutiny of the quality of evidence and thus the demand for evaluators to ensure the quality of their methods and analysis. The prominence of meta-evaluation was a product of this movement since it could provide recommendations in relation to how evaluation studies should be designed to produce technically adequate, useful and cost effective results. In addition the conducting of meta-evaluation serves to enhance the accountability of evaluators themselves, controlling potential evaluator bias, and increasing evaluation credibility.

It is important to distinguish between different types (forms and functions) of meta-evaluation. Figure 4.3 seeks to clarify the distinction between the related terms of meta-evaluation, meta-analysis, meta-synthesis and quality assurance in the form of evaluation of evaluations. The diagram incorporates a basic distinction between on the one hand,

evaluation synthesis (the synthesis of quantitative, or qualitative or mixed evidence of outcomes of particular policies / interventions), and on the other assessment of the process of evaluation adopted (the ontological and epistemological foundations of approaches adopted, internal logic, methods employed, treatment of data and robustness of conclusions). A comprehensive and rigorous meta-evaluation will be required to address both evaluation of process and synthesis of outcomes.

Figure 4.3: A comprehensive meta-evaluation graph



An example of a meta-evaluation of sport policy is provided by the London 2012 Meta-evaluation commissioned by the Department of Culture, Media and Sport, the first three reports of which outline the scope and methodology adopted (Grant Thornton et al., 2011b, c, 2012). The study, which is reputed to be the largest study to date of legacies of a mega-

event, seeks to evaluate and synthesise the results of the evaluations of individual projects or programmes which are a product of London's hosting of the 2012 Games.

The approach adopted in this study illustrates the potential but also a number of the practical difficulties in applying the meta-evaluation approach. The study as a whole is premised on a logic model designed to capture the intended legacy outcomes from the Games, and employs methods consistent with governmental advice on policy evaluation (Government Social Research Unit, 2007; HM Treasury, 2003) and requirements from the DCMS in terms specifically of the assessment of legacy (Department of Media Culture and Sport, 2008).

Reports 1 and 2 in the study (Grant Thornton et al., 2011b, c) outline the scope, research questions, strategy and methods adopted. The study aims to synthesise into a single overarching study "the findings of individual 'project-level' evaluations – commissioned outside of the meta-evaluation study – in order to provide a comprehensive initial evaluation of the additionality, outputs, results, impacts and associated benefits of the investment in the Games." (Grant Thornton et al., 2011a: p. 5). The four principal legacy themes or impacts identified for evaluation are: sport; the economy; community engagement; and the regeneration of East London. In each of these areas

Evaluations will ... be synthesised using a common set of output, result and outcome indicators, in order to answer a core set of research questions, paint a picture of the activity underway across each legacy theme and aggregate the impacts wherever possible. This 'bottom up' research approach will be supplemented with a combination of:

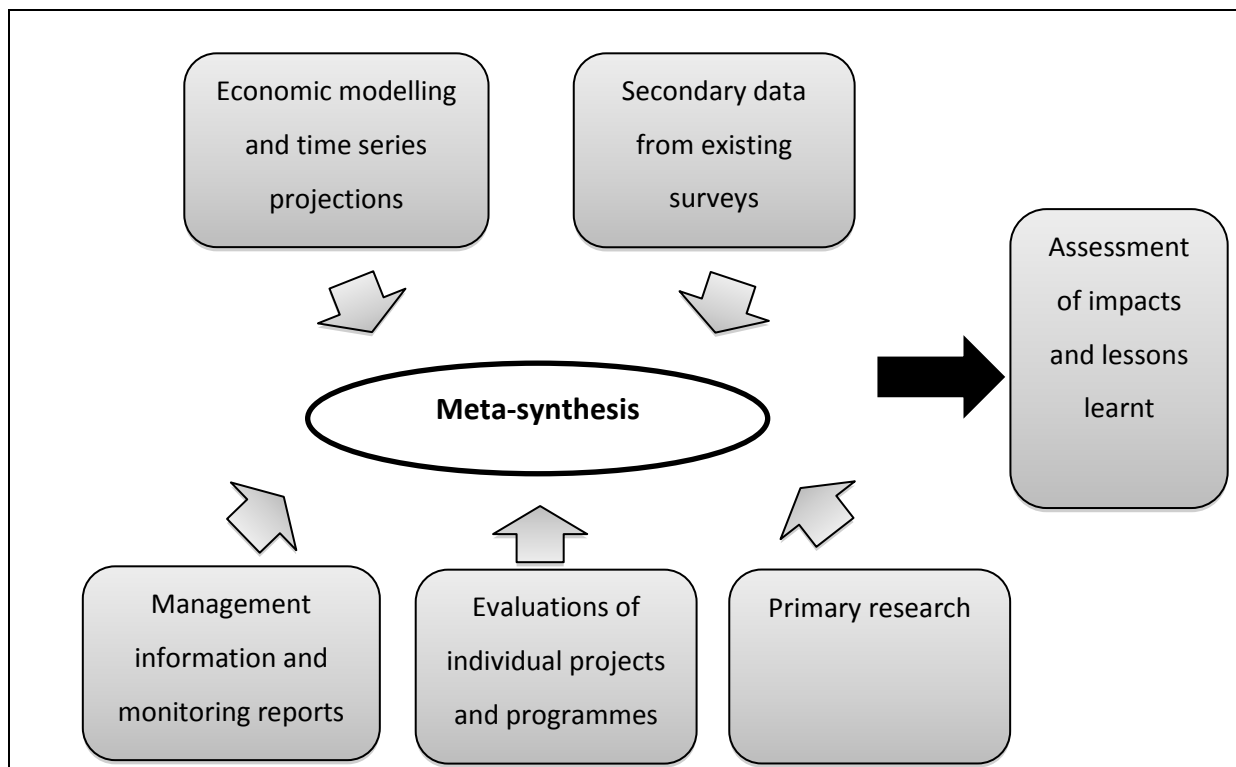
- Analysis of management information data, monitoring reports and case studies, particularly for major projects lacking evaluation;
- 'Top down' analysis of secondary data from National Statistics and established surveys, in some cases involving the inclusion of additional questions to aid the interpretation of the drivers of 'high level' trends;
- Economic modelling to assess wider and longer term economic impacts, including effects on nations and regions outside London;

- Limited primary research of different types, including both survey work and a programme of consultations.

(Grant Thornton et al., 2011a: p. 6)

Amongst the points to emphasise here is that meta-analysis which broadly attempts to aggregate the effect size of the impact of particular interventions, while useful in the aggregation of data on effect sizes in randomised control studies prevalent in the medical field, is inappropriate in this kind of context. The synthesis is thus a matter of integrating material from 'top down' data with 'bottom up', project-level evaluation data. To take a concrete example an intended legacy of the Games was that it would inspire greater levels of participation in sport and physical activity. National level cross-sectional data on participation across time from the Taking Part Survey developed by DCMS and the Active People Survey developed by Sport England provide a picture of participation using different operational measures of participation in sport. This can be compared with data on the impact of individual projects or programmes which have developed or perhaps intensified as a result of the winning of the bid to host the 2012 Games. Bottom-up evaluations are much more likely to reveal what types of intervention are effective in leveraging additional participation in sport and why they are successful, since they can incorporate qualitative analysis of the project and its impact. However the rigour and relevance of both top-down and bottom-up evaluations in terms of, for example, transparency, credibility, rigour, and accuracy need to be appraised before inclusion in the overall evaluation.

Figure 4.4: Meta-Synthesis Strategy



Source: adapted from Grant Thornton et al (2011a: p. 6)

Figure 4.4 above demonstrates the nature of the strategy adopted, which will involve the detailed appraisal of the nature and quality of the monitoring and management information systems, and of the evaluation of individual projects and programmes. It will also incorporate attempts to estimate in quantitative terms the impact of the Games on participation levels, not by meta-analysis but rather by an intelligent synthesis of the data.

At the time of writing (Report 4 is about to be published) a major concern is with reconciling the fact that top-down, national level studies indicate a fairly flat profile for sports participation (with little or no indication of increased participation from the period before the London Bid was initiated, to 2011, though this does not include the immediate post-Games period) and the evidence of bottom up data from individual (and sometimes quite large) interventions which suggest that these projects have provoked significant increases in participation. Several explanations might be mooted to reconcile such data. First it is unclear whether participation in new projects represents *substitution* for other forms of participation which thus does not result in increases in overall levels of participation. Second,

there may be displacement if the new 'Olympic initiatives' replaced other initiatives which would otherwise have been provided. Third, in many instances such projects may impact on those who already participate in sport, intensifying their rate of participation but not increasing the numbers of participants significantly. Where the project is intended to bring in new participants this attraction of existing participants may be described as a form of *leakage*. Third, the national data from the *Active People* and *Taking Part* surveys only capture adult participation (for those aged 16 or more) while the growth in participation may have been greatest among younger people. For this reason UK government has decided to lower the age range for which data is sought (and incidentally also to merge the two national surveys into a single data source).

