2012

Measuring the performance of partnerships: Why, what, how, when?

Kelly, C

10.1111/j.1749-8198.2012.00476.x
Geography Compass
Blackwell Publishing Ltd

THE FULL VERSION OF THIS PAPER, INCLUDING FIGURES AND TABLES, CAN BE FOUND AT: Geography Compass, (2012) 6(3), pp149-162

Measuring the performance of partnerships: Why, what, how, when?

Abstract

Partnerships have become increasingly prevalent across a wide range of sectors for the delivery of services and implementation of policy. Partnerships are seen as a more effective way of delivering policy interventions than state-led or 'top-down' approaches. Evaluating partnership performance is therefore crucial in order to determine whether partnerships really are better than more traditional methods of policy implementation. To date, however, partnership effectiveness has often been conceptualised as cumulative; the result of a set of variables acting in a one-dimensional, linear way which results in the ability (or not) of a partnership to achieve its goals. This paper highlights the shortcomings of such a linear conceptualisation of effectiveness and argues instead that when evaluating partnerships, effectiveness should be viewed as a non-linear, multi-faceted composite which changes in space and time.

Introduction

Partnership working is characterised by a coming together of organisations and individuals to resolve conflict or address specific issues which cannot be resolved by the organisations or individuals acting alone. Such approaches are also increasingly seen as a way of empowering individuals to take an active role in identifying and delivering

their own needs, and in improving the effectiveness of policy interventions (Cabinet Office, 2010).

The literature is clear about the proliferation of such approaches and yet recognises that, to date, there has been a lack of empirical evidence to suggest whether they are any better at achieving their aims than more traditional policy interventions (Dowling et al., 2004, Stojanovic and Ballinger, 2009). In addition, some authors have questioned whether these new institutions simply act as vehicles through which the state continues to enact its own policy goals in a top-down manner, rather than empowering more inclusive participation in policy formulation and implementation (Holzinger et al., 2006, Imrie and Raco, 1999, Jordan et al., 2005, Kearns, 1992). The need to evaluate the effectiveness of partnership approaches is therefore clear, yet there is little agreement in the literature on the theoretical and methodological frameworks that should be used. Indeed, a 'one size fits all' approach to evaluation is unlikely to be appropriate given the diversity of the types of partnerships and collaborations that exist and the complexity of the environments in which they operate.

The purpose of this paper is to comment on the various approaches that have been developed to evaluate partnership effectiveness, and to suggest an alternative conceptualisation of effectiveness which might offer a more accurate reflection of the dynamic nature of partnership performance. The paper begins with an examination of the theoretical basis which underpins approaches to evaluation. The various purposes of evaluation are discussed, as is the evolution of different methodological approaches. Questions of 'what', 'how' and 'when' to evaluate are addressed. Next, some of the practical challenges which arise in measuring the performance of partnerships are

discussed and the difficulties in choosing and applying appropriate measures of success are highlighted. An alternative conceptualisation, which recognises the importance of changes in the context and process of partnership action, is suggested.

Why evaluate?

Although the development of theories of evaluation is relatively young, its growth over the past twenty years has been exponential. Indeed, in their review of the development of different approaches, Pawson and Tilley (1997, p.1) suggest that the impulse to evaluate has become endemic:

> The enterprise of evaluation is a perfect example of what Kaplan (1964) once called the 'law of the hammer'. Its premise is that if you give a child a hammer then he or she will soon discover the universal truth that everything needs pounding. In a similar manner, it has become axiomatic [...] that *everything, but everything needs evaluating*. (original emphasis).

Definitions for the term 'evaluation' are as diverse as the approaches used and the environments and participatory mechanisms within which it is applied (Chess, 2000, Oels, 2006). Chelimsky and Shadish (1997, p. xii), however, neatly sum up the term as being '*about determining merit or worth*'. Evaluation can serve many different purposes, and the approach taken will depend on the motivation which lies behind the drive to evaluate. Capwell *et al.* (2000) note six primary reasons for evaluating:

(i) To determine the achievement of aims or objectives

A common way to measure performance is by assessing the achievement of stated aims or objectives, usually at the end of an intervention, programme or partnership (*ex ante* evaluation). *Ex ante* evaluation is a type of summative assessment and is one of the best understood purposes for evaluation (Chess, 2000). The success of a

programme is assessed in terms of its ability to deliver planned outcomes and the results are often used to compare programmes to determine which approach works best in any given situation.

