

In: Zeitschrift für Evaluation, Heft 1, 2005, S. 7-28

Selbstevaluation durch Leistungsvergleiche in deutschen Kommunen

Sabine Kuhlmann
(Universität Konstanz)

1 Hintergrund: Leistungsmessung in der deutschen Verwaltungsreformtradition

Leistungsmessung (*performance measurement*) ist in der OECD-Staatenwelt nahezu durchgängig zu einem vorrangigen Tagesordnungspunkt bei der Modernisierung des Öffentlichen Sektors geworden. Haben entsprechende Reformansätze zweifelsohne durch den betriebswirtschaftlichen Managerialismus, der unter dem Stichwort New Public Management (NPM) zunächst in den angelsächsischen Ländern Furore machte und seit den frühen 90er Jahren als „Neues Steuerungsmodell“ (NSM) auch die Modernisierungsdiskussion in der Bundesrepublik ergriffen hat (vgl. KGSt 1993; Bogumil/Kuhlmann 2004), einen deutlichen Auftrieb erfahren, findet sich der strategische Grundgedanke von Leistungsmessung und Evaluation auch bereits in früheren Reformanläufen (Wollmann 2004a):

Ging es unter den Bedingungen wirtschaftlicher Prosperität und wachsender Budgets in den 1960er Jahren vor allem darum, (kostspielige) Interventionsprogramme und Reformmaßnahmen, die im Hinblick auf ihre Wirksamkeit und ihre Innovationskraft als ungesichert schienen und dadurch neue Informationsbedürfnisse auslösten, mittels indikatorengestützter Wirkungskontrollen effektiver zu gestalten, bestand in den 1970er Jahren das Anliegen von Wirkungsanalysen angesichts veränderter ökonomischer Rahmenbedingungen und der Notwendigkeit, mit knappen Ressourcen zu haushalten, zunehmend darin, ineffektive Programme zu erkennen und so zu legitimierbaren Budgetkürzungen zu gelangen (Derlien 1997: 6 f.). Indem die New Public Management-Bewegung der 1980er und 1990er Jahre darauf gerichtet war/ist, die Leistungsfähigkeit öffentlichen Handelns durch eine Steuerung über Ergebnisse/Resultate (Output-Steuerung, *management by results*) zu steigern, greift sie die für die Planungsdiskussion und Programmevaluation der 1960er/1970er Jahre charakteristische Forderung nach Ergebnis- und Wirkungstransparenz (wieder) auf. In der Modellvorstellung des NPM/NSM gewinnen Performanzmessungen und Leistungsindikatoren dadurch noch zusätzlich an Bedeutung, dass

sie den (angemahnten) Übergang von der klassischen Regel- zur Ergebnis- und Produktsteuerung und von der hierarchischen zur kontraktuellen Verwaltungsführung (at arm's length) stützen sollen.

Mit dem Übergang von der betriebswirtschaftlich inspirierten Managementbewegung zu einer – unter dem Schlagwort „Governance“ – wieder stärker die Außenbeziehungen und das gesellschaftliche Umfeld von Politik und Verwaltung berücksichtigenden Reformdiskussion (vgl. Benz 2003; Klenk/Nullmeier 2003) hat sich in den letzten Jahren erneut der analytische Fokus und Gegenstandsbereich von Performance Measurement im öffentlichen Sektor verschoben. Angesichts der zunehmenden Kritik am ökonomischen Reduktionismus der NPM-Reformen, der dazu geführt hatte, dass sich Leistungsmessungen vielfach auf betriebswirtschaftliche Steuerungsgrößen beschränkten, treten nunmehr wieder die Verbindungslinien zwischen politisch-administrativem System und gesellschaftlichem Handlungs- und Wirkungsfeld in den Vordergrund. Darin, dass die Frage nach der Wirksamkeit (*impacts, outcomes*) politischen Handelns, nach den Möglichkeiten (und Grenzen) politisch-strategischer Steuerung und wirkungsorientiertem Politik- und Verwaltungshandeln vermittels Ziel- und Wirkungsindikatoren verstärkt auf die reformpolitische Agenda (weniger allerdings bislang in die konkrete praktische Umsetzung) tritt, deutet sich an, dass zumindest die Debatte um Performance Measurement und Evaluation wieder „politischer“ geworden ist (Jann 2002: 295).