4.3 'Additionality' and related concepts in the assessment of impact

Of course in order to assess the legacy of the 2012 Games for London and the UK, an assessment is required of the extent to which the staging of the Olympic Games in London produced inputs, throughputs, outputs and in particular outcomes that would not otherwise have occurred. In other words an assessment of the net impact of the Games requires an assessment of additionality.

The concept of additionality originally came from the evaluation of innovation and technology policy in which a justification for public support for technology development in private companies was needed to demonstrate that public funds did not simply displace private corporate investment in R&D, but were additional to that which would have happened anyway (Buisseret et al., 1995). The framework of additionality was developed and refined in the UK in the early 1980s (Luukkonen, 2000). In practical terms, additionality has become one of the key concepts employed in public sector policy evaluation studies together with, for instance, general impacts, effectiveness, efficiency, and value for money (English Partnership, 2008; HM Treasury, 2003).

In policy terms the focus of a concern with additionality is on distinguishing the *net* impact of a project or programme, that is, what additional impacts (and outputs/outcomes) were achieved exclusively as a result of the programme/project rather than those impacts/outcomes which would have occurred anyway. This is a critical issue for a number of reasons. First of all, in terms of the validity of findings, the final assessment of the impacts

created by a particular intervention does not stand up to scrutiny if only the *gross* impacts were captured. Without assessing additionality it is not clear what the intervention is adding over and above what would have happened anyway. As a result, it might present a misleading picture of the value of a programme due to the fact that only the direct impacts of a programme are measured, and the wider impacts, or how the project may have impacted on other activities, are not taken into account. In addition, the process of teasing out the additionality of an intervention helps programme/initiative developers and policy makers to gain a better understanding of all stages of an intervention's lifecycle, to make a comparison between actual achievements and the objectives of a programme, and to thereby identify unintentional outputs/outcomes.

Overall, the process of defining, or calculating / estimating additionality is crucial to maximise the impacts of an intervention, and to ensure it delivers real results. At the end of the process, it draws lessons from the evaluated programme to inform the work of stakeholders, and also the development and evaluation of future similar projects.

Measuring additionality

English Partnership (2008) provides practical guidance on the basic information of how to take into account the additionality of intervention with the purpose of ensuring the net impact could be assessed. The formula from Figure 4.5 displays how to assess the net impact, and also presents how to calculate the additionality of the intervention. Here, the application of the additionality formula does not focus on precise calculation of every element, but does require clarity about the likely scale and nature of an intervention's additional impacts.

Estimation of the net impact involves adjustments to be made for *leakage*, *displacement*, *substitution* and *multiplier effects*. The first step of calculating additionality is to set out the counterfactual scenario which means what would be the case if its antecedent were not true, in other words, what would have happened if the intervention had not gone ahead. For example, when calculating the impacts boosted by the UK hosting the London 2012 Games, the counterfactual scenario is defined as what would have happened without the London 2012 Games taken place. The counterfactual has two dimensions: (a) the policy counterfactual; (b) the outcome counterfactual.

- (a) **The policy counterfactual:** refers to an assessment of the key strategies, policies, and initiatives which would have been delivered in the absence of the Games. This would normally be done by reviewing the policy and strategic documents and conducting key stakeholder interviews to establish the nature, and direction of travel of policy, before a given baseline date. In the case of the 2012 meta-evaluation study for example a baseline date of 2003 was adopted because it was in that year that the British Olympic Association, the City of London, and the British government committed to submitting a bid to host the Games. There would clearly have been a range of interventions which would have been major drivers of changing people's behaviour in terms of sport participation even if the Games bid had been unsuccessful, and many of these policies would have been 'visible' before the baseline date. Even where policies were already in existence before the baseline, as a result of the London 2012 Games, the government may have paid more attention to the addressing of such issues resulting in larger injections of funding and policy effort into some existing initiatives, and also the implementation of some new interventions. Thus by contrasting the two policy scenarios the additional impact of policy change can more readily be captured.
- (b) **The outcome counterfactual:** relates to what results (outcomes) would have occurred on the ground in the absence of the Games. Thus for example the national and local levels of sport participation may be compared before and after the decision was taken to bid for hosting of the Games using data from national surveys. Only after the counterfactual analysis being identified can an estimation of the additional impact of legacy-related activities be made.