(ii) To improve programme implementation

Another of the important drivers for evaluation is the need to provide feedback and assess progress during the lifetime of a programme, in order to forecast the likelihood of achieving objectives and to make any necessary adjustments to ensure success (mid-term evaluation). This type of evaluation is also summative, as it is designed to measure performance against specific criteria. However, this approach also includes elements of formative evaluation by examining the way that a programme is being implemented and by seeking ways to improve delivery (Chess, 2000).

(iii) To provide accountability to funders, communities and other stakeholders

Providing measures of financial accountability is another well understood purpose of evaluation. In assessing performance, decisions can be made on the benefits of a programme relative to the costs associated with its implementation. In times of funding restriction, cost-benefit analysis can provide important insights into how limited resources can be used to maximum effect (Oels, 2006).

(iv) To increase community support for initiatives

Increasing community support can be an important mechanism for raising the profile of an initiative and thereby securing further funding and support for the future development of the initiative. Reflecting on and evaluating the performance of an initiative can provide

useful data, which can then be disseminated through various media to help engender support and widen the engagement and participation of stakeholders.

(v) To contribute to the scientific basis for interventions

Evaluation for information on the achievement of programme outcomes or long-term changes represents another type of assessment; that of impact evaluation. Tracking the long-term outcomes from an intervention or programme, however, can be difficult to achieve and expensive. Difficulties include showing that changes are achieved as a direct result of the intervention rather than other external variables. Although impact evaluation demands an ongoing commitment to continue monitoring activity long after the intervention has ceased, this type of evaluation can offer long-term data which can help to provide the basis for future policy decisions (Chess, 2000). In addition, this type of evaluation may also be driven by academic interest in establishing empirical evidence from which to refine and adapt theoretical models (Oels, 2006).

(vi) To inform policy decisions

Evaluation data can be used for policy development in two specific ways. Impact evaluation data can be used to '*move political will and make investments in particular areas more likely*' by providing empirical evidence of the success of particular types of intervention (Capwell et al., 2000, p.19). Evidence from evaluation can also be used in a reflexive manner to refine existing policy objectives and make them more effective in practice through a process of policy learning (Sanderson, 2002). This type of embedded evaluation forms a crucial element of adaptive management practice and is particularly prevalent in environmental policy initiatives (Day, 2008, Hockings et al., 2000)

The six reasons for evaluation listed above provide an indication of the wide variety of purposes for evaluation. These purposes, in turn, form the basis for the identification of criteria against which aspects of partnership should be evaluated.

What to evaluate, and how to evaluate it?

As the imperative to develop collaborative approaches to address social and environmental issues has burgeoned, so too has the variety of approaches to evaluation (Pawson and Tilley, 1997). Table 1 provides a range of examples to highlight the different contexts within which partnerships exist and some of the approaches taken in evaluating them. Some studies, such as those by Bramwell and Sharman (1999), Fletcher (2003) and Barker (2004) have evaluated elements of the process of collaboration. These authors have found that certain factors, such as the ability of stakeholders to participate in decision making and implement agreed actions, are crucial in enabling partnerships to achieve their goals. Other authors, such as Backstrand (2006), have highlighted the impact that the institutional context in which partnerships operate can have in enabling or preventing partnerships from achieving their goals. In practice, current evaluation programmes tend to draw on a range of tools from multiple approaches, in order to avoid the shortcomings associated with using one single approach.

Positivist approaches

Approaches centred on the positivist tradition attempt to isolate the specific 'ingredients' of programme success from the mass of potential variables. This type of evaluation

relates most closely to laboratory-based ‘experimental’ methods, whereby the impact of a single variable on performance is measured, and all other variables are excluded. The central objective of this type of evaluation is to demonstrate a causal relationship between the action and the output: namely did the application of *x* cause the observed change in *y*, for example. In this type of evaluation, quantitative indicators based on desired outputs may be selected. So, for example, in the evaluation of a coastal habitat management programme, an indicator might be selected to show the total area of wetland habitat restored by the end of the programme (Ehler, 2003). This type of positivist output indicator measures the level of attainment of a specific target or goal, generally at the end of a programme (Dixon and Sindall, 1994).