Vor dem Hintergrund dieser Entwicklungen wird im folgenden Beitrag der interkommunale Leistungsvergleich in Deutschland als eine Form von Performanzbeobachtung und interner Selbstevaluation näher betrachtet. Im Mittelpunkt steht dabei die (evaluative) Frage, welchen Nutzen und welche Erfolge interkommunale Vergleichsprojekte bislang hervorgebracht haben und in welchem Verhältnis diese zu den aufgewendeten Kosten stehen, die für die Konzipierung, Implementation und laufende Fortschreibung des Leistungsvergleichs erforderlich sind. Damit soll eine klaffende Lücke in der bisherigen Verwaltungs(reform)forschung, die bislang kaum empirische Hinweise auf die Ergebnisse, Wirkungen und Kosten der lokalen Verwaltungsmodernisierung geliefert hat¹, ein Stück weit geschlossen werden. Es wird in drei Schritten vorgegangen:

¹ Ein Forschungsprojekt, das an der Universität Konstanz im Verbund mit den Universitäten Marburg, Potsdam und der Humboldt-Universität zu Berlin durchgeführt und von der Hans-Böckler-Stiftung finanziert wird und an dem die Verfasserin mitwirkt, wendet sich derzeit unter der Themenstellung „Zehn Jahre Neues Steuerungsmodell – Evaluation kommunaler Verwaltungsmodernisierung“ dieser Frage zu (vgl. auch Kuhlmann 2004a; Bogumil/Kuhlmann 2004)

- Zunächst wird ein Überblick über abgeschlossene und laufende interkommunale Vergleichsprojekte in Deutschland gegeben (Abschnitt 2).
- Sodann wird der Versuch unternommen, die Ergebnisse und – soweit ersichtlich – Wirkungen des interkommunalen Leistungsvergleichs abzuschätzen und den von den lokalen Akteuren aufgewendeten Konzipierungs-, Implementations- und Transaktionskosten gegenüberzustellen (Abschnitt 3)
- Indem abschließend auf die Besonderheiten politischer Rationalität in konkurrenzdemokratisch verfassten Systemen Bezug genommen wird, soll der interkommunale Leistungsvergleich in Deutschland einer resümierenden Bewertung unterzogen und sollen Perspektiven für die Zukunft aufgezeigt werden (Abschnitt 4).

2 Interkommunale Vergleichsprojekte in Deutschland

Interkommunale Leistungsvergleiche stellen zum einen eine Form des nicht-marktlichen Wettbewerbs im öffentlichen Sektor dar (vgl. Reichard 1998), die von der (institutionenökonomisch inspirierten) Annahme geleitet sind, dass die Anreizwirkung des echten Marktwettbewerbs durch andere (intrinsische und extrinsische) Anreize ersetzt und Wettbewerb „simuliert“ werden kann (Wettbewerbssurrogat)². Zum andern sind sie als Form standardisierter, indikatoren- und kennzahlengestützter interner Selbstbewertung und –kontrolle in den weiteren Kontext von Evaluation und Wirkungsanalyse einzuordnen (vgl. Pitarelli/ Monnier 2000: 9 ff.; Wollmann 2000), deren Gegenstand die Beobachtung und der interorganisatorische Vergleich von Prozessen, Strukturen, Ergebnissen (outputs) und Wirkungen (outcomes) öffentlichen Handelns sein können.

Im deutschen Reformkontext wurde der Startschuss zu einer zeitweise um sich greifenden regelrechten „Vergleichseuphorie“ durch die Bertelsmann Stiftung im Jahre 1990 mit ihrem Projekt „Grundlagen einer leistungsfähigen Kommunalverwaltung“ gegeben (siehe Tabelle 1; vgl. auch Schuster 2001, 2003), gefolgt von der KGSt, die mit der Gründung IKO-Netzes im Jahre 1996 ein „internes Informationssystem der Kommunen“ etablierte. Dieses umfasste im Jahr 2001 immerhin 4% aller westdeutschen Kommunen und, bezogen auf das Jahr 2002, bundesweit 748 Kommunen (einschließlich Doppelmitgliedschaften) in 73 Vergleichsringen

² Von einem – in eher seltenen Fällen anzutreffenden - Benchmarking der Kommunen kann dort die Rede sein, wo der Leistungsvergleich auf die Identifikation von *best practice*-Lösungen zielt und Unterschiede zwischen den beteiligten Institutionen zum Anlass für institutionelle Lernprozesse und Innovationen genommen werden (vgl. Kuhlmann 2003a).

