

Toward a model of school inspections in a polycentric system

F.J.G. Janssens and M.C.M. Ehren

Abstract

Many education systems are developing towards more lateral structures where schools collaborate in networks to improve and provide (inclusive) education. These structures call for bottom-up models of network evaluation and accountability instead of the current hierarchical arrangements where single schools are evaluated by a central agency. This paper builds on available research about network effectiveness to present evolving models of network evaluation. Network effectiveness can be defined as the achievement of positive network level outcomes that cannot be attained by individual organizational participants acting alone. Models of network evaluation need to take into account the relations between network members, the structure of the network, its processes and its internal mechanism to enforce norms in order to understand the achievement and outcomes of the network and how these may evolve over time. A range of suitable evaluation models are presented in this paper, as well as a tentative school inspection framework which is inspired by these models. The final section will present examples from Inspectorates of Education in Northern Ireland and Scotland who have developed newer inspection models to evaluate the effectiveness of a range of different networks.

Introduction

Over the last three decades, many governments around the globe increasingly recognize the limitations of centralized policy. They acknowledge that hierarchical forms of coordination have distinct drawbacks in allowing schools limited flexibility in responding to external demands. Arvidsson (2003) also points to the information overload of central policy-makers when trying to implement and monitor (new) policy from one central core. As collaborative, partnerships and networks are, according to Gray et al. (2003), are expected to be more effective in creating an education system in which schools can be responsive to their context and provide innovative and affordable services they are being utilized to a greater extent. Networks as the dominant form of organizing and social coordination reflect the idea that one single government (such as in a hierarchical model) does not have all the knowledge required to solve complex, diverse, and dynamic problems, and that no single actor has the overview necessary to employ all the instruments needed to make regulation effective. Governments realize increasingly that they cannot solve complex social problems on their own and turn to networks and partnerships to provide better and less expensive services to citizens, according to Mayne (2003). Examples are from England where the Department of Education has introduced national, local and (subject) specialized leaders of education who support (groups of) schools in specific areas of improvement, has introduced consortia for professional development (Teaching School Alliances), and has established trusts that run chains of schools under a funding agreement with the Secretary of State; or the Netherlands where mainstream and special schools are now working under a new education authority to provide inclusive education to students in their region.

These changes fit theoretical conceptions of 'polycentricism' which signify 'a structural feature of social systems and refer to many centres of decision making that are formally independent of each other' (Ostrom, Tiebout & Warren, 1961, p. 831). 'Polycentric regulatory regimes are those in which the state is not the sole locus of authority, but where state and non-state actors are both regulators and regulated in highly complex and interdependent relations' (Black, 2008, p. 1-2). In these systems the relations between government and

schools are changed to address the insufficient knowledge of government to identify the cause of problems and design effective solutions that are adequately and fully implemented by schools. Changes have included a decentralizing of decision-making and inventing new ways to regulate the self-regulation of schools.

These changes towards a more network-oriented education system have far reaching consequences for the Inspectorates of Education as Ehren (2016) describes. Inspectorates of Education traditionally use a top down model of (single) school evaluation which is not suitable to deal with the dynamics of collaboration of schools within a network. Jenkins et al. (2003) and Gray et al. (2003) for example point out that such top-down systems provide limited insight into the value partners within a network add to services in a particular area over time. These systems emphasize individual agencies' performance targets and budgets and get in the way of them working together and may hinder cross-cutting work. As network outcomes are often the result of collaborative efforts and fragile compromises between partners with different political, social and economic aims who often also have to satisfy and negotiate conflicting stakeholder interests, a top down hierarchical model will have difficulty to find clear and simple evaluation criteria to evaluate aims and objectives of the network (Schwartz, 2003). Honingh and Ehren (2013) and Ehren and Perryman (2015) describe how most Inspectorates of Education predominantly use standardized inspection frameworks to judge quality of single schools, often ignoring the collaborative work of schools with others schools and stakeholders or their contribution to network-level outcomes. In a more polycentric and decentralized system, their centralized and standardized methods are however becoming increasingly obsolete. As Honingh and Ehren (2013) and Ehren and Perryman (2015) suggest, their roles and responsibilities need to change towards more agile and local methods of evaluation.

Such a shift is however no mean feat as the ambiguous nature of networks, differences in perceptions of connectedness, divergence in defining criteria for success, and the difficulty in identifying and attributing measurable outcomes make such network evaluations a challenging task (Dolinski, 2005; Popp et al., 2005, 2013; Provan, Veazie, Staten, & Teufel-Shone, 2005; Rose, 2004). The collaborative and often complex arrangements for decision-making, communication and reporting complicate how organizations can be held to account as questions such as 'who is accountable to whom and what kind of accountability is in play in such arrangements' are difficult to answer? Add to this, networks as dynamic 'moving targets' combined with difficulties identifying and understanding network effectiveness, and one can begin to understand the complexity of network evaluation (Popp et al., 2013). Evaluating a network requires studying how decisions and activities occur in a diffused decision-making model. It also involves recognizing that networks evolve through stages of development.

This paper proposes a range of evaluation models that can capture such decision, activities and stages of development to evaluate the effectiveness of networks. We will provide examples of how such models can be, and are (to some extent) used by Inspectorates of Education in their evaluations of school networks. Such an evaluation and 'polycentric' inspection model essentially starts with an outline of what effective networks in education look like, which will first be presented in the next section (see also Ehren et al., submitted). In the last section of this paper we conclude by describing a range of promising examples of Inspectorates of Education in Ireland and Scotland and discuss the changing role of inspectorates in the governance structure of networks of schools.

Network effectiveness

Defining network effectiveness: multilevel purposes of a network

Unlike organizations, networks create distinctive network effects, like rapid growth and transmission of information. As networks grow and new members provide access to additional connections, the network can diffuse information, ideas, and other resources more and more widely through its links and become more effective.

Network effectiveness may include open communication, strengthened network capacity and production of knowledge to solve problems that are relevant for the entire network and go beyond the remit of each individual organization. For education networks, such effects can for example include addressing low achievement orientation in communities, lack of homework support, or improved service provision and integration of services across the network such as access to specialized education programmes (e.g. for gifted students). Provan and Kenis (2008) emphasize that network effectiveness needs to be defined by looking at the network as a whole and whether it has been able to move forward in addressing the issue on which they came together to work. In order to justify investing in networks, there is a need to measure the overall impact of the network and demonstrate the added value of the network in terms of achieving new outcomes or improving efficiency or effectiveness, instead of looking at improved performance of individual members of the network (Popp et al., 2013). Network effectiveness is not a mere aggregation of the performance of its members but should be understood as outcomes that cannot be reached by each of the individual members, although there is an expectation that individual organizational participants may, and probably should, benefit as well from collaborating in the network.

Network effectiveness can therefore be defined “*as the attainment of positive network level outcomes that could not normally be achieved by individual organizational participants acting independently*” (Provan & Kenis, 2008, p. 230).

These outcomes will be somewhat unique to each network, and to each sector in which a network exists, depending on the purpose of a particular network (Provan et al., 2007). Following Provan and Milward (2001) and Kenis and Provan (2009), networks can be considered successful when they are able to achieve their expected objectives. Gray et al. (2003) categorize network effects as 1) creating synergy where partnership adds value by combining mutually reinforcing interests, 2) leading to transformation, where the partnership objective is to transform different views into an ideological consensus, and 3) enhancing (financial) efficiency when the use of resources is maximized across the partners in the network.