A key difficulty with the positivist approach is the heterogeneity of contexts within which partnerships operate. Given this heterogeneity, it can be extremely difficult to isolate the specific causal factors and then to apply them in a vacuum. Real-world situations are inherently complex and partnerships and programmes are subject to a range of dynamic endogenous and exogenous variables. The positivist approach therefore offers little benefit to partnership evaluation because it fails to take account of the spatial and temporal complexity in real-world situations and ignores the effect of this complexity on partnership performance.

Constructivist approaches

Constructivist approaches to evaluation emerged in the 1970s, as a reaction against the positivist experimental paradigm (Pawson and Tilley, 1997). Instead of attempting to find the generic principles behind an intervention, constructivist approaches focus instead on the actors and processes within a partnership and the impact that their

perceptions and understandings have on the success of partnership actions (Guba and Lincoln, 1981). In constructivist approaches, qualitative methods are the dominant paradigm. The constructivist approach led to one of the most important changes in evaluation research: namely the shift away from a focus on quantifying outputs towards a qualitative emphasis on the processes involved. As a result, evaluation research began to recognise the diversity of understanding and expectation about a programme and its likely performance, that exists between stakeholders, practitioners and policy makers and the resultant impact that those multiple views could have on the success of the intervention (Guba and Lincoln, 1981).

There are difficulties, however, in focussing almost exclusively on the process of partnership rather than the outcomes. By viewing programmes or interventions as sets of negotiated understandings between different groups of stakeholders, each context is argued to be entirely unique and so provides no help in establishing external validity (Pawson and Tilley, 1997, Blamey and Mackenzie, 2007).

Pragmatic evaluation

In response to the difficulties associated with isolating the specific factors for success in complex environments, and the need for evaluation which could inform policymaking, a new suite of evaluation approaches called ‘pragmatic’ evaluations emerged during the 1990s. Pragmatic evaluation links the choice of evaluation tools (quantitative, qualitative or a mixed method) directly to the purposes of the research (Patton, 1997). The ability of multi-sector partnerships to raise the profile of an issue beyond immediate partners and to bring it to the attention of the general public and policymakers is an example of a pragmatic evaluation goal. This type of evaluation has been used to evaluate the

success of development initiatives such as Local Strategic Partnerships and rural development programmes (Goodwin, 1998, Geddes et al., 2007). Evaluation of the performance of these types of partnerships provides insight into the reality of purported new governance approaches by indicating whether new policies have been developed as a result of the partnership process (Forsyth, 2005).

As with positivist and constructivist approaches, however, the narrow scope and focus of pragmatic evaluation, driven by the specific needs of the end user can result in strong internal validity but weak external validity. The consequence is often that the wider goal of understanding *why* a specific intervention works in a specific context is lost (Chen and Rossi, 1983, Chen, 1990).

Theory-based evaluations

The approaches to evaluation described above are characterised by their focus on methods. Given that none of these method-led approaches was fully able to meet the needs of evaluating multi-dimensional partnership interventions, an alternative pluralistic approach called 'theory-based evaluation' was developed during the 1990s (Sullivan and Stewart, 2006, Chen, 1990). Theory-based evaluations grew out of programme theory and attempt to map the entire process of partnership, rather than inferring that success is the result of specific inputs and outputs (Cronbach, 1982, Dickinson, 2006, Cronbach, 1963, Hall, 2004). Two main theory-based approaches predominate: 'realistic evaluation' and 'theory of change'.

'Realistic evaluation' and 'theory of change' approaches

The ‘theory of change’ and ‘realistic evaluation’ approaches use theoretical and contextual understandings of the drivers for collaborative action to inform the evaluation process by shaping the specific research priorities and guiding the questions that the evaluation will seek to address (Blamey and Mackenzie, 2007, Connell and Kubisch, 2002). Important differences exist, however, between the two approaches (Dickinson, 2006). The ‘theory of change’ approach is prospective: the evaluation process is embedded within the programme itself and is an iterative process. This type of evaluation is better suited to strategic evaluations of large-scale, multi-site or whole community programmes because of its stronger emphasis on programme outcomes and how change is being achieved. ‘Realistic evaluation’, on the other hand, is retrospective, with the evaluator remaining outside of the partnership being evaluated, and is better suited to micro-scale evaluations where the local conditions can provide important insights into why specific components of a programme work in a particular context.