The assessment of net impact thus expressed in the formula:

Table 4.2: Net Impact Calculation

Net Impact	=	Additionality of Intervention
Net Impact*	=	[Gross Impact x (1-Leakage) x (1-Displacement) x (1-Substitution) x (1+Multiplier effect)]

*after taking into account of the counterfactuals effects

In order to assess the net impact one has to take into account four principal elements:

- **Leakage**: the extent to which the gross impact of benefits generated and intended for a particular group, region or country, incorporates beneficiaries from other groups, regions and countries. An example here might be where the UK's coach development system training coaches at elite level, produces coaches who subsequently take up employment with teams from other countries.
- **Displacement**: where new provision displaced other activities or services which had previously been supplied, thus where a provider of services launches an 'Olympic Fitness' facility but this has an impact on the market displacing other providers.
- **Substitution**: when consumers of a service or beneficiaries of a project simply substitute the new service provision for what they had been previously using or benefiting from.
- **Multiplier effect**: the extent to which direct benefits from an intervention trigger further additional indirect benefits. For example attracting of a family member to a jogging club results in other family members participating informally in jogging.

We outline here an example of an approach to assessing additionality and demonstrating associated issues in relation to the Workplace Challenge Programme (WCP). The WCP evaluation formed part of another meta-evaluation research project of the impact of the London 2012 Games in a non-hosting English region (Leicestershire) (Chen and Henry, 2012) which incorporated a meta-synthesis of qualitative and quantitative evidence within a county of the impact of the 2012 Games.

In terms of assessing additionality, the first step was to set out the counterfactual scenario. It was indicated in interviews with key stakeholders that the existence of the WCP was not directly attributable to the London 2012 Games (in other words, it would have taken place anyway). Nevertheless, it was given a greater prominence because of the 2012 Games. In terms of the London 2012 impact, a survey of participants (n= 202 of whom 77 were reporting participation after the Games)¹ indicated that

¹ There were two waves of survey collected respectively in 2011 and 2012. The data reported here was only drawn from the 2012 survey as it collected respondents' participation rate before the Games and after, which we believe was more appropriate to the context and well reflected the London 2012 impact (if there is any).

- 18% (+/- 8%) of the respondents who either agreed or strongly agreed that '*the publicity material for the London 2012 Games had influenced their decision to participate in the WCP programme*', thus the Games appeared to impact upon participation rates for the WCP;
- 32% (+/-10%) participants either agreed or strongly agreed that '*the publicity surrounding the 2012 Games made them more aware of the benefits of taking part in sport and physical activity*';
- Quantitative evidence collected before and after the programme suggested there was a statistically significant increase in the level of physical activity participation [from Time 1 (i.e. before taking part in the programme) (M= 4.62, SD= 1.71) to Time 2 (post programme report) (M= 5.95, SD= 2.18), $t(60) = -5.81, p < .000$ (two-tailed)]². The mean increase in the level of physical activity participation was 1.33 units (30 minutes of moderate intensity exercise). The eta squared statistic (Eta squared = .36) indicated a large effect size. However this was an increase largely on the part of people who already participated in sport.

In terms of exploring the additional *outcomes* of the WCP there are a number of issues to consider. First of all, the issue of *displacement* was not assessed in this case study though interviews with stakeholders did review whether the WCP did displace or replace existing schemes or programmes. Secondly, there was leakage in the project in the sense that the major impact was on a group outside the target group since it was intended to get the less active to become more active but in fact a disproportionate element of the response was by the already active. Substitution seems to have played little role in decision-making by the target group. This is indicated by the fact that regardless of whether there was any substitution a net increase in participation was reported. However, attempting to pinpoint (a) whether increased participation in sport through WCP is of sufficient intensity to result in wider social, and economic benefits; or (b) whether there is any impact on participation rates for individuals who are not on the WCP but who are influenced to participate by those who are on the WCP (i.e. the existence of *multiplier effects*), is problematic. Thus in many respects the data on the WCP captures gross rather than net impacts.