For example, if the main purpose of a network is to improve the efficiency through better coordination of services, reducing both gaps in and duplication of services, then the ultimate outcome of interest will be more coordinated service delivery across the network. If the main purpose of a network of schools is to improve inclusive education, then the quality of joint provision of services to vulnerable students across the network is the outcome of interest (see Janssens & Maassen, 2015).

Recent analyses of effective networks in education indicate that strong networks of teachers and head teachers promote cooperative learning and improvement in, and across schools and enhance effective teaching practices and student achievement (Earl and Katz, 2006; Chapman and Hadfield, 2010; Hargreaves, 2012; Ainscow, 2015).

Isolating network effectiveness from individual member outcomes is however fraught with difficulties as activities and service delivery are often located within each member of the network (schools or youth services) and network-level outcomes are more difficult to distinguish from the contribution of its members. Likewise, improved performance of individual members caused by participation in the network is hard to isolate as there are often multiple contributing factors to member level outcomes, making it difficult to attribute changes to network activities alone (Popp et al., 2013). For example, it is the schools in the network that are providing education, making it difficult for networks to determine what the legitimate outcomes of the network are versus those of the individual schools.

Networks are complex entities that will have an impact at a number of levels within the network. Network evaluations need to take into account these multiple levels and chains of impact in understanding the outcomes of the network. That chain includes the network's impact on its members, the members' impacts on their local environments, and the members' combined impact on their broader environment. Evaluations designed to examine the effectiveness of the network must understand the relationship between these three and be clear about whether they are assessing performance of individual members of the network, or of the network as a whole.

Levels of analysis to consider in the evaluation of network effectiveness were described in some depth in Hill (2002), building on the work of Provan and Milward (2001) who identified three levels of analysis in their framework for evaluating public sector networks: *community*, *network*, and *organization/participant*. Hill (2002) broke this third level down into two levels, the organization and the individual. A brief description of the four levels of analysis, along with examples of outcomes measures for each of these levels is included in Table 1.

Table 1: Levels of analysis in Inter-organizational network evaluation (Popp et al., 2013, p. 70).

Level of analysis	Description	Sample outcomes
Individual	Assessment of the impact that the network has on the individuals who interact in the network on behalf of their respective organizations and on individual clients.	<ul style="list-style-type: none"> - Increased job satisfaction - Increased capacity - Increased client satisfaction with services - Improved client outcomes
Organization	Assessment of the impact that network has on the member organizations, as the success of network members is critical to overall network effectiveness.	<ul style="list-style-type: none"> - Agency/organization survival - Enhanced legitimacy - Resource acquisition - Improvement in referrals
Network	Assessment of the network itself can have a variety of foci, many of which depend on the relative maturity of the network. The strength of relationships across the whole network is always an important focus.	<ul style="list-style-type: none"> - Network membership growth - Relationship strength - Member commitment to network goals
Community	Assessment of the contributions that the network makes to the community it was established to serve	<ul style="list-style-type: none"> - Better integration of services - Less duplication of and fewer gaps in services - Services provided at lower cost to the community - Positive policy change - Improved population-level outcomes

Stakeholders' views in defining network effectiveness

Another fundamental problem in any effort to evaluate networks, according to Provan and Milward (2001, p. 422) might be that external stakeholder groups seldom exist for networks as they do for individual organizations within the network. That is, effectiveness tends to be seen by external groups as depending on what specific service providers either do or do not do, rather than how well services are provided as a result of network activities (Popp et al., 2013). Stakeholders tend to evaluate, reward, or punish individual agencies, regardless of the network's role in enhancing or limiting client outcomes. Despite this possible problem, the task for network organizers is to minimally satisfy the needs and interests of stakeholders at network and individual member levels, while emphasizing the broader needs of the community and the clients the network must serve (Provan & Milward, 2001).

Given the many different stakeholders (e.g., network members, service recipients, funders and decision-makers), with potentially differing or even conflicting ideas about a 'good' outcome, it is important to be able to show the impact of networks in areas that matter to varying groups (Newcombe, 2003; Brandon & Fukunaga, 2014).

Network characteristics and processes contributing to network effectiveness

Regardless of the purpose of a network, however, there are a number of known factors and processes or activities, based on the literature reviewed, that explain or predict the effectiveness of networks (Popp et al., 2013). Using knowledge about effective network characteristics, such as its available resources, governance, leadership, and structure will enhance the quality of a network evaluation. Activities undertaken during the initial formation of the network, as well as during the network's growth will all affect how the network will evolve and be sustained over time. Understanding which processes and ways of network development contribute to positive network-level outcomes helps to inform evaluation models that can capture the effectiveness of the network.

Structure of effective networks

As far as network structure is concerned, following the work of Provan and Milward (2001), networks can be considered successful when they are able to survive in the long term. Network survival allows network clients to access services in a stable way and gives network workers stable jobs, while the network partners can systematically exploit the advantages of the network. West (2010) for example explains that clear structures for collaboration are a key factor in ensuring networks have an impact on student achievement. According to West (2010), there should be clearly defined and commonly understood structures for leadership and decision-making. These structures need to be adapted to other contingencies in the network in order to be effective, according to Provan and Kenis (2008) (see also Ehren & Perryman, 2015).

Typical structural characteristics include the *governance structures* of networks, the *size* of the networks (number of participants) and the *geographical spread* of organisations in the network (Provan & Kenis, 2008). Governance structures can range from centrally directed to shared governance. Centrally directed networks have one lead partner or a coordinating administrative office, whereas networks that share governance collaborate on an informal basis. Another structural dimension is provided by Muijs et al. (2010) who reflect on the extent to which collaboration has been entered into voluntarily or, for at least one partner, under some form of coercion. Muijs et al. (2010) describe a theoretical continuum, where at one end one could find completely voluntary arrangements, whereby two or more schools form a network without any form of incentive. At the other end of the continuum we find

networks in which two or more schools have been compelled to collaborate with one another by the government or the local authority, for example, with one school charged with improving the other.

The structural contingency of networks has a great impact on the effectiveness of the network. West (2010) for example explains that *clear structures* for collaboration are a key factor in ensuring networks have an impact on student achievement. According to West (2010), there should be clearly defined and commonly understood structures for leadership and decision-making. An Ofsted report and survey¹ explains for example how distances between schools, especially in rural areas (such as the East Coast and South West of England), can limit the flexibility in the use of expertise and resources and therefore detract from the potential advantages of working together in a partnership. Schools in these rural areas often also have limited access to support services.

According to Provan and Kenis (2008), these structures need to be adapted to other contingencies in the network to be effective. These authors for example explain how larger networks will struggle to have effective forms of bottom-up shared governance as members will either ignore critical network issues or spend large amounts of time trying to coordinate across 10, 20, or more organizations, particularly when participants are spread out geographically. Larger networks often also face problems with the distribution of trust across the network and with ensuring goal consensus. Such large networks are therefore more effective, according to Provan and Kenis (2008) with brokered forms of network governance, where a separate administrative entity governs the network and its activities. Shared governance is most likely to be an effective form when trust is pervasive throughout the network and provides a strong basis for collaboration among network members. Such collaboration among all members is, according to Provan and Kenis (2008), less essential in more centralized networks where a lead organization coordinates collaboration through dyadic ties with individual members.