The ‘realistic evaluation’ approach divides programmes into three components; the context within which it operates (C), the mechanism used to deliver the programme (M) and the outcomes achieved (O). The same programme applied in differing contexts, it is theorised, can therefore lead to a variety of outcomes or CMO configurations. These different configurations provide a cumulative understanding of what works, for whom, and under what circumstances (Befani et al., 2007).

Hasnain-Wynia *et al.* (2003) provide an useful visualisation of the key characteristics and measures of a ‘realistic evaluation’ approach, as used in their evaluation of community care network partnerships (Figure 1). The framework shown in Figure 1 also

encapsulates elements of Waddock and Bannister's (1991) 'interaction amongst partners' by acknowledging the role of previous collaboration experience, and community perceptions and understanding of the need for partnership action, under 'environmental characteristics'. The various conditions which exist prior to the establishment of a partnership, together with the specific geographical context within which it will operate, will have a significant impact on multiple aspects of the process, and are therefore treated as an integral element of the evaluation.

When to evaluate?

An important limitation in all of the approaches to evaluation described above is their application as linear processes (Dickinson, 2006, Sanderson, 2002). This linearity presents particular difficulties for the evaluation of partnerships which have no specific time frame or life expectancy, and therefore no clearly defined or obvious point at which they should be evaluated (Rowe and Frewer, 2004). The point at which an evaluation of partnership performance is undertaken will clearly have an impact on the findings of that evaluation (El Ansari et al., 2001). Levels of effectiveness in partnerships may change in response to internal dynamics or external contextual changes. Therefore, the point at which evaluation takes place may be crucial in understanding the reasons for success or failure. For example, viewed from a single temporal standpoint, a partnership may seem efficient, networked and progressive when it may in fact have undergone a series of crises or flux based around specific issues, the *resolution* of which has resulted in the emergence of new collaborative cohesion and the achievement of stated goals.

Questions surrounding when to evaluate also affect the external validity of the findings of an evaluation. The difficulty of identifying the specific point in a partnership's life that evaluation should be undertaken is compounded if a comparative methodology is also used. For example, if a number of partnerships are selected for comparative study, they are unlikely to have begun to collaborate at exactly the same point in time, and will have taken differing lengths of time to reach maturity. Therefore, if partnerships with similar objectives, but operating in different contexts, are compared in order to determine the impact of context on partnership effectiveness, differences in their effectiveness may be because they are at different stages of development rather than because of differences in the way that they operate. Comparative evaluation of partnership performance therefore needs to draw on a detailed understanding of the endogenous and exogenous changes which have occurred during the lifetime of the partnership in order to draw robust conclusions about partnership effectiveness.

The way that partnership effectiveness is conceptualised over time also has important implications for the way that the effectiveness of policy intervention is evaluated. Figure 2 illustrates graphically different typologies of time. 'Clock' time (Figure 2 (a)) refers to:

> '*the continuum – that is, time as a non-spatial dimension in which events occur in apparently irreversible succession from the past through the present to the future*' (Ancona et al., 2001, p.514).

'Cyclical' time (Figure 2 (b)) refers to the recurring patterns which occur in the continuum of time, such as the seasons of the year, for example. 'Life-cycle' time (Figure 2 (d)) may include a cyclical process, but is delineated by clear start and end-points and, unlike cyclical time, is not necessarily repeated. Each of these typologies

views time as progressing in a specific linear direction. Current approaches to evaluation use this linear conceptualisation of time to view effectiveness as a cumulative attribute, increasing through accretion over time (Figure 3). However, this paper argues that such a cumulative and linear view of effectiveness is inaccurate, as variables within and outside of partnerships do not remain static, but can change (in a positive or negative way) according to internal and external dynamics.

Contrasting with notions of linear time is ‘event’ time. Event time may be predictable (Figure 2 (e)) in that an event is regularly repeated, or it may be repeated at irregular intervals (Figure 2 (c)) or it may be a singular event which is not repeated. This non-linear conceptualisation of time may be a more useful and accurate way to understand how the effectiveness of partnerships develops. Figure 4 illustrates the conceptualisation of effectiveness as a variable process.

