

Tilburg University

Towards an exogenous theory of public network performance

Kenis, P.N.; Provan, K.G.

Published in:
Public Administration

DOI:
[10.1111/j.1467-9299.2009.01775.x](https://doi.org/10.1111/j.1467-9299.2009.01775.x)

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Kenis, P. N., & Provan, K. G. (2009). Towards an exogenous theory of public network performance. *Public Administration*, 87(3), 440-456. <https://doi.org/10.1111/j.1467-9299.2009.01775.x>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

TOWARDS AN EXOGENOUS THEORY OF PUBLIC NETWORK PERFORMANCE

PATRICK KENIS AND KEITH G. PROVAN

This article offers insights into the complexity of assessing the performance of public networks. We have identified three so-called exogenous factors: form of the network, type of inception – whether the network was initially formed as voluntary or mandated – and developmental stage of the network. We argue that where a network stands on each of these factors will determine the appropriateness of specific criteria for assessing the performance of the network.

INTRODUCTION

Networks, consisting of three or more organizations that consciously agree to coordinate and collaborate with one another, are used to deliver services, address problems and opportunities, transmit information, innovate, and acquire needed resources. Consistent with the importance of networks to organizations in business, public, and nonprofit sectors, the study of networks has exploded in recent years. In the public sector, especially, networks are increasingly recognized as a viable mechanism for providing services and implementing policies and as an alternative to traditional hierarchical governance. The network form of governance has received considerable attention in the literature (Agranoff 1991, 2003; Baker *et al.* 1994; Jennings and Ewalt 1998; Israel *et al.* 1998; Chisholm 1998; Provan and Milward 2001, Goldsmith and Eggers 2004; Provan *et al.* 2005), especially as a mechanism for encouraging collaboration, building community capacity, and for enhancing organizational and client-level outcomes.

The presumed performance benefits of networks have attracted increased attention by policy-makers and network practitioners alike. After a period of network euphoria, during which the presence of networks was considered as something positive *per se*, questions have arisen as to whether and under what conditions networks are actually performing at a level that justifies the costs of collaboration, which can be substantial. Although many studies take the presumed benefits of network governance as their starting point (O'Toole 1997), the study of whether and under what specific circumstances networks are actually effective has received much less attention. This is especially the case when the network as a whole is taken as the unit of analysis (for a review, see Provan *et al.* 2007), as opposed to the actors participating in the network or to dyads of organizations (for studies of networks where the effectiveness of organizations and/or dyads is examined, see, for example, Wiewel and Hunter 1985; Miner *et al.* 1990; Baum and Oliver 1991; Stuart *et al.* 1999). Burt (1992, 1997), for example, used the aggregated rewards (namely, managers' promotion) of network participation as the indicator of performance, which means that the focus was on individual performance as a consequence of network involvement, as opposed to the performance of the network itself.

We are interested here in studying 'network level performance' or what Provan and Milward (2001) articulate as the 'joint production problem', where multiple agencies

Patrick Kenis is in TiasNimbas Business School and the Department of Organization Studies, Tilburg University. Keith G. Provan is in the Eller College of Management, University of Arizona and the Department of Organization Studies, Tilburg University.

are responsible for one or more components of a single service. This unit of analysis is especially appropriate for public sector networks. The critical issue, both for clients of services and system-level planners and funders, is often the performance of the entire network, not whether some organizations that are part of the network do a better job than others in providing a particular component of service (Provan and Milward 2001). Obviously, a network may be ineffective if individual organizations within the network do a poor job. But in addition, even though network organizations may provide excellent services on their own, overall network outcomes may be low (see Provan and Milward 1995). Thus, performance of the entire network needs to be treated as a separate issue – an idea that is central to the present paper. When defined as an entire system of inter-related service providers, which is the relevant unit of analysis for many public sector issues, networks become extremely complex entities that require separate theorizing in ways that go well beyond how networks have been traditionally discussed in the literature on organization theory and strategic management that addresses business networks.

Effectiveness has been one of the most controversial topics in the organization literature. The discussion thus far has been so inconclusive that some have even suggested that scholars stop addressing the topic altogether (Hannan and Freeman 1977). Our position regarding networks is that the study of effectiveness should not be dropped for the simple reason that managers, policy-makers, politicians, citizens and others have continued to question the effectiveness of specific interventions and the organizational mechanisms that deliver needed services. Rather than giving up on the topic of the effectiveness of networks, we consider it a challenge to find new avenues to contribute to our knowledge. That is what we will do in the present paper. We do not attempt a comprehensive review of the limited literature on network effectiveness. Rather, we examine a number of deficiencies in the literature addressing the topic. We then introduce one approach that we believe moves the discussion forward in a productive way.

DEFICIENCIES IN THE STUDY OF NETWORK PERFORMANCE

A first deficiency in the literature is that network performance is seldom studied as a dependent variable. Instead, most of the research has focused on explaining characteristics of networks, or, if performance is considered, the focus is typically to explain policy outcomes and service effectiveness at the level of the organization. Meier and O'Toole (2001), for example, used programme output as a measure of education network performance to assess the effectiveness of schools and school districts. This is also evident from the articles published in the Special Research Forum on Building Effective Networks, appearing in *The Academy of Management Journal* (D.J. Brass *et al.* 2004). All articles published used networks as the independent variable but, contrary to what the title of the Special Research Forum might suggest, not one study actually considered network performance as the dependent variable. In order to better understand why some networks perform well and others do not, we need studies where network performance is the dependent variable.

Another deficiency is that in many studies the concept of network performance is poorly specified. In 1967, Yuchtman and Seashore commented that 'Most of the study [of effectiveness] has been devoted to the *conditions* under which organizations are more or less effective' (1967, p. 891; emphasis in original) and that '. . . little attention, however, has been given to the concept of effectiveness itself. The latter remains conceptually a vague construct . . .' (1967, p. 891). Even today, scholars who address the issue of network

effectiveness do not properly define or operationalize what type of effectiveness (that is, which criteria) they have in mind, focusing instead on different conditions or 'success factors' contributing to effectiveness. This is problematic because it is likely that while the conditions studied may contribute to one type of performance, they do not necessarily contribute to another type.

Whether focusing on organizational or network performance, not specifying to which type of outcome the conditions or success factors contribute produces problematic recommendations. For example, in a study on networks of mental health delivery, Provan and Milward (1995) persuasively demonstrated that the four networks studied were differentially effective depending on whether they were assessed by the clients themselves, the clients' case managers, or the clients' families. Consequently, any study of the performance of organizations or networks should make fully explicit which performance criteria are being addressed. It makes a difference, for instance, whether a health and human services network is assessed on its productivity or on the quality of care it provides. It has been shown again and again that most common criteria of performance are statistically uncorrelated (see Meyer 2002, Ch. 2). This implies that the same object of evaluation can receive a positive assessment on one criterion while at the same time receiving a negative assessment on another criterion.