It is important to design evaluations that purposely build in ways to assess how the structure of the network supports the collaboration of network members. Popp et al. (2013, p. 68) suggest the following questions as a starting point for the design of such evaluations:

- Does the network have a clear vision and goals that are understood and supported by all members?
- Is the governance structure a good fit for this network?
- Is the network appropriately resourced to do its work?
- Does the leadership style fit with what we know about effective network leadership?
- Are important management tasks being attended to, and is the management focus evolving appropriately over time?
- Is attention being paid to both the management of the network, and management in the network?
- Does the network have both the internal and the external legitimacy it requires?
- Is the network/relationship structure evolving as expected and contributing positively to the work of the network?
- Is there an optimal mix of strong and weak ties among network members?
- Are the linkages targeted and appropriate?
- Is there trust among network members?
- Are power differentials being recognized and addressed as appropriate?
- Are there multiple levels of involvement?

¹ <http://www.ofsted.gov.uk/resources/unseen-children-access-and-achievement-20-years>; <http://www.publications.parliament.uk/pa/cm201314/cmselect/cmeduc/269/269we13.htm>

- Is there a balance of stability and flexibility?

Processes

Recent literature on the evaluation of networks stresses that evaluations also need to capture the processes within a network that contribute to network-level outcomes instead of only looking at structure (Gilchrist, 2006; Popp, et al., 2013). Evaluating 'how' results are achieved may be just as important as looking at 'what' results are achieved (Janssens & Dijkstra, 2012). A focus on processes as well as the outcomes of networks has the potential to make evaluations more fit for purpose.

Research shows that networks rely on trust and empathy, and thrive through the quality and reach of their relationships (Gilchrist, 2006, p. 29). Following Keast et al. (2004), networks can be considered successful when the collaboration between partners works and leads to the creation of a new organizational form that exists by itself, independently from the network partners. In this case, network participants can feel that they are part of a whole and no longer subject to pressure from their former organizations. Such networks have strong interpersonal relationships and social interaction.

Studies on networks of schools show how such networks promote continuous school improvement through the opportunities they provide for information transfer and development of new knowledge between individuals and levels in organizations (Daly et al, 2010; Moolenaar, 2010). West (2010) and a 2005 review study of the Centre for the Use of Research and Evidence in Education (CUREE) for example suggest that the balance of evidence seems to be that collaborative arrangements can impact on students, though not all do. The CUREE review cites 11 studies that have reported changes in teachers' knowledge and skills as a result of network 'interventions', the majority of which "led to clearly identifiable behaviour changes" such as increased involvement of parents in the life of the school and closer links with local communities. Evidence that collaborative arrangements have an impact on student achievement suggests, according to West (2010) that the following factors play a role:

- *Reciprocity*: At the heart of successful collaborations, there needs to be direct benefit to participating stakeholders.
- *Institutional relationships*: Relationships between partner organizations are stronger than relationships between individuals from those organizations.
- *Transparency*: There should be an open and honest articulation of aspirations and expectations and some process to ensure regular review of progress towards these.
- *Continuity and regularity*: consistent membership and regular communication, with clear timelines that are adhered to.
- *Acknowledgement of contributions*: The willingness to acknowledge individual contributions and to share credit should itself be a goal of collaboration.
- *Continual consultation*: New relationships demand the investment of time, energy, and goodwill.
- *Belief in the collaborative process*: Those involved should believe that more will be achieved by working together than working alone, and this perspective should frame interactions.

Evaluations need to generate knowledge about the status of these relationships, so they can be nurtured, repaired and shaped (Gilchrist, 2006). The evaluation of both the structures and processes which contribute to network-level outcomes is key to providing the network with information about its functioning and to allow the network to address membership, governance or structural issues and improve its overall functioning.

Evolution of networks

In addition, given what we know about the evolution of networks, and especially the challenges of attributing outcomes to networks in the early phases, "*evaluating networks appropriately requires some knowledge of the path of evolution and the particular life stage of the network being evaluated*" (Birdsell et al., 2003, p. 30). Indicators need to be developed against which to assess whether the network is being developed as planned, as well as leaving the flexibility for capturing unintended consequences and new directions resulting from the evolution of the network (Birdsell et al., 2003; Aviram, 2003) and changes in the context in which the network is operating.

Researchers have identified potential indicators, many of which are linked to the level of trust in a network, that relate to whether a network is evolving in maturity, such as (Birdsell et al., 2003, p. 33):

- Members being able to discuss money seriously;
- Achieving agreement about key issues (e.g., governing structure, criteria for success);
- Resolving a conflict successfully;
- Members voluntarily subjugating their own interests to those of the collective in the short term;
- Acknowledging that sustainability is about more than funding;
- Referral among members;
- Showing respect for various perspectives; and
- Using the network as a problem solving mechanism.

Additionally, an evaluation of how the network is enforcing and regulating its own norms and quality also provides an idea of its maturity. Aviram (2003) explains how networks can employ four mechanisms to enforce norms or quality. The first information mechanism described by Aviram (2003) includes the collection and dissemination of information on the credibility of (non)members of the network. Such information collection and dissemination can facilitate independent decisions on the feasibility of transaction and interactions within the network. The second mechanism of exclusion would follow from the collection of information to improve the transactions and interactions in the network by depriving members who are degrading the overall performance of the network from temporary or permanent access to the network. The third control mechanism refers to centralized control of transacting facilities and other members' assets, while the fourth 'switching mechanism' ensures that failed transactions between members (e.g. in sharing services or exchanging knowledge) are replaced with alternative, more effective ones. Mature networks have well-functioning mechanisms in place to ensure that each of its members effectively contributes to the performance of the entire network. A more holistic external evaluation of these mechanisms enhances our understanding of the longer term and system level impact of networks (Mandell & Keast, 2007).

Towards evaluation models of networks

As Mayne et al. (2003) explain, the reliance on networks to achieve policy aims yields specific concerns about their accountability and transparency. New network arrangements, such as the ones described above, often lack adequate accountability and transparency in understanding the arrangements between network partners to meet commitments to each other and to the common cause, and to understand whether specific collaborative arrangements are the best way to ensure the expected level of performance and results.

Evaluation of network performance and accountability of networks is important as partnering and collaboration are not without problems. Several authors (Gray et al, 2003) explain how an increased reliance on networks to implement policy aims may result in the fragmentation of delivery structures, self-protective behaviour and inter-organisational politics and struggles. Partners within a network often seek to protect their specific independence and identity, and traditional accountability and governance structures often support them in doing so. According to Mayne and Rieper (2003), collaborative arrangements often lose sight of the public objectives they are serving as there are many different levels of government involved in delivering the service, where each level may have different objectives. The complexity of managing the network and the partnership arrangement may push the public interest aside and create a range of opportunistic behaviours when individual partners destroy part of the cooperative surplus to secure a larger share of it. Aviram (2003) explains how network partners can default on obligations to other network partners when there is lack of complete control over each other's actions, or when large partners in a network degrade their services to small partners in the network who do not have the opportunity to opt out of the network (Aviram, 2003).