As can be seen in Figure 4, effectiveness changes in response to both positive stimuli (light stars) and negative stimuli (dark stars). In this non-linear conceptualisation, the level of effectiveness at any one point is the product of a suite of variables, or determinants of effectiveness, acting from both within and outside of the partnership, and which combine to produce a composite, termed ‘effectiveness’.

Practical challenges in measuring the performance of partnerships

Partnership approaches vary considerably in the way that organisations and individuals work together to achieve common goals. Partnerships exhibit differences in terms of their scale, structure, composition and agency (Rowe and Frewer, 2004, Selin, 1999). In addition as discussed above, partnerships operate within dynamic policy and

295 institutional contexts and are themselves subject to change in terms of stakeholder
296 engagement and resource availability (Sanderson, 2002). As a result, measuring
297 partnership performance is difficult.

298 As a first step, it is important to set clear and unambiguous criteria for assessing
299 success. However, achieving this goal in practice is not straightforward and will depend
300 on the chosen evaluation methodology. Difficulties exist in agreeing which indicators of
301 success to use and in reaching consensus on the level of achievement of each indicator
302 (Dixon and Sindall, 1994). In addition, conceptualisations of 'success' may vary
303 between individual stakeholders, particularly if the partnership has been established or
304 led by a top-down imperative and participants have had little or no opportunity to be
305 involved during the early stages of partnership formation (El Ansari et al., 2001,
306 Glendinning, 2002).

307 Partnerships may achieve a wide range of benefits which are not necessarily identified
308 as target outcomes. For example, a partnership in a coastal environment may not have
309 achieved its stated objective of publishing a management plan, but may still have
310 developed a shared sense of purpose and cohesion amongst divergent stakeholder
311 groups, which later facilitates the implementation of other partnership actions
312 (Stojanovic and Barker, 2008). In this respect, there may be value in the act of
313 partnership itself (Asthana et al., 2002).

314 Examining the effectiveness of a single local partnership in inherently complex 'real
315 world' environments carries its own set of specific difficulties. The challenge becomes
316 even greater when attempting to evaluate the same criteria across multiple case studies

in order to achieve comparability (Freeman and Peck, 2006). Specifically, the differing contexts and spatial scales within which partnerships work, can have a significant influence on the way that the same set of management principles are interpreted and implemented (Dahl-Tacconi, 2005). For example, differing legislative and policy environments can lead to significantly different ways of enacting agreed actions and interventions. Changes in one variable, such as staff turnover within partner organisations, for example, may lead to significant changes in other variables, such as the availability of resources or the confidence with which representatives can make decisions (Fletcher, 2007, Freeman and Peck, 2006). The differing cultural contexts of partner organisations, and mismatches in the spatial scales at which these organisations work, can also have an impact on the way that partnerships operate (Freeman and Peck, 2006, Evans, 2004). Evans (2004), for example, has described the degree to which individual or personal agendas and interests shape, promote or restrict the achievement of collective objectives.

An alternative approach: Mapping the 'determinants of effectiveness' through time

Given the challenges to evaluation noted above, dividing partnerships into three elements, namely: *context*, *process* and *outcome*, as espoused in the 'realistic' approach to evaluation, offers a useful framework within which to work (Figure 5). Associated with each of the three elements is a set of 'determinants of effectiveness' (Kelly, 2009, Kelly et al., in press). These determinants of effectiveness have been drawn from the literature on evaluation across a broad range of fields and are

considered to be the key generic ingredients for successful partnership working (Asthana et al., 2002, Dowling et al., 2004). An empirical study which tests this potential new approach has been undertaken and a discussion of the key findings can be found in Kelly et al. (in press).

Determinants of effectiveness associated with the context within which a partnership works include a 'pro-partnership' political and cultural climate, in which partnership action is seen as the most appropriate method for dealing with the identified issue, and which is particularly important in driving the early stages of partnership formation. Determinants of effectiveness associated with the process of partnership include; the degree to which all relevant stakeholders are identified and given an opportunity to participate in the process; the level of commitment that stakeholders have to remain actively engaged in partnership activity; and the degree to which levels of trust exist between stakeholders from different sectors. The important determinants of output and outcome effectiveness include; the extent to which stakeholders are prepared to abide by collectively agreed actions; the degree to which partnership objectives have been realised; and the ability of the partnership to shape and influence future policy (Oels, 2006).