Another deficiency often observed in the literature is that criteria for effectiveness are reduced to 'measurements'. In his book *Rethinking Performance Measurement*, Marshall W. Meyer writes in the Preface that he owes substantial debt to Robert Merton for persistently asking him whether he was confusing performance measures with performance (or in other words, whether he had fallen in the trap of 'operationalism') (Meyer 2002). This implies that scores on measurement instruments are used to evaluate performance without being clear about exactly what these measures actually operationalize (that is, measures of which specific criteria). It should be noted here that we distinguish between 'criteria' and 'indicators or measurements'. A criterion is a standard on which to base a judgement; 'indicators', on the other hand, serve to measure what a criterion is in operational terms and how to measure it. Consequently, although these concepts are used interchangeably in the literature, they actually have completely different meanings. It is both easy and tempting to measure something (for example, number of clients served) and label it 'performance', and then to rate and rank organizations on the measure and publicize the rankings so that the measure becomes synonymous with performance in people's minds. It is far more difficult to first answer the question, 'what is performance?' and, then, to answer the question, 'how should performance be measured?' For example, it is common in the literature on network performance to use proxies for network performance such as the development of trust, ways of communicating among members, development of commitment, and so on (Mandell 1994; Kickert *et al.* 1997; Koppenjan and Klijn 2004). In these examples, it is not the tasks accomplished in a network that are central in evaluating network performance, but rather, new ways of behaving and managing. Although we recognize the importance of use of proxies in evaluation in this type of literature, it is often not clear what these are proxies of.

What the above deficiencies have in common is that they either avoid, do not take seriously, or are rather careless when it comes to specifying performance criteria in the context of addressing the issue of performance. In some respects this is surprising given the fact that criteria are the key factors in assessing the performance of anything. It is also not surprising since it indicates that specifying criteria is such a difficult issue that scholars would rather avoid it. The choice of the criterion on which something is assessed

is likely to influence reported performance (or in other words, performance is a function of the assessment criterion used). Consequently, in what follows, rather than avoiding it, we propose to bring back into the discussion the issue of 'criteria' for the assessment of performance.

Consistent with these arguments, the key question we ask is which criteria should then be used when assessing or studying the performance of a network? Although criteria such as 'efficiency', 'effectiveness', or 'goal attainment' are mentioned most commonly, we believe that in principle, any criterion can be used to assess the performance of a network (efficiency, goal attainment, equity, quality, productivity, level of conflict, growth, survival, profit, stability, resilience, learning, and so on). A criterion is a 'norm,' and since it is a norm, this implies that any decision about the criterion used in assessing the performance of a network is a normative decision. According to Herbert Simon, assessment criteria are elements of value and not elements of fact (Simon 1976). Elements of value are implicit or explicit imperatives about the preferred state of a system that cannot be completely derived from the elements of fact; nor can they be proven empirically (Simon 1976). In contrast, elements of fact are statements about the observable world and its operation. These elements can be tested empirically to determine whether they are true or false. For example, we can empirically observe whether a company earns a 'profit' and we can test whether certain conditions are a causal explanation for the degree of profitability. The desirability of earning a profit cannot, however, be determined empirically. In a similar fashion, there is no scientific way to judge whether one criterion is 'better' than another in assessing the performance of a network.

The above implies that we favour the so-called multi-constituency approach to effectiveness. The multi-constituency approach argues that it is arbitrary to label one criterion *a priori* as the correct one because each presents a valid point of view (Connolly *et al.* 1980). Within this perspective performance is not treated as a single statement 'but as a set of several (or perhaps many) statements, each reflecting the evaluative criteria applied by various constituencies involved to a greater or lesser degree with the organization' (Connolly *et al.* 1980, p. 213). An overall judgment of performance is viewed as being neither possible nor desirable because the approach does not make any assumptions concerning the relative primacy of one's constituency's judgements over those of any other constituency. In this view, the role of the assessor is one of collecting information concerning the constituent's judgements of performance and reporting it to the individuals or agencies asking for the assessment. For instance, those interested in whether trains should run on time might get one assessment that is quite different from those interested in whether those same trains are comfortable.

Organizational scholars often feel uncomfortable with this approach because of its relativistic nature. Thus, it is impossible to have clear scientific directions regarding which criterion to use to assess either an organization or a network. We have identified several strategies used by researchers for dealing with this situation. A first strategy is to try to argue that one criterion is superior to another. For example, it may be argued that, ultimately, goal attainment is the only criterion that really matters (assuming it is possible to determine which of several goals is most critical). As argued before, this is a normative choice and in some cases, as is often the case in the initial stages of network development, goals are not yet present or made explicit. The same argument could be made for any assessment criterion which is considered by the researcher to be superior (for example, the goal model, the system-resource model, the internal process model, and the competing values model (for further discussion on these, see Boyne 2003).

A second strategy is to take a multidimensional approach towards effectiveness, such as the balanced scorecard approach (for a critical view, see Meyer 2002). With a multidimensional approach, there is recognition that different criteria might be at stake, each of which is important, but that they are not necessarily related. Here, however, a choice must be made as to which criteria should be part of the multidimensional space and which should not. Moreover, Jacobs and colleagues (2006) have demonstrated that, although composite measures may have a number of advantages, especially in terms of communicating to the public about the performance of public sector entities (in the form of comparative ratings and rankings of performance), the construction of a composite measure is not straightforward methodologically and leaves much open to misinterpretation as well as potential manipulation.

A third possible strategy is for the researcher to make a clear choice of one (or several) criteria in favour of others. This is a completely legitimate strategy as long as the researcher makes clear that they are conscious of and explicit about the fact that the choice has a normative character. An example of this approach is the study by Provan and Milward of networks serving adults with serious mental illness in which they clearly stated that: 'For this research, we felt it was critical to tie effectiveness measures to enhanced client well-being' (Provan and Milward 1995, p. 8).

In this paper we propose a fourth strategy, based on two assumptions. The first is that (based on the relativistic multi-constituency model) any criterion is, from a normative point of view, as legitimate as any other to assess a network. The second assumption is, however, that not every criterion is equally appropriate or reasonable for evaluating a network. The fact that a network scores low on a certain criterion may be related to the fact that that particular network, based on its structure, mission, and so on, cannot score highly on that criterion. The fact that a specific network performs well on service quality but poorly on financial performance may, for instance, be related to reasons that are intrinsic to that type of network; for instance, the type of clients served. Equally, the fact that a very young network cannot score highly on goal attainment is probably not related to the bad performance of that specific network but, instead, to reasons that are intrinsic to the network itself at that particular stage in its evolution. The challenge is thus to identify factors that prevent networks from performing well on certain criteria. The conclusion that a specific criterion is inappropriate should, thus, not be an element of value but an element of fact.