(Kuhlmann 2003a; vgl. auch Korte 2004)³. Weitere Impulse kamen wiederum von der Bertelsmann Stiftung, die im Jahr 2000 die – inzwischen abgeschlossenen - Vergleichsprojekte „kik“ und „kompass“ gründete (Tabelle 1; vgl. auch Tebbe 2004). Neben und teils – z.B. in Bayern und im Saarland - als bewusste Alternative zu diesen überregionalen Vergleichsprojekten der KGSt und der Bertelsmann Stiftung haben sich seit Mitte/ Ende der 90er Jahre zudem in fast allen Bundesländern interkommunale Vergleichsprojekte etabliert⁴, die teils von den kommunalen Landesverbänden moderiert, mitunter auch (temporär) vom Land bezuschusst (Bsp. Bayern, Nordrhein-Westfalen, Sachsen, Schleswig-Holstein), teils in der Eigenregie der Kommunen durchgeführt werden (vgl. unten Tabelle 2). Diese landesspezifischen Vergleichsprojekte, die in einigen Fällen (z.B. Bayern, Sachsen und Schleswig-Holstein) das Ergebnis von Erfahrungsaustausch und Konzepttransfer zwischen den kommunalen Landesverbänden sind, umfassen nach einer ersten Schätzung bereits knapp ein Drittel der deutschen Kreisverwaltungen. In den beiden untenstehenden Tabellen⁵ wird ein knapper Überblick über die wichtigsten überregionalen sowie länderspezifischen Projekte in der Bundesrepublik gegeben, die sich mit interkommunalen Leistungsvergleichen befassen.

Wie generell bei der Modernisierung des Öffentlichen Sektors in Deutschland, haben wir es auch bei der Etablierung von interkommunalen Leistungsvergleichen nicht mit einem zentralen federführenden Akteur, sondern mit einer Vielzahl von Akteuren und Arenen zu tun. Hierin findet die Vielgestaltigkeit des dezentral-fragmentierten Politik- und Verwaltungssystems der Bundesrepublik ihre Widerspiegelung (vgl. Wollmann/Schröter 1997: 188; Wollmann 2004a). Die Vergleichsprojekte werden – durchweg auf freiwilliger Basis - von unterschiedlichen Akteuren, zu unterschiedlichen Zeitpunkten, mit unterschiedlichen Instrumentierungen sowie teilweise gezielt in lokaler Eigenregie und Abgrenzung zu staatlichen Kontrollinstanzen (Fachaufsicht) ins Leben gerufen⁶. Sie stellen

³ Zwischen 1996 und 2002 sind innerhalb des IKO-Netzes insgesamt 127 Vergleichsringe gebildet worden und haben 651 Kommunen an Vergleichsringen teilgenommen, wobei diese Zahlen auch jene Vergleichsringe und Kommunen enthalten, die die Vergleichsarbeit nicht fortgesetzt haben (IKO-Netz der KGSt, Rundbrief 1/2003: 1).

⁴ Lediglich in den Bundesländern Hessen, Rheinland-Pfalz und Thüringen waren zum Erhebungszeitpunkt (Ende 2003) keine landesspezifischen Vergleichsprojekte außerhalb von KGSt und Bertelsmann Stiftung vorzufinden.

⁵ In den Tabellen nicht enthalten sind Vergleichsprojekte, die unter dem Dach von kommunalen Fachverbänden (z.B. Verband kommunaler Unternehmen, Verband kommunale Abfallwirtschaft, Deutscher Volkshochschulverband, Abwassertechnische Vereinigung) durchgeführt werden, ferner „reine“ Qualitätswettbewerbe bzw. Qualitätsvergleiche (bspw. nach dem CAF-Modell, vgl. Saatweber 2004). Auch sind nicht alle von den kommunalen Landesverbänden moderierten Vergleichsprojekte enthalten (z.B. Landkreistag Baden-Württemberg). Für ausführliche Quellennachweise siehe Kuhlmann 2004b.