Each individual determinant plays an important role in contributing to the overall effectiveness of the process and to the perceptions of effectiveness held by stakeholders within and outside of the partnership (Hasnain-Wynia et al., 2003). There can be elements of overlap of course between the determinants, as benefits which emerge from the *process* (such as increased levels of trust and understanding between stakeholders, for example), may also be viewed as partnership achievements or

outcomes (Chess, 2000). This interconnectedness is shown in Figure 5 by thin black arrows which link the determinants of process effectiveness to the determinants of output/outcome effectiveness. In addition, a large arrow links the achievements of the partnership back to the context within which it operates, to highlight the notion that partnership activity is embedded within the places and spaces in which it operates, and will therefore have an impact on that context throughout its actions and achievements. This influence may be both positive and negative. For example, if there has been little experience of partnership working prior to the new collaboration, as the partnership progresses, stakeholder perceptions of the purpose and value of the partnership change (either positively or negatively), altering their willingness to continue to participate and influencing the opinions and perceptions of other stakeholders outside of the partnership who may be considering joining.

Attaining consistently high levels of achievement of the determinants of effectiveness is difficult and unrealistic. It is much more likely that performance will be fluid and dynamic, with good levels of achievement of different determinants at different times. Successful partnerships may therefore be characterised by the maintenance of good levels of achievement across multiple determinants over a prolonged period of time (Kelly, 2009, Kelly et al., in press). The goal of evaluating partnership effectiveness should therefore be to identify why partnerships have failed to achieve high levels of the key determinants of effectiveness and to provide insight into how any decline in performance can be improved. Each of the challenges noted above highlights the need for evaluation approaches which reflect not only the context, mechanism and outcomes of a particular partnership, but also *changes* in those three elements over time.

Conclusions

Although collaboration has become one of the most common ways of delivering policy and managing complex environments, there is little empirical evidence to demonstrate whether such approaches are more effective than more traditional top-down methods of intervention. No single research paradigm is universally applicable. The choice of theoretical perspective and practical evaluation method will depend on the purposes, users and sponsors of the evaluation. Early evaluative techniques were rooted in the positivist experimental paradigm, whereby researchers looked for the 'universal truths' of causality by isolating and testing specific 'factors for success'. Later, a constructivist paradigm emerged and the focus shifted away from an evaluation of programme outputs onto the processes of collaboration itself. More recently, a pluralist approach has become the central paradigm, whereby the context within which programmes and partnerships operate is examined, alongside the more traditional measures of processes, outputs and outcomes.

A particularly useful approach is that provided by 'realistic evaluation'. 'Realistic evaluation' offers a number of advantages over alternative strategies in that it acknowledges the need to build on knowledge of 'what works' in order to progress understanding, whilst also accepting that differing contexts can lead to important differences in outcomes. In this way, 'realistic evaluation' provides an opportunity to define 'what works when'.

Existing approaches, however, have failed to acknowledge the impact of change on the various components which comprise partnership effectiveness. Changes in these

variables are likely to result in changing levels of effectiveness over time and this perspective has significant implications for the point at which evaluation is undertaken, particularly for those partnerships which are not delimited by specific time frames. An alternative approach is posited, based on analysis of changes over time in key 'determinants of effectiveness'. Tracing changes in the achievement of the determinants of effectiveness can highlight not only whether a partnership is likely to achieve its goals, but also why it has performed as it has, and what is needed in order to improve performance in the specific local context within which the partnership operates.

References

ANCONA, D. G., OKHUYSEN, G. A. & PERLOW, L. A. 2001. Taking time to integrate temporal research. *Academy of Management Review,* 26**,** 512-529.

ASTHANA, S. N. M., RICHARDSON, S. & HALLIDAY, J. 2002. Partnership working in public policy provision: a framework for evaluation. *Social Policy & Administration,* 36**,** 780-795.

BACKSTRAND, K. 2006. Multi-stakeholder partnerships for sustainable development: rethinking legitimacy, accountability and effectiveness. *European Environment,* 16**,** 290-306.