² For detailed data analysis please refer to Chen and Henry (2012).

The importance and issues of assessing the additionality

This brief example illustrates how, assessing additionality is not always a straightforward process. It requires a good understanding, judgement and knowledge of the intervention, together with sufficient information with which to assess claims along these four dimensions. Nevertheless the establishing of additionality remains critical in relation to the question of whether there is a rationale for a given intervention (Georghiou, 1998).

The addressing of the issue of additionality in evaluation studies, where it is undertaken at all, is reported by McEldowney (1997) as tending to rely on only one or two simple counterfactual-type questions; and when reporting the results of additional impacts of a programme, often only crude measures of additionality (i.e. high or low) are employed. Furthermore, information is often collected on these dimensions through key stakeholder interviews, but given stakeholder interests this data may be prone to bias, especially when interviewees are members of organisations in receipt of funding assistance.

4.4 Conclusion

This chapter has sought to provide an outline of a range of policy evaluation techniques that have recently begun to be adopted more frequently in the sport industry, illustrating their application with practical examples, together with critical analysis of the strengths and weaknesses of each.

In the case of the application of Logic Models, these are regarded in the field as being effectively employed to expose the causal assumptions of policy makers and programme deliverers.

In terms of meta-evaluation, its nature and peculiarities make it suitable for large-scale and comprehensive evaluation projects, which seek to develop evidence-based approaches to outcomes by gathering evidence from a range of projects or policy programmes. Such an approach invites the assessment of process (how adequate are the methods adopted to evaluate the policy programme?) and in terms of outcome (how robust are our assessment of the size and significance of impact). However, this approach has still not been widely adopted in sport-related evaluation, and certainly the use of meta-analysis is rarely likely to

be appropriate in the open system, dynamic policy contexts in which most sport policy interventions take place.

Similarly, in considering issues of additionality, while its importance has been underlined in the above discussion, it should be noted that, operationalising the key concepts has proved difficult. Nevertheless despite such problems of operationalisation, taken together these principles, mechanisms and models, provide conceptual tools for identifying the questions which it is crucial to address in providing realistic evidence-based policy evaluation in practice.

References

- Alter, C. and S. Murty 1997. Logic modelling: A tool for teaching practice evaluation. . *Journal of Social Work Education* 33: 103-117.
- Archer, M 1995. *Realist Social Theory: The Morphogenetic Approach* Cambridge Cambridge University Press.
- Ashworth, K., A. Cebulla, D. Greenberg and R. Walker 2004. Meta-evaluation: discovering what works best in welfare provision. *Evaluation* 10: 193-216.
- Bhaskar, R 1986. *Scientific Realism and Human Emancipation* London: Verso.
- Bhaskar, R 1978. *A Realist Theory of Science*. Hassocks: Harvester Press.
- Bickman, Leonard 1989. Barriers to the use of program theory. *Evaluation and Program Planning* 12: 387-390. doi: 10.1016/0149-7189(89)90056-6
- Biddle, S, M Hagger, N Chatzisarantis and S Lippke 2007. Theoretical Frameworks in Exercise Psychology. In *Theoretical Frameworks in Exercise Psychology*, eds. G Tenenbaum and R Ecklund, *Handbook of Sport Psychology*. Hoboken, N.J.: John Wiley & Sons.
- Buisseret, T.J., H.M. Cameron and L. Georghiou 1995. What difference does it make? Additionality in the public support of R&D in large firms. *International Journal of Technology Management* 10: 587-600.
- Bustelo, María 2002a. Meta-evaluation as a tool for the improvement and development of the evaluation function in public administrations. In *Meta-evaluation as a tool for the improvement and development of the evaluation function in public administrations, The 2002 EES Conference: three movements in contemporary evaluation: learning, theory and evidence*: University Complutense de Madrid.
- Bustelo, María 2002b. Meta-evaluation as a tool for the improvement and development of the evaluation function in public administrations. In *Meta-evaluation as a tool for the improvement and development of the evaluation function in public administrations, The 2002 EES Conference: three movements in contemporary evaluation: learning, theory and evidence*: University Complutense de Madrid.
- Chen, Shushu and Ian Henry 2012. A realist approach to policy evaluation for London 2012 Olympic legacy in a non-hosting region: a pilot study - an evaluation of the Workplace Challenge Programme In *A realist approach to policy evaluation for London 2012 Olympic legacy in a non-hosting region: a pilot study - an evaluation of the Workplace Challenge Programme The 20th EASM conference: Sport between business and civil society*, 207. Aalborg, Denmark: University College of Northern Denmark (in co-operation with the Danish Institute for Sports Studies) for the European Association for Sport Management.
- Conrad, Kendon J. and Frances L. Randolph 1999. Creating and using logic models: Four perspectives. *Alcoholism Treatment Quarterly* 17: 17-32.
- Cooksy, Leslie J. 1999. The Meta-Evaluand: The Evaluation of Project TEAMS. *American Journal of Evaluation* 20: 123-136. doi: 10.1177/109821409902000114
- Cooksy, Leslie J., Paige Gill and P. Adam Kelly 2001. The program logic model as an integrative framework for a multimethod evaluation. *Evaluation and Program Planning* 24: 119-128. doi: 10.1016/s0149-7189(01)00003-9
- Department of Media Culture and Sport 2008. *London 2012 Olympic and Paralympic Games Impacts and Legacy Evaluation Framework*. In *London 2012 Olympic and Paralympic Games Impacts and Legacy Evaluation Framework*. London: DCMS.
- Dodd-Butera, Teresa and Molly Broderick 2011. Health Policy and Poison Control Centers: Providing Analysis Utilizing a Logic Model. *Clinical Toxicology* 49: 603-603.