TOWARDS AN EXOGENEOUS THEORY OF NETWORK PERFORMANCE

An exogenous theory of network performance claims that the performance of a network is a function of the external criteria used to assess the network, and that network participants and managers may have little control over these criteria. The same network can score high on a certain criterion but at the same time low on another criterion. For example, a mental health delivery network can score high on client satisfaction but low on the satisfaction of the professionals working in the network. Consequently, poor performance of a network can be the result of two different types of causes. First, performance can be weak because the network performs poorly on a criterion on which it theoretically could and should perform well. There are many possible reasons for this poor performance, including poor network management, unsecured or poorly deployed resources, erroneous task integration, poor cooperation between the actors in the networks, and so on (see Mandell and Keast 2006). These are all issues over which network management should be able

to exert control. Generally speaking, the literature that addresses network performance discusses performance from this – what we call endogenous – perspective.

A different type of reason for a network to perform poorly is related to the fact that the network may be assessed on the basis of inappropriate or unreasonable criteria. The idea is that if a network is assessed on an inappropriate criterion, then it is most likely that the network will score low on that criterion, since even under the best endogenous conditions it will still under-perform. For example, we know from research that the performance of networks of health and social service agencies regarding improved client outcomes for people with serious illness may not occur for at least three to five years (Annie E. Casey Foundation 1995) after initial network formation. Consequently, it would not be appropriate to assess a two-year-old network of this kind on whether it has achieved its client service goals. Conversely, a network that is established to accomplish a specific, time-dependent outcome, should be assessed based on its accomplishment of that outcome. This is typically the case in networks that could be considered to be temporary organizations, like those set up to complete a construction project or to make a film. More generally speaking, in order to evaluate the performance of networks, the time needed to comply with the specific task should be taken into consideration. The opposite reasoning, of course, does not hold. It is not because a network scores low on a certain criterion that the criterion is inappropriate or unreasonable by definition. The network might perform poorly for endogenous reasons. Consequently, the crucial question to answer is when a criterion is appropriate and reasonable and when it is not. This is our focus.

To answer the above question, we first have to define what exactly is meant by appropriate criteria; secondly, we have to point out what *determines* whether it is appropriate or inappropriate to assess a given network on a certain criterion. Although often implicitly taken into account in assessing social systems – and also occasionally mentioned in the evaluation literature (see, for example, Weiss 1998) – we do not know of any systematic treatment of the concept of appropriateness of criteria and its consequences. We propose to define the degree of appropriateness of a criterion as the degree to which the object being assessed (in our case a network) can instrumentally manage how it scores on that criterion. Only when a network has the theoretical capacity to actually influence the score on a criterion is it appropriate or reasonable to assess the functioning of the network on that criterion.

There are different reasons why a network may not be able to instrumentally manage how it scores on a criterion (for a similar discussion, see Gouldner 1957, pp. 96–8). It might be related to the fact that the network has no control over the conditions contributing to the score on a certain criterion (for example, the existing level of trust in a sector or society or the costs involved in changing certain conditions). It might also be related to the values of a network or to what is acceptable, or legitimate, for a particular network. For instance, it would be inappropriate to assess a network in ways that would violate the values of the network (for example, assessing a network on a criterion of efficiency when one of the prime values of the network is consensus decision making). Since it would be extremely difficult to generalize about network values, given the fact that each particular network can develop its own idiosyncratic value considerations, we will not consider this in what follows. What we can do, however, is to identify those network characteristics that cannot be instrumentally managed by the network itself (what we call exogenous factors) but nevertheless influence the type of performance standards the network can accomplish. This implies that no criterion is inappropriate by definition but that the appropriateness of a criterion depends on the characteristics of the network being assessed.

On the basis of the limited literature available on the issue addressed here, and based on our own research in the field of network performance, we have identified three exogenous performance factors: the form of the network, whether the network is mandatory or voluntary, and the developmental stage of the network. We make no claim that these are the only factors that are relevant. Rather, we argue that based on the network literature and on our understanding of networks and network governance, these particular factors are relevant when it comes to understanding why some networks might perform better than others and, thus, that these factors should be taken into consideration by researchers and public policy officials when attempting to assess network performance.

THE FORM OF THE NETWORK

Networks come in different forms. They differ in terms of types of actors involved, the boundaries of the network, and the presence and absence of different types of links. It has been demonstrated that the structural form of a network has consequences for what the network can actually achieve (Baker and Faulkner 1993; Provan and Milward 1995; Cross *et al.* 2002; Burt 2005). Many networks simply lack the functionality to produce certain types of outcomes. This is a matter of design. This does not imply that certain networks are inherently dysfunctional but, rather, that they are designed and organized in ways that are best suited to achieving some types of outcomes, but not others. It is certainly true that network forms can be changed. However, this can be a difficult and time-consuming process, thereby limiting what each network form can reasonably accomplish.

The impact of network design on performance can be amply demonstrated by presenting a typology of network governance forms the present authors recently introduced (see Provan and Kenis 2008). Based on a review of the literature on networks, coupled with our own extensive observations from research and consulting, we identified three basic network governance models: shared or participant governance, lead organization governed, and network administrative organization governed. Each of these models differs regarding their structure. None of these structures turns out to be universally superior. Rather, we argue here that each form has its own particular functionality or, in other words, each differs in what it can do well.

Shared governance form

The simplest form of a network is one that has shared or participant governance. Networks having a shared governance form consist of multiple organizations that work collectively as a network but with no distinct governance entity. Governance of collective activities resides completely with the network members themselves. Figure 1(a) offers a graphic illustration of this model. Here, it is the network participants themselves that make all the decisions and manage network activities. There is no distinct, formal administrative entity, although when there are more than a handful of network participants, some administrative and coordination activities may be performed by a subset of the full network. The strength of this model is the inclusion and involvement of all network participants and its flexibility and responsiveness to network participant needs. Its weakness is its relative inefficiency. It is a model that seems best suited to small, geographically concentrated networks where full and active face-to-face participation by network participants is possible.