⁶ Beispielsweise wurde vom Landkreistag in Sachsen-Anhalt das bewusst eigenständige und von der Landesregierung (auch finanziell) unabhängige Vorgehen im interkommunalen Leistungsvergleich damit

im Kern eine *Selbstorganisation der Kommunen* dar, in der sich die deutsche Politik- und Verfassungstradition der kommunalen Selbstverwaltung widerspiegelt. Diesem dezentral-fragmentierten Vorgehen können als markante Gegenbeispiele die zentralistisch veranlassten und weithin als staatliche Kontrollinstrumente genutzten Leistungsvergleiche unter den britischen Kommunen gegenüber gestellt werden, die seit den 80er Jahren unter Federführung der eigens zu diesem Zweck eingerichteten Nationalen Audit Commission durchgeführt werden (vgl. Bovaird/Löffler 2001; ferner Wollmann 2004a; Wegener 2004)⁷. Dagegen lassen sich in den skandinavischen Ländern, aber auch in den USA⁸ eher solche Leistungsvergleiche vorfinden, die dezentral initiiert und durchgeführt werden und damit eher Ähnlichkeiten mit den im deutschen Kontext gebräuchlichen, auf Freiwilligkeit und Selbstorganisation beruhenden bottom up-Strategien aufweisen (vgl. Strid 2004; Askim 2004).

begründet, dass man sich nicht von den Aufsichtsbehörden abhängig machen und kontrollieren lassen möchte (Interviewprotokoll, 16.12.2003).

⁷ Allerdings gibt es neuerdings auch unter den britischen Kommunen in Zusammenarbeit mit ihrem Verband „Local Government Association“ Versuche, dem zentralistischen Vorgehen der Regierung durch die Bildung eigener „benchmarking clubs“ entgegenzutreten (vgl. Wollmann 2004b).

⁸ Das 1995 gestartete US-amerikanische *Comparative Performance Measurement Program* wurde – zunächst 44 Städte und Counties mit jeweils über 200.000 Einwohnern umfassend – von der International City/County Management Association (ICMA), einer Selbstorganisation der Gemeinden, ins Leben gerufen. Seit 1997 nehmen auch eine Reihe mittlerer und kleinerer Gebietskörperschaften am Projekt teil (vgl. *International City/County Management Association – ICMA, About the ICMA Center für Performance Management*, Internet-Seite: <http://www.icma.org/abouticma/programs/performance/about.cfm>, 1999; zitiert nach Schuster 2001: 213).

Tabelle 1 : Bundesweite/ Länderübergreifende Projekte interkommunaler Leistungsvergleiche in Deutschland (Stand: 2003)

Vergleichsprojekt	Trägerinstitution/ Initiative	Start (Ende)	Zahl der Vergleichskommunen	Zahl d. Vergleichs- ringe	Zahl der Aufgaben- bereiche	Anzahl definierter Kennzahlen	Vergleichs- ebene
„Grundlagen einer leistungsfähigen Kommunalverwaltung“	Bertelsmann Stiftung; Deutscher Beamtenbund	1992 (1998)	„Kernring“ aus 5 Städten ⁹ ; weitere 150 Teilnehmer-kommunen	ca. 30	8	Bsp.: Kennzahlenset Sozialwesen mit 221 Kennzahlen (für 16 Produkte)	Produkte
IKO-Netz	KGSt	1996	Ca. 400 Kommunen (2001)	73 (2002)	25 (2002)	Bsp.: Kennzahlenset Sozialhilfe mit 80 Kennzahlen (für 2 Produkte)	Produkte
Kik	Bertelsmann Stiftung	2000 (2004)	Ca. 50 Kommunen	20 (2003)	9	Bsp.: Kennzahlenset Sozialhilfe mit 14 Kennzahlen	Produkte/ teils Wirkungen
kompass	Bertelsmann Stiftung	2000 (2004)	8 Kommunen	1 (2003)	5	Bsp.: „Erste Steuerungsebene“: 140 Kennzahlen in 6 Zielfeldern	Wirkungen
Benchmarking für soziale Dienstleistungen	Sozialämter/ überörtl. Sozialhilfe-träger und con_sens Consulting	1995/1999 ¹⁰	84 deutsche Kommunalverwaltungen ¹¹	5 (2003)	2 (Arbeitsmarkt- und Sozialpolitik)	Kennzahlenset „Hilfe zum Lebensunterhalt“ mit ca. 40 Kennzahlen (zzgl. Unterkennzahlen)	Produkte/ Wirkungen / teils Prozesse

⁹ Zwischenzeitlich waren noch 3 weitere Städte (Böblingen, Potsdam, Dessau) hinzugekommen, die aber bis 1997 alle den Vergleichsring wieder verließen.