Accountability and inspections are therefore needed to provide checks on whether the collaborative mechanisms of networks are working in a cost-effective way to achieve their objectives and do not have unintended side effects and lead to dysfunctional behaviour. Evaluation of networks can bring order to the potential complexity of relationships within the network and assess the value of these relationships in delivering the agreed standards of outcomes and means (see Gray et al, 2003). Such an evaluation should provide insight into the compatibility of collaborative actors and to inform the design as well as the suitability of collaborative structures and procedures. Accountability approaches are required which focus on assessing improvements in the effectiveness and value for money of whole systems, instead of only their constituents parts (Jenkins et al. 2003, p.76). Such approaches can support the development of networks in identifying errors, understand why things occurred and what was learned as a result (Mayne, 2003). Evaluation can, according to Gray et al. (2003), also facilitate knowledge building and inform debates and choices about alternative forms of collaboration or ways to strengthen the partnership work. In such forms of accountability, measurement shifts from being a technique to determine the precise magnitude of things and to prove and judge the level of achieved performance to developing a credible argument to help clarify complex phenomena by gathering relevant information to enhance understanding about what a network is accomplishing.

We know from the evaluation literature that ideally evaluation planning should begin at the same time as the initial planning and design of the network, and evaluation should begin as soon as the network is operational (Mertens & Wilson, 2012). This is critical given the importance of using early process evaluation results to improve ongoing network development. In addition, since the substantive outcomes of interest are as wide ranging as the purposes of the various networks (Birdsell, Matthias, & colleagues, 2003), it may well be important early on to identify and agree on how effectiveness is defined for a particular network, as well as to decide what shorter term outcomes can be identified to help track progress.

A number of authors have suggested evaluation models that can be used to understand the dynamics of the interrelationships in a network; using bottom-up approaches to evaluation which take into account the complex and sometimes vague roles and powers between parties in a network which are crucial in the success of the network (see Arvidsson, 2003; Mayne & Rieper, 2003). Bemelmans-Videc (2003) and Segsworth (2003) for example sug-

gest 'meta-evaluation' and theory-driven evaluations using logic models. Others (e.g. Patton, 2010; Hill, 2002) have suggested developmental evaluation and system dynamic approaches. These will be described briefly below with the purpose of suggesting ways in which they can inform inspection frameworks.

Meta-evaluation

Bemelmans-Videc (2003) suggests that collaborative constructions have complex accountability relationships and therefore require an increased amount of self-evaluation by the partners. This puts greater strain on the external-internal controller relationship and brings 'meta-evaluation' on the accountability agenda where external evaluators (such as inspectors) will increasingly rely on the audit and evaluation of the network and will have an interest in instructing the network on relevant standards and guidelines to follow in its own (required) (self)evaluation. What is needed, according to Bemelmans-Videc (2003), is a form of coordination of external and internal evaluations, and a form of evaluation synthesis in which results on a set of shared evaluation criteria are made comparable and compatible. Evaluation criteria can be designed around network objectives which act as reference points for performance indicators, requiring network partners to be clear about their intentions, standards and created expectations. This will in turn enhance informal control within the network and in anchoring the partnership. External accountability however also needs to ensure that 'first-order activities' are assessed (such as the actual collaboration between partners and achieved network-level outcomes), instead of only checking on the internal control systems of networks.

Theory-driven evaluation

Meta-evaluation requires some standardization of network processes and outcomes to inform evaluation criteria for internal and external evaluations and are less suitable for capturing a variety of different purposes and collaborative arrangements. Theory-driven evaluations allow for a more localized approach in taking the purposes of the object of evaluation, a specification of what must be done to achieve the network's desired goals, the important aspects that may be anticipated, and how these goals and impacts would be generated, as a starting point. The foundations for theory-driven evaluation were laid by Peter Rossi, along with Carol Weiss and Huey-Tsych Chen who explained how programme theories and logic models can be constructed to guide an evaluation (Christie & Alkin, 2013, p. 25; Mertens & Wilson, 2012, p. 62; see also Astbury & Leeuw, 2010).

Logic models depict linear and fixed processes, where inputs lead to particular outputs, and these outputs in turn lead to the development of short, interim and long-term outcomes (Astbury & Leeuw, 2010; Mertens & Wilson, 2012). A logic model (also known as a logical framework, theory of change, or realist matrix) is a tool used to evaluate the effectiveness of a program (McLaughlin & Jordan, 1999; Funnell & Rogers, 2011; Astbury & Leeuw, 2010); they usually include a graphical depiction of the logical relationships between the resources, activities, outputs and outcomes of a program (Mertens & Wilson, 2012, p. 244). While there are many ways in which logic models can be presented, the underlying purpose of constructing a logic model should be to assess the "if-then" relationships between the elements of a program, based on the causal mechanisms for explaining how and why a program works. As Astbury and Leeuw (2010, p. 368) describe, 'mechanisms are underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest'.

Logic models can guide network evaluations in articulating how a network is expected to collaborate and generate network-level outcomes (Funnell & Rogers, 2011). These assumptions can be tested via both qualitative and quantitative methods of data collection.

Mayne (2008, 2011) particularly suggests 'contribution analysis' as an emerging approach in logic modeling, when attributions of how outcomes were caused are difficult to make. Contribution analysis is an approach for assessing causal questions and inferring causality in program evaluations. Various perspectives are sampled to gather different perceptions about the degree of impact an effort has made on observed results. While not perfect, it can offer a general perspective about the influences each members' efforts are having in a given area and to a particular network-level outcome (Mayne, 2008).

Following the outline of structure, processes and outcomes in the previous section, the following logic model can be used to inform an evaluation of network effectiveness (see also Popp et al, 2013) (see Figure 1).

Insert Figure 1 about here

Figure 1: Components of a tentative inspection framework to evaluate networks of schools

System dynamics

Logic models are particularly relevant for capturing linear cause and effect relations, but have limitations in understanding performance and more cyclical/dynamic processes of change at multiple levels and across multiple contexts. As networks are essentially complex phenomena and 'living systems', system dynamics provide useful models to understand the functioning of networks. Tang and Vijay (2001) quote Sterman (2000) who writes: "system dynamics is a perspective and set of conceptual tools that enable us to understand the structure and dynamics of complex systems". "Systems approaches have historically emphasized the need to understand dynamic interrelations between various components. Because the effect of a given input depends on other conditions in the system, emphasis shifts from isolating the causal effect of a single factor to comprehending the functioning of the system as a whole" (Diez Roux, 2011, p.).

Following Provan and Milward (2001) and Hill (2002) the system to evaluate can be described on four levels of analysis: 1) community; 2) network; 3) organization; and 4) the individual (see Table 1). When we apply the latter two levels to a network of schools the organization level deals with the evaluation of the impact that the network has on the member schools and stakeholders. The individual level deals with the impact that the network has on the individuals who interact in the network on behalf of their respective schools and on stakeholders.

System dynamics was originally developed during the mid-1950s by Forrester at the Massachusetts Institute of Technology, but many scholars have contributed to the development of suitable techniques. Luna-Reyes and Andersen (2003) and Geomakers (<http://geomakers.org>) summarize a range of tools that can capture the nonlinear behavior of complex systems over time, such as 'landscape scans', 'systemic action research', 'systems mapping', discourse analysis, grounded theory methodology, ethnographic decision models, and participant-observer research. These tools can be used in an iterative process of building and testing models which explain the performance and development of networks over time.