Lead organization form

Lead organization governance, our second network form, is common in vertical relationships between buyers and suppliers or between funders and recipient

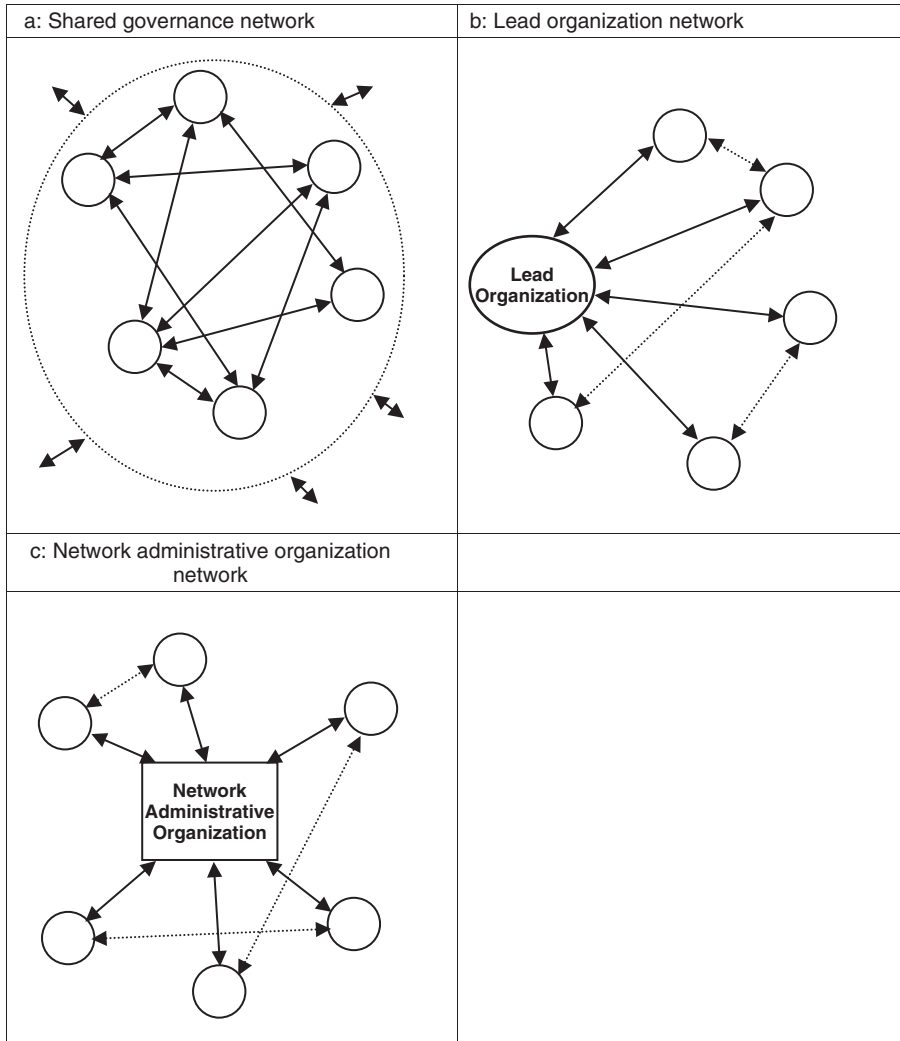


FIGURE 1 *Forms of network governance*

organizations. Examples of this form in a business context include the Keiretsu models of Japanese manufacturing (Gerlach 1992) and similar models of cooperative buyer-supplier relations (Uzzi and Lancaster 2003), where there is a single powerful/large buyer and a number of weaker and smaller suppliers. It can also occur in horizontal multilateral networks, most often when one organization has sufficient resources and legitimacy to play a lead role. This is a model that occurs often in publicly funded health and human services where there may be a core provider agency that assumes the role of network leader because of its central position in the flow of clients and key resources. The lead organization in a community health setting may be a hospital or a health clinic (Weiner and Alexander 1998); in mental health, a community mental health centre may fill this role (Provan and Milward 1995); and a city's police department is likely to be the lead organization in a community policing network.

Figure 1(b) visually represents the lead organization model. In this structure, network members all share at least some common purpose (as well as maintaining individual goals) and they may interact and work with one another (the dotted lines). However, all activities and key decisions are coordinated through and by one of the members acting as a lead organization. This organization provides administration for the network and/or facilitates the activities of member organizations in their efforts to achieve network goals.

The functionality of the lead organization model is its efficiency and the legitimacy provided by the lead agency. Because of its capacity to take on most of the responsibilities of running and coordinating network activities, most of the complexity and messiness inherent in the self-governed model can be avoided, and the network can be governed quite efficiently. The size and importance of the lead organization to outside parties, such as large funders, typically provides significant legitimacy to network participants. The weakness of the model is that the lead organization may have its own agenda and can readily dominate the other network members, causing resentment and resistance. In addition, because the lead organization takes on many of the activities of governing the network, network members can readily lose interest in network-level goals and focus instead on their own self-interest, undermining the viability of the network. While this model may result from a bottom-up process of network building, it may also result from mandate, especially when a government entity provides major funding to a single organization which then controls the network.

Network administrative form

Because of the inefficiency of shared governance or the problems of dominance and resistance inherent in lead organization networks, a network administrative organization (NAO) governance model may be used as an alternative. The basic idea is that a separate administrative entity is set up specifically to manage and coordinate the network and its activities. Like the lead organization model, the NAO plays a key role in coordinating and sustaining the network. Unlike the lead organization model, however, the NAO is not another network agency, providing its own set of services. Instead, the NAO is established with the exclusive purpose of network governance. It may be a government entity or, more likely, a nonprofit, which is typically the case even when the network members are for-profit firms (cf. Human and Provan 2000).

NAO's may be relatively informal structures, consisting of a single individual who acts as the network facilitator or broker, or they may be much more formalized and complex, consisting of executive director, staff, and board operating out of a physically distinct office. This more formalized version is especially likely when the NAO seeks official recognition to enhance its legitimacy among both internal and external stakeholders. Government run NAOs are generally set up (by mandate) when the network first forms, in order to stimulate its growth through targeted funding and/or network facilitation, and to ensure that network goals are met.

Figure 1(c) is a graphic depiction of the network administrative organization governance model. With this structure, network participants and groups may interact and work with one another (the dotted lines) but activities and key decisions are coordinated through and by a separate, independent entity. This governance entity may be a nonprofit corporation, an unincorporated nonprofit, or a governmental organization. The strengths of this model are its sustainability and legitimacy, especially to outsiders and, to a lesser extent, its efficiency. Its weaknesses are that network participants may rely on the governance entity too heavily and it may adopt decision-making processes that seem overly bureaucratic.