BARKER, A. 2004. Capacity building for sustainability: towards community development in coastal Scotland. *Journal of Environmental Management,* 75**,** 11-19.

BEFANI, B., LEDERMANN, S. & SAGER, F. 2007. Realistic evaluation and QCA: conceptual parallels and an empirical application. *Evaluation,* 13**,** 171 - 192.

BLAMEY, A. & MACKENZIE, M. 2007. Theories of change and realistic evaluation: peas in a pod or apples and oranges? . *Evaluation,* 13**,** 439 - 455.

BRAMWELL, B. & SHARMAN, A. 1999. Collaboration in local tourism policymaking. *Annals of Tourism Research,* 26**,** 392-415.

CABINET OFFICE 2010. The Compact. The Coalition Government and civil society organisations working effectively in partnership for the benefit of communities and citizens in England. *In:* OFFICE, C. (ed.). London: Cabinet Office.

CAPWELL, E. M., BUTTERFOSS, F. & FRANCISCO, V. T. 2000. Evaluation in practice: why evaluate? *Health Promotion Practice,* 1**,** 15-20.

CHELIMSKY, E. & SHADISH, W. (eds.) 1997. *Evaluation for the 21st century,* Thousand Oaks: Sage.

CHEN, H.-T. 1990. *Theory-driven evaluaions,* Newbury Park, Sage Publications.

CHEN, H.-T. & ROSSI, P. 1983. Evaluating with sense: the theory-driven approach. *Evaluation Review,* 7**,** 283-302.
CHESS, C. 2000. Evaluating environmental public participation: methodological questions. *Journal of Environmental Planning and Management,* 43**,** 769-784.
CONNELL, J. P. & KUBISCH, A. C. 2002. Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects and problems. *In:* CONNELL, J. P., KUBISCH, A. C. & FULBRIGHT ANDERSEN, K. (eds.) *New approaches to evaluating community initiatives Volume 2: theory, measurement and analysis.* Washington D.C.: Aspen Institute.
CRONBACH, L. 1963. 'Course improvement through evaluation'. *Teachers College Record,* 64**,** 672-683.
CRONBACH, L. 1982. *Designing evaluations of educational and social programmes,* San Francisco, Jossey-Bass.
DAHL-TACCONI, N. 2005. Investigating information requirements for evaluating effectiveness of marine protected areas - Indonesian case studies. *Coastal Management,* 33**,** 225-246.
DAY, J. 2008. The need and practice of monitoring, evaluating and adapting marine planning and management - lessons from the Great Barrier Reef. *Marine Policy,* 32**,** 823 - 831.
DICKINSON, H. 2006. The evaluation of health and social care partnerships: an analysis of approaches and synthesis for the future. *Health and Social Care in the Community,* 14**,** 375 - 383.
DIXON, J. & SINDALL, C. 1994. Applying logics of change to the evaluation of community development in health promotion. *Health Promotion International,* 9**,** 297-309.
DOWLING, B., POWELL, M. & GLENDINNING, C. 2004. Conceptualising successful partnerships. *Health and Social Care in the Community,* 12**,** 309-317.
EHLER, C. N. 2003. Indicators to measure governance performance in integrated coastal management. *Ocean & Coastal Management,* 46**,** 335-345.
EL ANSARI, W., PHILLIPS, C. J. & HAMMICK, M. 2001. Collaboration and partnerships: developing the evidence base. *Health and Social Care in the Community,* 9**,** 215-227.
EVANS, J. 2004. What is local about local environmental governance? Observations from the local biodiversity action planning process. *Area,* 36**,** 270-279.
FLETCHER, S. 2003. Stakeholder representation and the democratic basis of coastal partnerships in the UK. *Marine Policy,* 27**,** 229-240.
FLETCHER, S. 2007. Influences on stakeholder representation in participatory coastal management programmes. *Ocean & Coastal Management,* 50**,** 314-328.
FORSYTH, T. 2005. Building deliberative public-private partnerships for waste management in Asia. *Geoforum,* 36**,** 429-439.
FREEMAN, T. & PECK, E. 2006. Evaluating partnerships: a case study of integrated specialist mental health services. *Health and Social Care in the Community,* 14**,** 408-417.
GEDDES, M., DAVIES, J. & FULLER, C. 2007. Evaluating local strategic partnerships: theory and practice of change. *Local Government Studies,* 33**,** 97-116.