Network mapping

When a network involves multiple partners working in collaboration or when the development of a network is a goal of the project, network mapping can provide insight into the dynamics and health of these relationships. Network mapping particularly supports an evaluation of the process indicators (e.g. reciprocity and trust), as well as structural features (e.g. centralized coordination of the network). Tracking how ideas are shared and spread and where participants take joint actions can help support developmental processes.

Being able to generate data about a network can inform the development of strategies. Mapping a network can reveal that certain individuals are particularly influential, as sources of expertise or as connectors. It can also outline the strengths or vulnerabilities of the system and can reveal how densely connected a network is or whether there are peripheral connections that could stimulate innovation (Newcombe, 2003). Analysis may suggest strategies for communicating and organizing within the network. Network mapping may also provide an indicator of how different strategies are unfolding. Monitoring a network over time can reveal how the network responds to various interventions.

Mapping a network is a process of identifying connections between people and graphically displaying those connections. This can be done by hand, although increasingly powerful and accessible software is enabling a more comprehensive analysis of networks and their behavior (Provan & Lemaire, 2012; Popp et al., 2013).

Developmental evaluation

Some recent advances in the discipline of evaluation show particular promise in increasing our ability to understand the development, and ultimately the impact, of complex entities such as networks of schools. Most specifically we are referring to recent work by Michael Quinn Patton (2010) on a new approach to evaluation, called *developmental evaluation*.

Developmental evaluation is about helping people to learn to think and act as evaluators with a goal of ensuring that evaluations have a lasting impact (Patton, 2006, 2010; Gamble, 2008). Patton (2010) describes developmental evaluations as learning evaluations, where the aim is to encourage people involved in innovation initiatives to be constantly assessing what is working as intended, what is not, and using what they learn to make necessary adjustments to the initiative. This is critically important in innovative networks, as precisely what activities and approaches are going to work best in a particular context is often unclear (Popp et al., 2013). This makes ongoing evaluation necessary.

Also, as has been discussed, a number of network researchers are suggesting that evaluations that take a traditional approach to output performance measurement on one level are unlikely to be helpful, in that there is a missed opportunity to gather multiple level process information that would be useful in informing the future directions for a network (Popp et al., 2013). This may make developmental evaluation a particularly good fit for networks that have some element of innovation in their vision.

A logic model can be a useful tool for developmental evaluations of networks. Developmental evaluation requires that the model be updated periodically, given changing priorities and new understandings. In a developmental mode, we move from a logic model as a static instrument, to one that we expect to change and evolve over time. One technique is to build the model from scratch more than once over a period of time, through the systematic testing and refinement of the model (Astbury & Leeuw, 2010). In this way the implemented model becomes at the end a model that explains how and why a network works (or fails to work).

Implications for an inspection framework

Ehren et al. (submitted) discuss how inspections need to shift their methodology of evaluation and judgement of school quality and involvement of users to more 'polycentric approaches'. The models they propose include an assessment of the effectiveness of networks (evaluating the structure and processes described in the previous section), as well as validating specific network-level practices and purposes. The evaluation models described in the previous section can be seen as an addition to the inspection methods, judgements and user involvement they describe. As more Inspectorates of Education are developing a menu of types of inspections, the use of meta-analysis, logic models, system dynamics, network mapping and developmental evaluations are useful additions to understand the performance of a more polycentric education system. Here we describe a number of examples of Inspectorates of Education that have incorporated these approaches in their inspections of schools, areas and themes. These examples are tentative as many Inspectorates of Education are constantly developing their methodology. They however provide a valuable overview of ways in which networks can be held accountable for their work, and how evaluations can support their further development.

Ireland: meta-evaluation in West Belfast area-based inspections

Area-based inspection in West Belfast have similar approaches to evaluation of networks of schools and service providers as those described under 'meta-evaluation'. Brown et al.(2015) describe how in West Belfast, clusters of all mainstream post-primary schools, have been set up to plan collaboratively to meet the needs of pupils in an area and for focusing on quality and sharing good practice. These voluntary clusters are called Area Learning Communities (ALC) and allow schools to meet legislative requirements to provide students with access to a minimum number of 24 to 27 courses (depending on key stage) by collaborating with other institutions within the ALC.

These ALC's are inspected by the Inspectorate of Education in Northern Ireland, ETI, during area-based inspections. As Brown et al. (2015) explain, self-evaluations form the starting point of these inspections; the ETI will ask the inspected organisations to provide them with a self-evaluation of the specific topic that will be inspected (e.g. transition of students) and requests that each organisation completes a self-evaluation report on the strengths and areas of improvement in their own organisation prior to the inspection taking place. This will inform the data collection and analysis during the visit.

The process of internal evaluation in the ALCs is supported by district inspectors who are also part of the external inspection team. As Brown et al. explain in Ehren et al. (submitted):

“Additionally, the ETI has district inspectors who are responsible for a number of schools in a geographical area and who are part of an inspection team of individual schools. They also quality assure school self-evaluations through district visits that are outside of the formal inspection programme. Their expertise and oversight of a geographical area provides valuable opportunities to support the ALC's in acting on full area inspections and, as Brown et al. (2015) describe, they have played a valuable role in shaping collective self-evaluations in the ALC and in connecting these self-evaluations to full area inspections.

According to Brown et al. (2015), these evaluations have been highly successful in West Belfast because of the strong and frequent interconnections and the high trust environment in which inspections and self-evaluations are carried out and in which collaborative agreements are made on improvement.”

Scotland: system dynamics in Place-based Scrutiny – East Perthshire

A second example is the use of system dynamics in 'place-based scrutiny' in East Perthshire in Scotland. Education Scotland, responsible for inspection of Scottish schools, was involved in this scrutiny exercise which included a number of Inspectorates overseeing services within that area (e.g. health, education etc.). The 'place-based scrutiny' took place 2015 with a scoping exercise in February and fieldwork during a week in March, and is an interesting example of how a team of external Inspectorates and internal evaluators in the region used a holistic and system dynamics approach to answer the following questions:

1. What is it like to live in this community?
2. How well are services collaborating to improve outcomes for people living there?
3. Is our collective activity addressing/tackling inequalities?

The exercise aimed to identify issues that need to be addressed to improve the lives of the people living in the area.

The External Evaluation Team was made up of Education Scotland (Lead Facilitator), the Care Inspectorate, the Scottish Housing Regulator, Audit Scotland, Healthcare Improvement Scotland, HMICS, the Fire Inspectorate, and the Academic Advisor, while the internal evaluation team involved key stakeholders from the local community, such as the environment and consumer services, Housing & Community Care, Children and Families' Services, Scottish Fire & Rescue Service, Police Scotland, Senior Community Capacity Worker and the National Health Service (NHS).

The scrutiny was coordinated by Audit Scotland and started with a locality profile which included data about specific subgroups and sub-communities and their needs (e.g. poor pensioners, countryside communities, striving families), according to area/street, using local knowledge to identify families and groups of residents to engage with. Analysis of the profile data during a scoping exercise resulted in the external and internal evaluation team forming into five mixed teams, made up of staff and members of the various scrutiny bodies in East Perthshire. The teams organised their fieldwork to map the Strategic Objectives for the area against the different subgroups and sub-communities that were the focus for further enquiry. The teams were also formed to ensure that external evaluators and inspectors were enquiring into areas that were not necessarily their own area of expertise or background. The scoping exercise was supported by researchers from the University of Glasgow (Professor Chris Chapman).