¹⁰ Zunächst wurde das Vergleichsprojekt der 16 großen Großstädte unter der Moderation der Kienbaum Unternehmensberatung GmbH 1995 gegründet, welches 1999 von con_sens Consulting übernommen wurde. Seit diesem Zeitpunkt kamen weitere Benchmarking-Projekte hinzu.

¹¹ Hier ist auch der (landesspezifische) Benchmarking-Kreis der 18 Städte in NRW enthalten.

Tabelle 2 : Länderspezifische Projekte interkommunaler Leistungsvergleiche in Deutschland (Stand: 2003)

Bundesland	Vergleichsprojekt	Trägerinstitution/ Initiative/ Begleitung	Start (Ende)	Zahl der Vergleichs- kommunen	Zahl d. Vergleichs- ringe	Zahl der Aufgaben- bereiche
Baden- Württemberg	AG Produktkennzahlen/ Vergleichsringe Baden-Württemberg	Kommunen und Städtetag Baden-Württemberg	2002	Ca. 40 Kommunen	10 (2003)	10
Bayern	Innovationsring Bayerische Landkreise	Bayerischer Landkreistag (Startphase vom Freistaat Bayer bezuschusst)	1996	10 Landkreise	1	3 (7 Produkte)
Brandenburg	Vergleichsring „Kreisfreie Städte Brandenburg“	IKO-Netz der KGSt und Städte- und Gemeindenbund Brandenburg	2003	4 Kreisfreie Städte	1	8
Mecklenburg- Vorpommern	Controlling im Aktionsprogramm zur Integration von Sozialhilfeempfängern in den allgemeinen Arbeitsmarkt	Land MV (Versorgungsamt Rostock); Landkreise/ kreisfreie Städte; Trägergesellschaft des Landes mbH (Controlling-Durchführung)	2001	12 Landkreise, 6 kreisfreie Städte	1	1
Niedersachsen	(1) Niedersächsischer Kosten- und Leistungsvergleich	Niedersächsischer Städte- und Gemeindebund/ Gemeinden/ Diekwisch Consulting	1994	28 kleinere/ mittlere Gemeinden (2002)	1	12
	(2) Kennzahlenvergleiche der Niedersächsischen Landkreise	Landkreistag Niedersachsen	1999	37 Landkreise und Region Hannover	keine	36
Nordrhein- Westfalen	Nordrhein-Westfälischer Kosten- und Leistungsvergleich ¹²	Nordrhein-Westfälischer Städte- und Gemeindebund/ Gemeinden/ Diekwisch Consulting	1995-96	24 Gemeinden	1	12
Saarland	Interkommunale Vergleichsringe Saarland	Landkreistag Saarland	1999	6 Landkreise	6	6
Sachsen	Innovationsringe Sachsen	Sächsischer Landkreistag/ Sächsischer Städte- und Gemeindetag (80% Zuschuss vom Land)	2002	19 Kommunen	5	5
Sachsen- Anhalt	Vergleichsring Bauordnungsämter ¹³	Landkreistag Sachsen-Anhalt; Landkreise	2002-03	21 Landkreise	1	1 (6 Produkte)
Schleswig- Holstein	Innovationsring Schleswig-Holstein (Leistungsvergleich noch in Vorbereitung)	Landkreistag Schleswig-Holstein; Landkreise (Zuschuss vom Land)	2002	11 Landkreise	1	3

¹² Mangels Resonanz bei den Kommunen (nur 24 von 358 Gemeinden hatten sich beteiligt) wurde der Leistungsvergleich vom nordrhein-westfälischen Städte- und Gemeindebund in den Folgejahren nicht wieder angeboten.