- English Partnership 2008. *Additionality Guide - A standard approach to assessing the additional impact of interventions* In *Additionality Guide - A standard approach to assessing the additional impact of interventions*
- Georghiou, Luke 1998. *Issues in the Evaluation of Innovation and Technology Policy*. *Evaluation* 4: 37-51. doi: 10.1177/13563899822208374
- Government Social Research Unit 2007. *The Magenta Book: Guidance Notes for Policy Evaluation and Analysis*. In *The Magenta Book: Guidance Notes for Policy Evaluation and Analysis*. London: HM Treasury.
- Grant Thornton, ECORYS and Centre for Olympic Studies and Research Loughborough University 2011a. *Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games Summary of Reports 1 and 2: 'Scope, research questions and strategy' and 'Methods'*. In *Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games Summary of Reports 1 and 2: 'Scope, research questions and strategy' and 'Methods'*. London: Department of Culture Media and Sport.
- Grant Thornton, ECORYS and Centre for Olympic Studies and Research Loughborough University 2011b. *Report 1 – Scope, research questions and strategy: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games*. In *Report 1 – Scope, research questions and strategy: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games*. London: Department of Culture Media and Sport.
- Grant Thornton, ECORYS and Centre for Olympic Studies and Research Loughborough University 2011c. *Report 2: Methods: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games* In *Report 2: Methods: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games* London: Department of Culture Media and Sport.
- Grant Thornton, ECORYS and Centre for Olympic Studies and Research Loughborough University 2012. *Report 3: Baseline and Counterfactual: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games* In *Report 3: Baseline and Counterfactual: Meta-Evaluation of the Impacts and Legacy of the London 2012 Olympic Games and Paralympic Games* London: Department of Culture Media and Sport.
- Gullickson, A., L. Wingate, F. Lawrenz and Chris L. S. Coryn 2006. *The NSF advanced technology education program: Final evaluation report*. In *The NSF advanced technology education program: Final evaluation report*. Kalamazoo:: Western Michigan University: The Evaluation Center.
- Head, B 2009. *Evidence-based Policy: Principles and Requirements*. In *Evidence-based Policy: Principles and Requirements*, ed. Productivity Commission, *Strengthening Evidence-based Policy in the Australian Federation*. Canberra: Productivity Commission,.
- HM Treasury 2003. *The Green Book: Appraisal and Evaluation in Central Government Treasury Guidance*. In *The Green Book: Appraisal and Evaluation in Central Government Treasury Guidance*, ed. HM Treasury. London.
- Jordan, Gretchen B. 2010. *A theory-based logic model for innovation policy and evaluation*. *Research Evaluation* 19: 263-273. doi: 10.3152/095820210x12827366906445
- Leeuw, F and J Vaessen 2009. *Impact Evaluations And Development: NONIE Guidance on Impact Evaluation* In *Impact Evaluations And Development: NONIE Guidance on Impact Evaluation* ed. Network of Networks on Impact Evaluation: World Bank.
- Lenihan, Helena 2011. *Enterprise policy evaluation: Is there a 'new' way of doing it?* *Evaluation and Program Planning* 34: 323-332. doi: 10.1016/j.evalprogplan.2011.03.006
- Luukkonen, Terttu 2000. *Additionality of EU framework programmes*. *Research Policy* 29: 711-724. doi: 10.1016/s0048-7333(99)00041-4