After the scoping exercise, a week of fieldwork was scheduled where the five teams spent time in the community, engaging with residents across all life stages. Collaborative enquiry was used to collate and analyse data and to engage with residents and representatives of the council to build a profile of the area. According to Education Scotland², members of the local community were invited to share their own experiences and to participate and have ownership of the process through their help in devising meaningful questions to jointly reflect on the achievements and successes, issues and challenges of living in this locality, asking local people directly about their experiences and challenges living in this place. The emphasis was on spending time in the locality and engaging with the community, without the use of quality or performance indicators, but using an open format of questioning.

According to Education Scotland, the evaluation team was 'expected to have fun, take risks with the methodology and to be bold enough to make the paradigm shift from always

² This section was informed by an internal working paper 'Place-based Scrutiny – East Perthshire; Methodology and Learning from the Process' (2015), and a follow-up interview with a representative from Education Scotland

being consistent to being valid'. The internal and external evaluation team jointly interviewed residents and spend time in the community, using a relatively open ended framework to organise findings, to identify the questions that needed to be asked and of whom, and to organize the process and allow the team to make sense of large amounts of data. During the fieldwork, the teams kept a log of the process to assist with professional reflection and an evaluation of the efficacy of this approach to place-based scrutiny. The teams came together three times over the course of the week to check in and continue to develop templates and methods to capture the findings.

The overall methodology was flexible and could be customised for use in individual settings (for example focusing on individual themes/clusters of themes, or on single household groups/clusters of groups). It also allowed for flexibility in terms of the timing and scheduling of activities.

Scotland: developmental evaluation in School Improvement Partnership Programme³

Another interesting example from Scotland is the School Improvement Partnership Programme (SIPP) which provides ideas for the use of developmental evaluation by Inspectorates of Education. In March 2013, the Scottish Cabinet Secretary for Education and Lifelong Learning announced the SSIP as a solution focused approach to tackle the steadfast link between socio-economic deprivation and low educational attainment. The aim of the programme is to make explicit links to strategic improvement planning in schools and local authorities.

The School Improvement Partnerships are an approach based on action research and process of collaborative inquiry. Schools have been asked to take the lead in developing projects around a number of key themes (e.g. differences in achievement by gender, improving transition, differences between small and large schools); projects are expected to operate across local authority boundaries (cross-sectoral, multi-disciplinary, partnerships with independent sector), and involve a partnership with local authorities, Education Scotland and other agencies. Each partnership is expected to share and try out effective approaches and to indicate what success will look like, with a strong focus on impact in making a difference to young people's achievement and ultimately life chances. Each of the partnerships will develop a structured, collaborative enquiry approach. The methodology used in each of the action research projects is called 'collaborative enquiry' and follows the following three phases:

Phase 1: Preparing the ground

1. Analysis of context (where are we now?)
2. Agreeing enquiry questions (what are our key concerns?)
3. Agreeing purposes (what would success look like?)

Phase 2: Exploring the evidence:

1. Using the available expertise (how do we exploit internal and external knowledge?)
2. Collecting data (what further evidence do we need?)
3. Making sense of the evidence (what new insights do we have?)

Phase 3: Testing change

1. Deciding on actions to be taken (What changes do we need to make?)
2. Implementing a strategy (How do we lever and embed change?)
3. Monitoring outcomes. (How do we know we have made a difference?)

³ This section was taken from the website of Education Scotland: http://www.educationscotland.gov.uk/Images/SIPPpurposeandrational_e_tcm4-826981.pdf (retrieved February 2016)

Each project is supported by a trio of named individuals from Education Scotland, local authorities and university researchers who each undertake their own enquiry to explore how the partnership project contributes to the overarching programme enquiry.

The overall programme is coordinated by Education Scotland who lead the development of this programme and manage its implementation. Education Scotland sits on the panel (with ADES and Glasgow University) to select partnership projects that will be funded; partnerships may be visited by the Education Scotland Area Lead Officer. Education Scotland's role in the Programme is also about brokering national partnering and links across authorities and university researchers.

Education Scotland collaborates with researchers and will identify key challenges in and monitor developments of the programme. Its role is also to provide assistance in collating statistical information about the schools and partnerships to inform their decision-making; supporting a database and communication system to support the sharing of knowledge and resources within and between partnerships, and bringing partnerships together to share experiences and practice at appropriate points within the programme.

According to Education Scotland: 'This ambitious programme seeks to harness the professional creativity and innovation that exists within the Scottish education community. The programme provides exciting opportunities to rethink roles and relationships within the system and generate and share new practice.' (see: http://www.educationscotland.gov.uk/Images/SIPPpurposeandrational_e_tcm4-826981.pdf)

Conclusion and discussion

This paper described some emerging and promising approaches to evaluation, which may be used by Inspectorates of Education in their evaluation of networks, geographical areas and themes. The models we described include meta-evaluations, theory-driven logic models, systems dynamics, network mapping and developmental evaluation; they allow Inspectorates of Education to understand the effectiveness of networks of schools and other service providers, and to understand the performance of education systems that have many centers of decision-making that are formally independent of each other. We argued that such evaluations are a mechanism by which practice based knowledge can be co-created by network practitioners, researchers and school inspectors, and then shared to improve the overall functioning of a polycentric education system.

The examples we described of Inspectorates of Education in Northern Ireland and Scotland using these evaluation models highlight how the use of such models implies a dramatic shift in the position of Inspectorates of Education, schools and their stakeholders. They now become equal partners in a more interactive and ongoing evaluation of education quality, implying a different set of consequences to motivate improvement. The examples from Northern Ireland and Scotland indicate how Inspectorates of Education can develop a set of intelligent and more flexible evaluation models and intervention strategies that would improve the performance of the entire education system by purposefully providing relevant actors with feedback, to improve relations in the network and increase openness to external stakeholders and information.

An important comment to make is that a development towards more lateral polycentric structures of school governance and evaluation needs to be informed by research on effective networks and involvement of stakeholders in decision making processes. Many authors (Kenis & Provan, 2009; Popp et al., 2013) have proposed various conceptualizations

and measures of network effectiveness according to different criteria. Some authors conceptualize and measure network performance based on the stakeholder perspective. Some focus on network performance from the partner organizations' perspective; some take into account the entire network and evaluated the benefits for people working within the network; others look at matters from the community perspective, recognizing the contribution to the pool of clients served by the network. Other studies stress the conceptualization and measures of network effectiveness based on the intrinsic characteristics of the network itself. Some authors evaluate network performance by looking at the network's ability to survive in the long term. According to others, networks can be considered successful when collaboration between partners works. Yet others focus on the ability of the network to achieve its expected objective.

The resulting landscape is still perplexing and implies the need to understand the intrinsic meaning of network performance and its measurement criteria. The approaches we suggested take this multidimensional nature of networks into account and the fact that different networks will have different purposes, structures and processes of change. Expanding the menu of types of inspections allows for greater flexibility in implementing evaluations that are fit for purpose and inform system-wide improvement in an ever changing education landscape.