Consistent with the fact that each of the three network governance forms is unique, each has characteristics that make it more or less likely that a particular performance criterion will be more appropriate. For instance, efficiency is not an effectiveness criterion that appropriately fits a shared governance form. Conversely, high levels of multi-organizational collaboration should not be expected from the lead organization form. The leads us to the following general proposition:

Proposition 1: The performance criteria that are most appropriate for evaluating a network will depend, in part, on the type of network governance form adopted.

MANDATED VERSUS VOLUNTARY INCEPTION OF THE NETWORK

'Voluntary networks' are created bottom-up by the professionals and organizations that will participate in the network, whereas 'mandated networks' are created by policy dictate, typically by a government agency (see Laumann *et al.* 1978; Selsky and Parker 2005). We do not know of any study that has specifically looked into whether the appropriateness of performance criteria differs across mandated and voluntary networks. However, the fact that mandated and voluntary networks are significantly different phenomena has already been convincingly argued by Mandell (1990). In particular, she demonstrated that strategic management in a mandated network is different from strategic management in a voluntary network. Some indirect evidence can be presented, however, for the fact that whether a network is voluntary or mandated will affect the type of performance criteria that is most appropriate.

For example, Van Raaij (2006) studied norms, or the criteria network members used to assess the performance of their networks. She studied four sub-regional service delivery networks for patients with non-congenital brain injuries (NBI). One of her findings was that there seemed to be a clear difference in the type of outcomes that can realistically be attained by voluntarily versus mandatory networks. The three outcome criteria – network legitimacy, activating capacity (the ability to keep the network going), and network climate (referring to the balance between the accomplishment of organizational interests and the willingness to go beyond their own organization's interests, making decisions beneficial for the network as a whole) – were seen by participants as appropriate for networks that were formed voluntarily but not for mandated networks.

The finding – that these criteria were deemed to be inappropriate for mandated networks – is related to the fact that not all relevant levels in the participating organizations were convinced of the value of use of a network. In the mandated networks, Van Raaij found that while the operating-level professionals were themselves convinced of the need for the network, it was extremely difficult to convince their managerial directors. In contrast, networks formed voluntarily already had a history of recognizing the need to coordinate their activities with other organizations. It was not difficult to have service professionals in these organizations convince their fellow organizational members of the need to collaborate. When network members are not actively involved in the network's inception (as is the case in mandatory networks), the network will be less self-activating; thus, members will have difficulties balancing organizational and network interests. Consequently, assessing mandatory networks based on their level of self-activation is not appropriate.

Another study (Addicott and Ferlie 2006) indirectly points to the fact that the appropriateness of type of performance criteria differs between mandated and voluntarily formed networks. In a study of London Cancer Networks they found a considerable

discrepancy between what the Department of Health considered to be criteria of success, versus what those delivering direct patient care regarded as important. From the eight criteria considered important by the stakeholders, only one turned out to also be a criterion which the 'remote centre' considered important – a timely diagnosis as treatment for all patients. What this study demonstrated is that the criteria used to evaluate the network by those outside the network (that is, those who are responsible for the mandate) were not likely to be consistent with what the actors in the network actually believed was important.

We argued above that a network needs a specific form in order to meet certain functionality. In the same way, one can argue here that a network needs a common foundation of norms in order to achieve network-level outcomes such as a positive network climate, network legitimacy and activating capacity. Because they lack a common normative system, at least at initial formation, it is inappropriate to assess mandated networks on these criteria. It seems that policy-makers cannot at the same time mandate the inception of networks while also expecting that these contain a normative system. To assess these networks after inception as if they were formed voluntarily is, thus, not appropriate.

What we describe here is also related to the importance of the legitimacy of any form of coordination. As stated by Meyer and Rowan: 'Organizations that . . . lack acceptable legitimated accounts of their activities . . . are more vulnerable to claims that they are negligent, irrational or unnecessary' (Meyer and Rowan 1991, p. 50). In a study of two small-firm multilateral networks, Human and Provan (2000) made a distinction between internal and external legitimacy. Internal legitimacy refers to the fact the network is positively assessed by the criteria of those participating in the network. Consistent with our argument above, it is unlikely that internal legitimacy will be high in mandated networks. Human and Provan (2000) also demonstrated that the sequence of legitimacy building efforts of a network may affect its success. Having conducted two longitudinal case studies, they found that the network that focused on building its internal legitimacy first, and then turned to building legitimacy with its external constituents, managed to withstand a major crisis and ultimately succeeded. In contrast, the network that first built its external legitimacy never managed to overcome its internal legitimacy problems and ultimately failed.

External legitimacy refers to the fact that a network is positively assessed by the criteria of outside constituents such as funders, regulators, the public, the media, and so on, something that is most likely in mandated networks. As noted in the study by Addicott and Ferlie (2006), these can be very different from the criteria used as a basis for assessing internal legitimacy. Although it is clear that external legitimacy is crucial in attracting customers, securing funding, dealing with government, and so on, it becomes clear from the above that external legitimacy criteria might be more feasible for mandatory than for voluntary networks, especially at early stages of network evolution. This argument is especially important in the public sector, where government entities may use their power to fund and regulate to mandate formation of networks.

This leads to a proposition suggesting that there is a relationship between the type of network formed at initial inception and the appropriateness of different effectiveness criteria, especially legitimacy:

Proposition 2: The performance criteria that are most appropriate for evaluating a network will depend, in part, on whether a network is voluntary or mandated.

THE DEVELOPMENTAL STAGE OF THE NETWORK

Quinn and Cameron, in their 1983 article entitled 'Organizational Life Cycle and Shifting Criteria of Effectiveness', observed that the 'appropriateness of different criteria at different times in the organizational life cycle has seldom been considered' (p. 41). Their claim seems still to hold today for organizations and is certainly true for networks.

As with any type of coordination, networks evolve across different developmental stages. For example, D'Aunno and Zuckerman (1987) proposed a four-stage life cycle model of network development and identified factors that are critical in the transition of the network from one stage to the next. Largely based on the literature on life cycle models of individual organizations (Kimberly *et al.* 1980; Quinn and Cameron 1983), they proposed a four-stage model: emergence of a coalition, transition to a federation, maturity of federation, and critical crossroads. Although their research was on hospital federations, their life cycle model could also be applied to the development of most public networks (for the broader literature on the evolutionary dynamics of networks, see Ring and Van de Ven 1994; Human and Provan 2000; Kenis and Knoke 2002; Isett and Provan 2005; Powell *et al.* 2005). The issue at this point is not to present a new model of network developmental stages or to discuss the key factors present at each stage. Rather, it is to develop an argument that different assessment criteria can be more or less appropriate at different developmental stages.