GLENDINNING, C. 2002. Partnerships between health and social services: developing a framework for evaluation. *Policy & Politics,* 30**,** 115-127.
GOODWIN, M. 1998. The governance of rural areas: some emerging research issues and agendas. *Journal of Rural Studies,* 14**,** 5-12.
GUBA, Y. & LINCOLN, E. 1981. *Effective evaluation: improving the usefulness of evaluation results through responsive and naturalistic approaches,* San Francisco, Jossey Bass.
HALL, J. 2004. Pluralistic evaluation: a situational approach to service evaluation. *Journal of Nursing Management,* 12**,** 22-27.
HASNAIN-WYNIA, R., SOFAER, S., BAZZOLI, G., ALEXANDER, J. A., SHORTELL, S., CONRAD, D. A., CHAN, B., ZUKOSKI, A. P. & SWENEY, J. 2003. Members' perceptions of community care network partnerships' effectiveness. *Medical care research and review,* 60**,** 40s - 60s.
HOCKINGS, M., WITH STOLTON, S. & DUDLEY, N. 2000. Evaluating effectiveness: a framework for assessing the management of protected areas. *In:* PHILLIPS, A. (ed.) *Best practice protected area guidelines series.* Gland: IUCN.
HOLZINGER, K., KNILL, C. & SCHAFER, A. 2006. Rhetoric or reality? 'new governance' in EU environmental policy. *European Law Journal,* 12**,** 403-420.
IMRIE, R. & RACO, M. 1999. How new is the new local governance? Lessons from the United Kingdom. *Transactions of the Institute of British Geographers,* 24**,** 45-63.
JORDAN, A., WURZEL, R. K. & ZITO, A. 2005. The rise of 'new' policy instruments in comparative perspective: has governance eclipsed government? *Political Studies,* 53**,** 477-496.
KAPLAN, A. 1964. *The conduct of inquiry: methodology for behavioral science,* San Francisco, Chandler.
KEARNS, A. J. 1992. Active citizenship and urban governance. *Transactions of the Institute of British Geographers,* 17**,** 20-34.
KELLY, C. L. 2009. *Partnerships in the management and development of marine nature-based tourism: an analysis of effectiveness.* Doctor of Philosophy PhD, University of Plymouth.
KELLY, C. L., ESSEX, S. & GLEGG, G. in press. Reflective Practice for Marine Planning: A Case Study of Marine Nature-based Tourism Partnerships. *Marine Policy*.
OELS, A. 2006. Evaluating stakeholder dialogues. *In:* STOLL-KLEEMAN, S. & WELP, M. (eds.) *Stakeholder dialogues in natural resources management: theory and practice.* Berlin: Springer.
PATTON, M. 1997. *Utilisation-focused Evaluation,* Beverly Hills, Sage.
PAWSON, R. & TILLEY, N. 1997. *Realistic evaluation,* Thousand Oaks, Sage.
ROWE, G. & FREWER, L. J. 2004. Evaluating public participation exercises: a research agenda. *Science, Technology & Human Values,* 29**,** 512-556.
SANDERSON, I. 2002. Evaluation, policy learning and evidence-based policy making *Public Administration,* 80**,** 1-22.
SELIN, S. 1999. Developing a typology of sustainable tourism partnerships. *Journal of Sustainable Tourism,* 7**,** 260-273.
STOJANOVIC, T. & BALLINGER, R. 2009. Integrated coastal management: a comparative analysis of four UK initiatives. *Applied Geography,* 29**,** 49-62.

534 STOJANOVIC, T. & BARKER, N. 2008. Improving governance through local coastal
535 partnerships in the UK. *Geographical Journal,* 174**,** 344-360.
536 SULLIVAN, H. & STEWART, M. 2006. Who owns the theory of change? *Evaluation,* 12**,**
537 179 - 199.
538 WADDOCK, S. A. & BANNISTER, B. D. 1991. Correlates of effectiveness and partner
539 satisfaction in social partnerships. *Journal of Organizational change,* 4**,** 64-79.

540

541