References

- Ainscow, M. (2015). *Towards self-improving school systems: Lessons from a city challenge*. London: Routledge.
- Arvidsson, G. (2003). Collaboration by Contract and Pooling Resources: The Implications for Evaluation. (p.131-155). In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- Astbury, B. & Leeuw, F.L. (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, 31(3) 363-381.
- Aviram, A. (2003). Regulation by networks. *Brigham young university law review*, (4), pp. 1179-1238. Retrieved from <http://digitalcommons.law.byu.edu/cgi/viewcontent.cgi?article=2181&context=lawreview>.
- Bemelmans-Videc, M. (2003). Auditing and Evaluating Collaborative Government: the Role of Supreme Audit Institutions. In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- Birdsell, J., Matthias, S., & colleagues (2003). *Networks and their role in enhancing research impact in Alberta*. Alberta, Canada: Alberta Heritage Foundation for Medical Research. Retrieved from http://evaluationcanada.ca/distribution/200312_birdsell_judy_mathhias_sharon.pdf
- Black, J. (2008). *Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes*. LSE Law, Society and Economy Working Papers 2/2008 London School of Economics and Political Science. Law Department. Retrieved from www.lse.ac.uk/collections/law/wps/wps.htm and the Social Sciences Research Network electronic library at: <http://ssrn.com/abstract=1091783.0>
- Brandon, P. R., & Fukunaga, L. (2014) The State of the Empirical Research Literature on Stakeholder Involvement in Program Evaluation, *American Journal of Evaluation*, 35 (1),26-44.

- Brown, M., McNamara, G. and O'Hara, J. (2015). *School Inspection in a Polycentric Context: The Case of Northern Ireland*. Dublin: (EQI) Centre for Evaluation, Quality and Inspection.
- Christie, A.C. & Alkin, M.C. (2013). An Evaluation Theory Tree. In: Alkin, M.C., ed. *Evaluation Roots: a wider perspective of theorists' views and influences*. Thousand Oaks, Cal: Sage
- Centre for the Use of Research and Evidence in Education (CUREE) (2005). *Systematic research review: the impact of networks on pupils, practitioners, organisations and the communities they serve*. Nottingham: National College. Retrieved from http://www.curee.co.uk/files/publication/1219841354/Impact_of_Networks_Leaflet.pdf
- Chapman, C., & Hadfield, M. (2010). Supporting the middle tier to engage with school-based networks: Change strategies for influencing and cohering. *Journal of Educational Change*, 11, 221-240. doi: 10.1007/s108333-009-9125-y.
- Daly, A. J., Moolenaar, N., Bolivar, J., & Burke, P. (2010). Relationships in reform: The role of teachers' social networks. *Journal of Educational Administration*, 48(3), 20-49.
- Diez Roux, A. V. (2011). Complex systems thinking and current impasses in health disparities research. *American Journal of Public Health*, 101(9), 1627–1634. <http://doi.org/10.2105/AJPH.2011.300149>. <http://europepmc.org/articles/pmc3154209>
- Dolinski, C. (2005). *Southern Alberta Child and Youth Health Network – Evaluation report*. Report prepared for the Southern Alberta Child and Youth Health Network (SACYHN). Calgary, AB: SACYHN.
- Earl, L. and Katz, S. (2006). *How Networked Learning Communities Work*. Victoria: Centre for Strategic Education Seminar Series Paper No. 155.
- Ehren, M. & Perryman, J. (2015). *School inspections in a polycentric context; Ofsted and a self-improving school system*. London: UCL Institute of Education (unpublished manuscript).
- Ehren, M.C.M. (Ed.) (2016). *Methods and Modalities of Effective School Inspections*. Dordrecht: Springer.
- Ehren, M.C.M., Janssens, F.J.G., Brown, M., McNamara, G., O'Hara, J. and Shevlin, P. (submitted). *Emerging models of school inspections: Shifting roles and responsibilities of Inspectorates of Education in a polycentric system*.
- Funnell, S. & Rogers, P. (2011). *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco, CA: John Wiley & Sons.
- Gamble, J.A.A. (2008). *A Developmental Evaluation Primer*. Montreal, Quebec: The J.W. McConnell Family Foundation. <http://cpo.wisc.edu/wp-content/uploads/2014/07/A-Developmental-Evaluation-Primer-EN.pdf>
- Gilchrist, A. (2006). Maintaining relationships is critical in network's success. *Healthcare Papers*, 7(2): 28- 31.
- Gray, A., Jenkins, B. & Leeuw, F. (2003). Collaborative Government and Evaluation: The Implications of a New Policy Instrument (p. 1-29). In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- Hargreaves, D. (2012), *A self-improving school system: towards maturity*, Wolfson College, Cambridge, Nottingham: NCSL.
- Hill, C. (2002). *Network literature review: Conceptualizing and evaluating networks*. Calgary, AB: Southern Alberta Child and Youth Health Network. Retrieved from http://sls.royalroads.ca/sites/default/files/files/network_literature_review_conceptualizing_and_evaluating_networks_2002.pdf

- Honingh, M. and Ehren, M.C.M. (2013). *Onderwijstoezicht in een polycentrisch sturingsmodel; Dilemma's bij het vaststellen en verbeteren van de onderwijskwaliteit*. Bestuurskunde, (4).
- Janssens, F.J.G. & Dijkstra, A.B. (2013). *Positionering van de evaluatiemethodologie van het onderwijstoezicht: Uitgangspunten voor de inrichting van het toezicht op de kwaliteit en deugdelijkheid van het onderwijs [Positioning of the evaluation methodology of school inspections]*. Amsterdam/ Enschede: Universiteit van Amsterdam/Universiteit Twente.
- Janssens, F. & Maassen, N. (2015). *School inspections in a polycentric context: The Dutch Inspectorate of Education*. Enschede: University of Twente (unpublished manuscript).
- Jenkins, B., Leeuw, F. & Van Thiel, S. (2003). Quangos, Evaluation and Accountability in Collaborative Government. (p.53-83). In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- Keast, R., Mandell, M.P., Brown, K. & Woolcock, G. (2004), Network Structures: Working Differently and Changing Expectations, *Public administration review*, 64, 3, 363-371.
- Kenis, P. & Provan, K.G. (2009). Towards An Exogenous Theory Of Public Network Performance, *Public Administration*, 87, 3, 440-456.
- Luna-Reyes, L. F., & Andersen, D. L. (2003). Collecting and analyzing qualitative data for system dynamics: methods and models. *System Dynamics Review*, 19(4), 271-296.
- Mandell, M., & Keast, R. (2007). Evaluating network arrangements: Toward revised performance measures. *Public Performance & Management Review*, 30(4), 574-597.
- Mayne, J., Wileman, T., & Leeuw, F. (2003). Networks and partnering arrangements: New challenges for evaluation and auditing (p. 29-51). In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- Mayne, J. (2008). *Contribution Analysis: An approach to exploring cause and effect*, ILAC methodological brief, available at http://www.cgiar-ilac.org/files/ILAC_Brief16_Contribution_Analysis_0.pdf
- Mayne, J. (2011) Contribution Analysis: Addressing Cause and Effect. In: K. Forss, M. Marra and R. Schwartz (Eds.), *Evaluating the Complex*. Transaction Publishers; Piscataway, New Jersey.
- Mayne, J. and Rieper, O. (2003). Collaborating for Public Service Quality: the Implications for Evaluation. In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers.
- McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65-72.
- Mertens, D.M. & Wilson, A.T. (2012). *Program Evaluation Theory and Practice: a comprehensive guide*. New York: The Guilford Press.
- Muijs, D., West, M., & Ainscow, M. (2010). Why network? Theoretical perspectives on networking. *School Effectiveness and School Improvement*, 21(1), 5-26.
- Moolenaar, N. M. (2010). *Ties with potential: Nature, antecedents, and consequences of social networks in school teams*. University of Amsterdam: Dissertation.
- Newcombe, R. (2003). From client to project stakeholders: a stakeholder mapping approach. *Construction Management and Economics*, 21(8), 841-848.
- Ostrom, V., Tiebout, C.M. and Warren, R. (1961). The Organization of Government in Metropolitan Areas: A Theoretical Inquiry. *American Political Science Review*, 55(4) (December): 831-842.