The importance of evolution as a factor in what can and cannot be accomplished has already been described in the organizational literature. At certain stages some criteria can simply not be achieved and are thus not appropriate. It is clear, for instance, that when a network is newly emergent, goal attainment will be problematic (see Annie E. Casey Foundation 1995). Most of the time and energy of network members will instead be spent developing network structures and processes, rather than on achieving outcomes (Dockery 1996; Goss 2001). In contrast, mature networks should be expected to be able to attain network-level goals and to be relatively efficient in operation. At this point, a process of interaction and collaboration should already have developed, along with clear structure for performing, so measuring these outcomes will be less relevant. During the early growth stage, most networks should be expected to develop legitimacy, but not necessarily to *be* legitimate. Following this line of reasoning, the following proposition can be stated:

Proposition 3: The performance criteria that are most appropriate for evaluating a network will depend, in part, on the developmental stage of the network.

CONCLUSIONS AND DISCUSSION

We have identified three exogenous dimensions of network effectiveness, namely, governance form, type of inception (voluntary versus mandated), and developmental stage. These are exogenous factors because while they are key determinants of the extent to which a particular evaluation criterion is appropriate for assessing network performance, they are largely outside the control of network members.

Our approach is based on three principles. First, when talking about the 'functioning', 'effectiveness', 'performance', and so on of public networks, exactly what criteria are being considered should be clearly specified. Is it efficiency, survival, goal attainment, equity, client satisfaction? Each of these has different implications and should be considered as indicators of distinct aspects of network performance.

Second, each of these criteria is equally valid from a normative point of view. We cannot think of any normative argument to suggest why a public network should be assessed on its innovativeness rather than on its efficiency, for example. In line with Czarniawska (2003), we are equally critical of the tendency in organization science to start research from an 'ideal point' as perceived by scientists. Instead, we believe it is most appropriate to consider a wide range of potential effectiveness criteria, such as the many points identified by Provan and Milward (2001) in their article 'Do Networks Really Work?'

Third, and most central to the theme of the paper, is that the evaluation criteria selected must be appropriate. In other words, criteria have to be closely connected to what the network can realistically achieve. The challenge in this paper has been to identify those key factors that determine whether certain network evaluation criteria can be achieved or not or, in other words, which evaluation criteria are appropriate and when.

We have identified several broad conclusions that can be drawn from our approach. For one thing, we have demonstrated that conflicts about whether or not a public network performs effectively can have three causes. First, the different parties addressing performance (researchers, network members, key stakeholders, and so on) might actually be referring to different things when assessing a network. For example, one may be referring to productivity and another to goal attainment. Clarity about which criteria are being considered is critical if network performance is to be evaluated in a meaningful way. Second, there can be conflict about norms; for example, should the network be efficient or should it contribute to client satisfaction? This is a discussion about elements of value, not elements of fact, leaving little room for social scientists to contribute anything apart from observing the phenomenon. Third, there may be conflict in terms of what the assessing party expects and what the party assessed can actually achieve. An interesting case in point here is the NHS (National Health Service) in Britain. The recently published report entitled 'What is Productivity?' was presented in order to communicate to the outside world that the NHS should not be assessed on productivity since it is inconsistent with their mission (NHS 2006). Instead, the report argued, the NHS should be assessed on their contribution to the quality of care.

A second conclusion is that our approach forces both network researchers and public policy officials who design and fund networks to distinguish between exogenous and endogenous performance factors. This difference is crucial. If endogenous factors are regarded as the cause for underperformance, responsibility would lay with the network itself. In such a case the reason may be plain mismanagement. In contrast, if exogenous factors are regarded as the cause of underperformance, responsibility would lie outside the network and would have to do with the fact that the network assessment may be based on inappropriate criteria related to its governance form, evolutionary stage, and so on. Referring to the NHS example above, the question is thus whether the fall over the past decade by between 0.6 per cent and 1.3 per cent a year in the productivity of the NHS is a case of mismanagement or whether the NHS has been assessed using an inappropriate criterion (that is, productivity). Although we believe our point is valid and important, we understand that at the same time it provokes a number of issues which need further discussion. In particular, we have assumed that inappropriate criteria are more likely to be set by outsiders (for instance, by government through mandate) than by insiders, who are expected to have a better sense for what is feasible and what is not. But it may also be that insiders have conflict over network goals and outcomes. Pointing out that network managers are constrained in what they can instrumentally manage or control, and making explicit what these constraints are, are two different things. We have

concentrated on the former and given some evidence regarding the latter. However, far more discussion is needed, especially in determining exactly what criteria are appropriate under what specific conditions.

Another conclusion we can make that builds on our basic argument is that the more constituencies (or stakeholders) a network must be responsive to, the more likely it is that the network will be perceived as poorly performing. If we assume that constituencies are defined by the specific criteria they use to assess a situation, then more constituencies imply a broader range of diverse evaluation criteria. Moreover, it is unlikely that constituencies will take notice of the importance of the factors that determine whether a network *can* achieve something. Rather, they will tend to base their views on the principle that the network *should* achieve what they consider to be important. Thus, it is likely that the more constituencies a network has, the more the network will, on an aggregate level, be perceived as under-performing. All this indicates that the perceived performance of a network is a function of the types of independent criteria on which the network is assessed and whether or not these criteria are appropriate. Multiple network constituencies are especially likely in a public sector context.

While we would like to generalize about our arguments, and consistent with this last point, our approach is most likely to hold for public sector networks. In part, this is because in the public sector, discussions about what performance means are much more open to political interpretation ('should trains run on time, be cheap or safe?') and, consequently, the normative aspects of our work are especially relevant. There is, of course, nothing wrong with political discourse about what a society wants to achieve, but the greater the number of achievement criteria that are considered, the higher the chance that inappropriate criteria will be used to assess outcomes. In addition, public networks (including nonprofits) are much more likely to have a collective goal than in the private sector, where the performance of individual organizations is paramount. The nature of public services is such that the issues addressed are typically highly complex (O'Toole 1997), necessitating collective action. Thus, consideration of overall network performance is likely to be much more salient than in the private sector.

Our work leaves a number of issues unaddressed. In particular, it is certainly possible that use of a certain criterion in the assessment of a network might be appropriate based on one exogenous factor but not on the basis of another. For example, a criterion such as efficiency might be appropriate for a network in a later stage or age of maturity, but might at the same time be inappropriate because the network also has a shared governance form, which is inappropriate for efficiency. If we radicalize this argument it might lead to a situation where a network cannot be assessed on any performance criterion at all because for any criterion one could point to an exogenous factor that makes the criterion inappropriate. Another possibility could be that these exogenous factors cluster in some way or the other and that only a limited set of configurations is possible. For instance, it may be that only voluntary networks can be shared-governed or that mature networks are more likely to be governed either through a lead organization or NAO. Consequently, it might be useful to look for configurations, or in other words 'commonly occurring clusters of attributes [. . .] that are internally consistent, such that the presence of some attributes can lead to reliable prediction of others' (Miller and Mintzberg 1983, p. 57). It might be the case that the appropriateness of a specific criterion is part of a multidimensional constellation of conceptually and analytically distinct characteristics that commonly occur together (Drazin and Van de Ven 1985; Miller 1987; Meyer *et al.* 1993).