- Madzivhandila, T. P., G.R. Griffith, E. Fleming and A.E. Nesamvuni 2010. Meta-evaluations in government and government institutions: a case study example from the Australian Centre for International Agricultural Research. In *Meta-evaluations in government and government institutions: a case study example from the Australian Centre for International Agricultural Research*, the Annual Conference Australian Agricultural and Resource Economics Society.
- McEldowney, James J. 1997. Policy Evaluation and the Concepts of Deadweight and Additionality. *Evaluation* 3: 175-188. doi: 10.1177/135638909700300204
- Mulroy, E.A. and H. Lauber 2004. A user-friendly approach to program evaluation and effective community interventions for families at risk of homelessness. *Social Work* 49: 573-586.
- Nilsson, Neil and Donald Hogben 1983. Metaevaluation. *New Directions for Program Evaluation* 1983: 83-97. doi: 10.1002/ev.1346
- Patton, M 1978. *Utilisation-Focused Evaluation*. Beverly Hills: CA: Sage.
- Patton, M.Q. 1997. *Utilization-Focused Evaluation-The New Century Text*. Thousand Oaks: Sage.
- Pawson, Ray 2006. *Evidence-based policy : a realist perspective*. London: SAGE.
- Pawson, Ray and Nick Tilley 1997. *Realistic evaluation*. London: SAGE.
- Russ-Eft, Darlene and Hallie Preskill 2008. Improving the quality of evaluation participation: a meta-evaluation. *Human Resource Development International* 11: 35-50. doi: 10.1080/13678860701782311
- Scott-Little, Catherine, Mary Sue Hamann and Stephen G. Jurs 2002. Evaluations of After-School Programs: A Meta-Evaluation of Methodologies and Narrative Synthesis of Findings. *The American Journal of Evaluation* 23: 387-419. doi: Doi: 10.1016/s1098-2140(02)00234-5
- Scriven, M. 1991. *Evaluation Thesaurus*. Newbury Park: SAGE.
- Scriven, M. 1969. An introduction to meta-evaluation. *Educational Products Report* 2: 36-38.
- Stake, R.E. 1967. Countenance of educational evaluation. *Teachers College Record* 68: 523-540.
- Stufflebeam, D. 2001. The Metaevaluation Imperative. *American Journal of Evaluation* 22: 183-209. doi: 10.1177/109821400102200204
- Stufflebeam, Daniel L. 1974. *Meta-evaluation*. In *Meta-evaluation*. Kalamazzo: MI: Western Michigan University Evaluation Center.
- Stufflebeam, Daniel L. 1971. The Relevance of the CIPP Evaluation Model for Educational Accountability. *Journal of Research and Development in Education* 5: 19-25.
- Taylor-Powell, E. 1999. Logic model notes. In *Logic model notes*, Paper presented at the University of Wisconsin's providing leadership for program evaluation. Madison: Wisconsin.
- Weiss, Carol H. 1997. How Can Theory-Based Evaluation Make Greater Headway? *Evaluation Review* 21: 501-524. doi: 10.1177/0193841x9702100405
- Wholey, Joseph. 1994. Assessing the feasibility and likely usefulness of evaluation. In *Assessing the feasibility and likely usefulness of evaluation*, eds. Joseph. Wholey, H.P. Hatry and K.E. Newcomer, *Handbook of practical program evaluation* 15-39. San Francisco:: Jossey-Bass.
- Yin, R.K. 2009. *Case study research: Design and methods*. California: Sage Publications, Inc.
- Yin, Robert K. 1998. The abridged version of case study research: Design and method. In *The abridged version of case study research: Design and method*, ed. L. Bickman D. J. Rog, *Handbook of applied social research methods*, 229-259. Thousand Oaks, CA, US: Sage Publications, Inc.