- Patton, M. Q. (2006). Evaluation for the way we work. *Non-Profit Quarterly*, 13(1), 28-33.
- Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York: The Guilford Press.
- Popp, J. K., L'Heureux, L. N., Dolinski, C. M., Adair, C. E., Tough, S. C., Casebeer, A. L., Douglas-England, K. L., & Morrison, C. C. (2005). How do you evaluate a network? A Canadian child and youth health network experience. *Canadian Journal of Program Evaluation*, 20(3), 123-150.
- Popp, J., MacKean, G., Casebeer, A., Milward, H. B., & Lindstrom, R. (2013). *Inter-organizational networks: A critical review of the literature to inform practice*. Retrieved from http://www.businessofgovernment.org/sites/default/files/Inter-Organizational%20Networks_0.pdf
- Provan, K. G., & Milward, H. B. (2001). Do networks really work? A framework for evaluating public-sector organizational networks. *Public Administration Review*, 61(4), 414-423.
- Provan, K. G., Veazie, M. A., Staten, L. K., & Teufel-Shone, N. I. (2005). The use of network analysis to strengthen community partnerships. *Public Administration Review*, 65(5), 603-612.
- Provan, K. G., Fish, A., & Sydow, J. (2007). Interorganizational networks at the network level: A review of the empirical literature on whole networks. *Journal of Management*, 33(3), 479-516.
- Provan, K., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory*, 18(2), 229-252.
- Provan, K. G., & Lemaire, R. H. (2012). Core concepts and key ideas for understanding public sector organizational networks: Using research to inform scholarship and practice. *Public Administration Review*, 72(5), 638-648.
- Rose, T. R. (2004). *Research project report: Evolution, development, and maintenance of the Southern Alberta Child and Youth Health Network*. Centre for Health and Policy Studies at the University of Calgary.
- Schwartz, R. (2003). The Politics of Evaluating Government Collaboration with the Third Sector. In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers
- Segsworth, B. (2003). Evaluation, Accountability, and Collaboration. In: A. Gray, B. Jenkins, F. Leeuw, J. Mayne (Eds). *Collaboration in Public Services: The Challenge for Evaluation*. New Brunswick: Transaction Publishers
- Tang, V. and Vijay, S. (2001). *System Dynamics; Origins, development, and future prospects of a method*.
 ESD.83, Fall 2001. Research Seminar in Engineering Systems.
web.mit.edu/esd.83/www/notebook/System%20Dynamics%20final.doc (retrieved February 2016).
- West, M. (2010). School-to-school cooperation as a strategy for improving student outcomes in challenging contexts. *School effectiveness and school improvement*, 21(1), 93-112.

Network processes and structures

Network level outcome

Outcomes at other levels

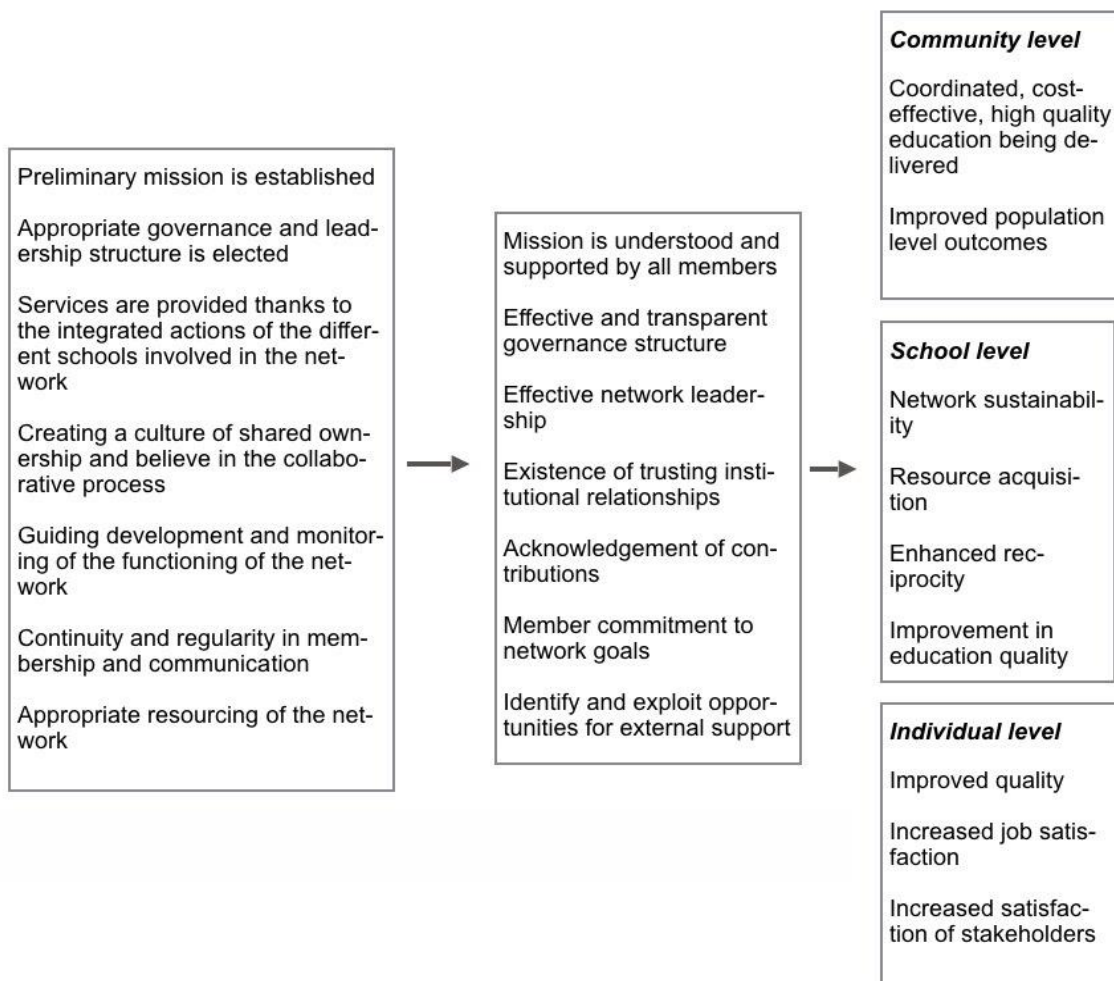


Figure 1: Components of a tentative inspection framework to evaluate networks of schools