Overall, we have tried to demonstrate that understanding the performance of networks is a complex issue. It is much more than just a matter of starting from a preferred criterion and trying to uncover the conditions that lead to higher performance. Consequently, researchers should be skeptical about the many lists of so-called 'success' factors presented in the literature on the performance of networks. Our paper is an attempt to move thinking well beyond this simplistic view. Apart from pointing to the complexity of the performance issue for researchers, we hope also to contribute to the practice of assessing network performance. By introducing the concept of exogenous factors and by introducing ideas that will, hopefully, stimulate an alternative perspective for conducting research on network performance, we hope to demonstrate to policy-makers that it is not useful to assess all types of networks on any single criterion or to focus solely on endogenous factors related to network leadership and management.

ACKNOWLEDGEMENT

An earlier version of this paper was presented at *The Determinants of Performance in Public Organizations II* workshop, December 2006, at the University of Hong Kong.

REFERENCES

- Addicott, R. and E. Ferlie. 2006. 'A Qualitative Evaluation of Public Sector Organizations: Assessing Organizational Performance in Healthcare', in G.A. Boyne, K.J. Meier, L.J. O'Toole and R.M. Walker (eds), *Public Service Performance: Perspectives on Measurement and Management*. Cambridge: Cambridge University Press.
- Agranoff, R. 1991. 'Human Services Integration: Past and Present Challenges in Public Administration', *Public Administration Review*, 51, 6, 533–42.
- Agranoff, R. 2003. *Leveraging Networks: A Guide for Public Managers Working Across Organizations*. Arlington, VA: IBM Endowment for the Business of Government.
- Annie E. Casey Foundation. 1995. *The Path of Most Resistance*. Baltimore, MD: Annie E. Casey Foundation.
- Baker, E.L., R.J. Melton, P.V. Stange, et al. 1994. 'Health Reform and the Health of the Public: Forging Community Health Partnerships', *Journal of the American Medical Association*, 272, 16, 1276–82.
- Baker, W.E. and R.R. Faulkner. 1993. 'The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry', *American Sociological Review*, 58, 6, 837–60.
- Baum, J. and C. Oliver. 1991. 'Institutional Linkages and Organizational Mortality', *Administrative Science Quarterly*, 36, 2, 187–218.
- Boyne, G.A. 2003. 'What is Public Service Improvement?', *Public Administration*, 81, 2, 211–27.
- Brass, D., J. Galaskiewicz, H. Greve and W. Tsai (eds). 2004. 'Building Effective Networks', *Academy of Management Review*, special issue.
- Burt, R.S. 1992. *Structural Holes*. Cambridge, MA: Harvard University Press.
- Burt, R.S. 1997. 'The Contingent Value of Social Capital', *Administrative Science Quarterly*, 42, 339–65.
- Burt, R.S. 2005. *Brokerage and Closure. An Introduction to Social Capital*. Oxford: Oxford University Press.
- Chisholm, R.F. 1998. *Developing Network Organizations*. Reading, MA: Addison-Wesley.
- Connolly, T., E.J. Conlon and S.J. Deutsch. 1980. 'Organizational Effectiveness: A Multiple-constituency Approach', *Academy of Management Review*, 5, 2, 211–17.
- Cross, R., A. Parker and S. Borgatti. 2002. *A Bird's-eye View: Using Social Network Analysis to Improve Knowledge Creation and Sharing*. Somers, NY: IBM Corporation.
- Czarniawska, B. 2003. 'Forbidden Knowledge: Organization Theory in Times of Transition', *Management Learning*, 33, 353–65.
- D'Aunno, T. and H. Zuckerman. 1987. 'A Life-cycle Model of Organizational Federations: The Case of Hospitals', *Academy of Management Review*, 12, 3, 534–45.
- Dockery, G. 1996. 'Rhetoric or Reality? Participatory Research in the NHS', in K. De Koning and M. Martin (eds), *Participatory Research in Health*, 164–76.
- Drazin, R. and A.H. Van de Ven. 1985. 'Alternative Forms of Fit in Contingency Theory', *Administrative Science Quarterly*, 30, 514–39.
- Gerlach, M.L. 1992. *Alliance Capitalism: The Social Organization of Japanese Business*. Berkeley, CA: University of California Press.
- Goldsmith, S. and W.D. Eggers. 2004. *Governing by Network: The New Shape of the Public Sector*. Washington, DC: Brookings Institution Press.