¹³ Die vom Landkreistag Sachsen-Anhalt moderierten interkommunalen Vergleichsprojekte laufen jeweils ein Jahr und beziehen sich auf wechselnde Aufgabenfelder. So gab es in den Jahren vor 2002 bereits Vergleichsprojekte in den Bereichen KfZ und Kitas, an denen ebenfalls alle 21 Landkreise beteiligt waren.

Es wird deutlich, dass in Deutschland - ähnlich wie in den skandinavischen Ländern - den kommunalverbandlichen Akteuren eine besondere „Promotorenrolle“ bei der Etablierung von Leistungsvergleichen zukommt. Insbesondere die KGSt, aber auch die kommunalen Landesverbände und andere Non Profit Organisationen (vor allem die Bertelsmann-Stiftung; vgl. Tebbe 2004) haben die Verbreitung von Performanzindikatoren und Vergleichssystemen maßgeblich vorangetrieben. Darüber hinaus konnten sich auch zunehmend private Beratungsunternehmen im Bereich interkommunaler Leistungsvergleiche etablieren und sich entsprechende „Marktsegmente“ erschließen. Dagegen nehmen die Landesregierungen derzeit noch überwiegend eine lediglich (finanziell) unterstützende, zum Teil moderierende Funktion im Hinblick auf lokale Vergleichsprojekte wahr (z.B. Innenministerien von Baden-Württemberg und Sachsen). Neuerdings gibt es jedoch vermehrt Hinweise darauf, dass die Landesregierungen im Verhältnis zu den Kommunen zunehmend verwaltungspolitisch aktiv und auch gesetzgeberisch tätig werden, was etwa aus den Gesetzesänderungen im Haushaltsrecht, die sich in einigen Bundesländern (z.B. NRW, Hessen) in der Vorbereitung befinden, ablesbar ist. Diese im Zeichen der Haushaltskrise verfolgten Novellierungen könnten in den kommenden Jahren bewirken, dass der traditionelle Input-Haushalt durch den Produkthaushalt abgelöst wird (hierzu Budäus 2004) und sich die Kommunen auf dieser Grundlage zunehmend veranlasst, wenn nicht in die Pflicht genommen sehen, Leistungsvergleiche auf Produktbasis durchzuführen¹⁴.

3 Kosten und Nutzen von Leistungsvergleichen

Einerseits kann kein Zweifel daran bestehen, dass interkommunale Leistungsvergleiche eine Vielzahl von – wenn auch überwiegend kleinteilig-inkrementellen – Organisations- und Prozessveränderungen in den beteiligten Verwaltungen bewirkt oder zumindest begünstigt haben, die von der Schnittstellenoptimierung über die Verbesserung der IT-Ausstattung bis hin zur Umstellung der Personalplanung und Mitarbeiterschulung reichen (vgl. Korte 2004; Kuhlmann 2003a). Durch den interkommunalen Erfahrungsaustausch in den Vergleichsringen wurde auch die Verbreitung des Bürgeramtsmodells gefördert, was daraus ersichtlich wird, dass sich viele Kommunen, die am Leistungsvergleich mitwirkten, aufgrund der positiven Erfahrungen der anderen beteiligten Kommunen, ebenfalls entschlossen, Bürgerämter

¹⁴ So etwa die Prognose des (z. Zt. noch nicht im Leistungsvergleich aktiven) Thüringischen Gemeinde- und Städtebundes, wonach die bevorstehenden landesrechtlichen Normierungen im kommunalen Haushaltsrecht einen klaren Impuls für produktbezogene interkommunale Leistungsvergleiche mit sich bringen werden (Interviewprotokoll, 16.12.2003).

einzurichten¹⁵. Die Durchführung von Vergleichsprojekten hat zudem den Aufbau interkommunaler Netzwerke, die – neben indikatorengestützten („harten“) Leistungsvergleichen – einen qualitativen („weichen“) Erfahrungsaustausch zwischen den kommunalen Akteuren ermöglichen, begünstigt und vorangetrieben. Der Austausch von „Anwenderwissen“ über sektoral organisierte Netzwerke stellt einen wesentlichen, wenn nicht – aus Sicht der Akteure – überhaupt den wichtigsten Nutzen interkommunaler Leistungsvergleiche dar¹⁶. Gerade im Zusammenhang mit der aktuellen – unter dem Stichwort „Governance“ (vgl. Benz u.a. 2003; Klenk/Nullmeier 2003) firmierenden - Diskussion um neue Formen von Selbstorganisation durch Netzwerkbildung, die jenseits von Markt und Hierarchie angesiedelt sind, verdienen diese Bemühungen der Kommunen, sich zu Vergleichsringen und Netzwerken zusammenzuschließen als innovative Form von *local governance* besondere Beachtung (Heinelt 2003).