- Goss, S. 2001. 'Making Local Governance Work – Networks, Relationships and Management of Change'. Basingstoke: Palgrave.
- Gouldner, A.W. 1957. 'Theoretical Requirements of the Applied Social Sciences', *American Sociological Review*, 22, 92–102.
- Hannan, M.T. and J. Freeman. 1977. 'Obstacles to Comparative Studies', in P.S. Goodman and J.M. Penning (eds), *New Perspectives on Organizational Effectiveness*. San Francisco, CA: Jossey-Bass.
- Human, S.E. and K.G. Provan. 2000. 'Legitimacy Building in the Evolution of Small-firm Networks: A Comparative Study of Success and Demise', *Administrative Science Quarterly*, 45, 2, 327–65.
- Isett, R. and K.G. Provan. 2005. 'The Evolution of Dyadic Interorganizational Relationships in a Network of Publicly Funded Nonprofit Agencies', *Journal of Public Administration and Theory*, 15, 1, 149–65.
- Israel, B.A., A.J. Schulz, E.A. Parker and A.B. Becker. 1998. 'Review of Community-based Research: Assessing Partnership Approaches to Improve Public Health', *Annual Review of Public Health*, 19, 173–202.
- Jacobs, R., M. Goddard and P.C. Smith. 2006. 'Public Services: Are Composite Measures a Robust Reflection of Performance in the Public Sector?', CHE Research Paper 16. York: Centre for Health Economics.
- Jennings, E.T. and J.A.G. Ewalt. 1998. 'Interorganizational Coordination, Administrative Consolidation, and Policy Performance', *Public Administration Review*, 58, 5, 417–28.
- Kenis, P. and D. Knoke. 2002. 'How Organizational Field Networks Shape Interorganizational Tie-formation Rates', *Academy of Management Review*, 27, 2, 275–93.
- Kickert, W.J.M., E.-H. Klijn and J. Koppenjan. 1997. *Managing Complex Networks: Strategies for the Public Sector*. London: Sage.
- Kimberly, J.R., R.H. Miles and Associates. 1980. *The Organizational Life Cycle: Issues on the Creation, Transformation, and Decline of Organizations*. San Francisco, CA: Jossey-Bass.
- Klitgaard, R. and P.C. Light (eds). 2005. *High Performance Government: Structure, Leadership, Incentives*. Santa Monica, CA: RAND Corporation.
- Koppenjan, J. and E.-H. Klijn. 2004. *Managing Uncertainties in Networks*. London: Routledge.
- Laumann, E.O., J. Galaskiewicz and P. Marsden. 1978. 'Community as Interorganizational Linkages', *Annual Review of Sociology*, 4, 455–84.
- Mandell, M.P. 1990. 'Network Management: Strategic Behavior in the Public Sector', in R.W. Gage and M.P. Mandell (eds), *Strategies for Managing Intergovernmental Policies and Networks*. New York: Praeger.
- Mandell, M.P. 1994. 'Managing Interdependencies Through Program Structures: A Revised Paradigm', *American Review of Public Administration*, 24, 1, 99–121.
- Mandell, M.P. and R. Keast. 2006. *Evaluating Network Arrangements: Toward Revised Performance Measures*, paper presented at the Conference 'A Performing Public Sector: The Second Transatlantic Dialogue', Leuven 1–3 June.
- Meier, K.J., L.J. O'Toole, Jr. 2001. 'Managerial Strategies and Behavior in Networks', *Journal of Public Administration Research and Theory*, 11, 271–95.
- Meyer, A.D., A.S. Tsui and C.R. Hinings. 1993. 'Configurational Approaches to Organizational Analysis', *Academy of Management Journal*, 36, 1175–95.
- Meyer, J.W. and B. Rowan. 1991. 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', in W.W. Powell and P.J. DiMaggio (eds), *The New Institutionalism in Organizational Analysis*. Chicago, IL: University of Chicago Press.
- Meyer, M.W. 2002. *Rethinking Performance Measurement*. Cambridge: Cambridge University Press.
- Miller, A.D. and H. Mintzberg. 1983. 'The Case for Configuration', in G. Morgan (ed.), *Beyond Method: Strategies for Social Research*. Beverly Hills, CA: Sage.
- Miller, A.D. 1987. 'The Genesis of Configuration', *Academy of Management Review*, 12, 686–701.
- Miner, A.S., T.L. Amburgey and T. Stearns. 1990. 'Interorganizational Linkages and Population Dynamics: Buffering and Transformational Shields', *Administrative Science Quarterly*, 35, 689–713.
- NHS. 2006. 'What is Productivity?' (<http://www.nhsconfed.org/docs/confidence2.pdf>), accessed 1 November 2006.
- O'Toole, Jr., L.J. 1997. 'Treating Networks Seriously: Practical and Research Based Agendas in Public Administration', *Public Administration Review*, 57, 45–52.
- Piore, M. and C. Sabel. 1984. *The Second Industrial Divide: Possibilities for Prosperity*. New York: Basic Books.
- Powell, W.W., D.R. White, K.W. Koput and J. Owen-Smith. 2005. 'Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences', *American Journal of Sociology*, 110, 1132–205.
- Provan, K.G., A. Fish and J. Sydow. 2007. 'Interorganizational Networks at the Network Level: A Review of the Empirical Literature on Whole Networks', *Journal of Management*, 33, 479–516.
- Provan, K.G. and P. Kenis. 2008. 'Modes of Network Governance: Structure, Management, and Effectiveness', *Journal of Public Administration Research and Theory*, 18, 2, 229–57.
- Provan, K.G. and H.B. Milward. 1995. 'A Preliminary Theory of Network Effectiveness: A Comparative Study of Four Community Mental Health Systems', *Administrative Science Quarterly*, 40, 1–33.
- Provan, K.G. and H.B. Milward. 2001. 'Do Networks Really Work? A Framework for Evaluating Public Sector Organizational Networks', *Public Administration Review*, 61, 4, 400–9.

- Provan, K.G., M.A. Veazie, L.K. Staten and N.I. Teufel-Shone. 2005. 'The Use of Network Analysis to Strengthen Community Partnerships', *Public Administration Review*, 65, 5, 603–13.
- Quinn, R. and K. Cameron. 1983. 'Organizational Life Cycles and Shifting Criteria of Effectiveness: Some Preliminary Evidence', *Management Science*, 29, 33–41.
- Ring, P.S. and A.H. Van de Ven. 1994. 'Developmental Processes of Cooperative Interorganizational Relationships', *Academy of Management Review*, 19, 90–118.
- Selsky, J.W. and B. Parker. 2005. 'Cross-sector Partnerships to Address Social Issues: Challenges to Theory and Practice', *Journal of Management*, 31, 849–73.
- Simon, H.A. 1976. *Administrative Behavior*, 2nd edn. New York: The Free Press.
- Stuart, T.E., H. Hoang and R.C. Hybels. 1999. 'Interorganizational Endorsements and the Performance of Entrepreneurial Ventures', *Administrative Science Quarterly*, 44, 2, 315–49.
- Uzzi, B. and R. Lancaster. 2003. 'Relational Embeddedness and Learning: The Case of Bank Loan Managers and their Clients', *Management Science*, 49, 4, 383–99.
- Van Raaij, D. 2006. 'Norms Network Members Use: An Alternative Perspective for Indicating Network Success or Failure', *International Public Management Journal*, 9, 3, 249–70.
- Weiner, B.J. and J.A. Alexander. 1998. 'The Challenges of Governing Public-private Community Health Partnerships', *Health Care Management Review*, 23, 2, 39–55.
- Weiss, C. 1998. *Evaluation Research. Methods of Assessing Program Effectiveness*. Englewood Cliffs, NJ: Prentice Hall.
- Weller, J. and E.L. Quarantelli. 1973. 'Neglected Characteristics of Collective Behavior', *American Journal of Sociology*, 79, 66–85.
- Wiewel, W. and A. Hunter. 1985. 'The Interorganizational Network as a Resource: A Comparative Case Study on Organizational Genesis', *Administrative Science Quarterly*, 30, 4, 482–96.
- Yuchtman, E. and S. Seashore. 1967. 'A System Resource Approach to Organizational Effectiveness', *American Sociological Review*, December, 891–903.

Date received 11 April 2007. Date accepted 12 July 2007.