Andererseits gibt es eine Reihe von Hinweisen, die zu eher skeptischen Prognosen hinsichtlich der längerfristigen Aufrechterhaltung und Fortführung interkommunaler Vergleichsvorhaben Anlass geben. So wird das auf den ersten Blick ermutigende Bild des interkommunalen Leistungsvergleichs in Deutschland dadurch getrübt, dass die Bereitschaft der Kommunen, an Vergleichsprojekten teilzunehmen, bei weitem nicht flächendeckend ist. Auffällig ist vor allem, dass die ostdeutschen Kommunen entsprechenden Vorhaben eher zurückhaltend, wenn nicht ablehnend gegenüberstehen¹⁷. Des Weiteren deutet manches darauf hin, dass sich auch die Bereitschaft, den interkommunalen Leistungsvergleich längerfristig aufrechtzuerhalten und fortzuschreiben vielfach in Grenzen hält. Wiederum sind es vor allem die ostdeutschen Verwaltungen, die sich besonders häufig von den Vergleichsprojekten

¹⁵ Mit Blick auf das Bertelsmann-Grundlagen-Projekt war sogar von einer „regelrechten Welle von Neugründungen solcher Bürgerbüros“ (z.B. in Borken, Celle, Herford, Pforzheim, Werne) die Rede (Adamaschek 1997: 121). Mittlerweile verfügen in Nordrhein-Westfalen 85% der Städte ab 20.000 Einwohner und in Baden-Württemberg 78% der Städte über Bürgerämter (vgl. Bogumil u.a. 2003). Allerdings sei darauf hingewiesen, dass der interkommunale Leistungsvergleich die Etablierung von Bürgerbüros lediglich gefördert, nicht aber also solche ausgelöst hat.

¹⁶ Zu dieser Vermutung verhält sich auch die Feststellung konsistent, dass sich eine überwältigende Mehrheit (80%) der in einer IKO-Netz Kundenbefragung angeschriebenen Teilnehmer-Kommunen von der Mitarbeit in den Vergleichsringen eine Unterstützung beim interkommunalen *Erfahrungsaustausch* erwartet und in diesem das zentrale Teilnahmemotiv erblickt (siehe Korte 2004).

¹⁷ Während beispielsweise Anfang des Jahres 2000 noch immerhin 5% (=4 VR) aller (insgesamt 79) IKO-Netz Vergleichsringe von ostdeutschen Kommunen gebildet worden waren, sank ihr Anteil bis Ende 2002 auf 3%, was einer Anzahl von nur noch 2 Vergleichsringen, die im Land Brandenburg ansässig waren, entspricht (vgl. Kuhlmann 2003a m.w.N). Von den Ende 2001 insgesamt 672 am IKO-Netz beteiligten Kommunen (einschließlich Doppelmitgliedschaften) waren ganze 94% aus Westdeutschland (=633 Kommunen), wohingegen nur 39 Kommunen, also 6%, ostdeutscher Herkunft waren (ebd.).

verabschieden¹⁸. Nicht nur im IKO-Netz der KGSt scheint die Quote der „Aussteiger“, die nach Ablauf der „Probezeit“ die Projekte verlassen, beachtlich zu sein, sondern auch die von den kommunalen Landesverbänden ins Leben gerufenen Vergleichsprojekte (siehe Tabelle 2) treffen auf zum Teil erhebliche Mobilisierungsprobleme, und es ist vielerorts bereits vom „Einschlafen“ des interkommunalen Leistungsvergleichs die Rede¹⁹.

Lassen sich somit einerseits eine Reihe von Erfolgen und positiven Effekten interkommunaler Leistungsvergleiche benennen, werden andererseits die Schwachstellen und Kosten solcher Vorhaben immer sichtbarer. Vor diesem Hintergrund soll im Folgenden versucht werden, eine Art Kosten-Nutzen-Bilanz des bisherigen interkommunalen Leistungsvergleichs aufzustellen, indem anhand vorliegender empirischer Hinweise die tatsächlichen Auswirkungen von Vergleichsprojekten auf das Verwaltungshandeln (*performance*) abgeschätzt und den – soweit bekannt - aufgewendeten Kosten gegenübergestellt werden. Dabei werden in Anlehnung an einschlägige Typologien die folgenden „Performanzfelder“ unterschieden (vgl. Pollitt/Bouckaert 2000; Bogumil/Kuhlmann 2004; Kuhlmann 2004a):

- Input-Veränderungen: Kosteneinsparungen, Effizienzgewinne
- Output- und Prozessveränderungen: Servicequalität, Kundenfreundlichkeit
- Outcomes und System-Veränderungen: politische Steuerung, Mitarbeitermotivation

3.1 Input-Veränderungen

In der Frage, inwieweit interkommunale Leistungsvergleiche zu Kosteneinsparungen und Wirtschaftlichkeitsgewinnen in den beteiligten Verwaltungen geführt haben, ergibt sich ein ambivalentes Bild. Einerseits wird vielfach konstatiert, dass das Kostenbewusstsein in den Fachämtern deutlich gestiegen sei, was vor allem darauf zurückgeführt wird, dass Kosten und Leistungen transparenter sind und auf der Basis vorliegender Vergleichswerte überhaupt erst sinnvoll bewertet werden können. Auch gibt es einige Hinweise darauf, dass die am interkommunalen Leistungsvergleich beteiligten Kommunen gewisse Kostenersparnisse oder

¹⁸ Beispielhaft hier der Austritt von 13 thüringischen Landkreisen und 4 brandenburgischen Städten aus dem IKO-Netz. Zwischen März 2000 und Dezember 2001 reduzierte sich in mehr als der Hälfte der im IKO-Netz vertretenen kommunalen Aufgabenbereiche (19 von 34) die Zahl der teilnehmenden Kommunen, so dass es in dieser Zeitspanne zu einem Gesamtrückgang von 11% (84 Kommunen) kam (Kuhlmann 2003a m.w.N.).

¹⁹ So wörtlich die Einschätzung des Geschäftsführenden Direktors beim Landkreistag Rheinland-Pfalz (Interviewprotokoll, 16.12.2003). Ähnlich auch die Einschätzung des Hessischen Städtetages, nach dessen Auskunft der überwiegende Teil der Städte, die sich noch bis Ende der 90er Jahre in Vergleichsringen engagiert hatten, mittlerweile aus den entsprechenden Projekten ausgetreten ist, so dass die nach wie vor im IKO-Netz aktiven hessischen Kommunen eher Ausnahmefälle darstellen (Interviewprotokoll, 16.12.2003). Auch in Nordrhein-Westfalen wurde das vom Städte- und Gemeindebund (unter Mitwirkung von Diekwisch

Einnahmezuwächse zu verzeichnen haben, etwa weil bestimmte „kostensparende“ bzw. „einnahmewirksame“ Organisationsmaßnahmen implementiert oder aus anderen Vergleichskommunen übernommen wurden²⁰. Andererseits fällt bei näherem Hinsehen auf, dass die konstatierten „Gewinne“ in den Teilnehmerkommunen häufig aus einmaligen, punktuell anfallenden Mehreinnahmen resultieren, die kaum geeignet scheinen, eine längerfristige und nachhaltige Steigerung der Wirtschaftlichkeit des Verwaltungshandelns auszulösen. Auch sei daran erinnert, dass bislang, soweit ersichtlich, nicht der Versuch unternommen worden ist, die mit der Konzipierung, Implementation, Fortschreibung und Evaluation von Leistungsvergleichen verbundenen Transaktions- und Opportunitätskosten²¹ (vgl. Abschnitt 3.4) zu messen und ins Verhältnis zum erreichten Nutzen zu setzen. Es kann plausibel argumentiert werden, dass eine solche Analyse die reformbezogene Kosten-Nutzen-Bilanz, vor allem in Hinsicht auf Effizienzkriterien, erheblich verschieben, resp. verschlechtern würde.

3.2 Output- und Prozessveränderungen

Zum einen lässt sich zeigen, dass es bei den an interkommunalen Leistungsvergleichen beteiligten Kommunalverwaltungen zu messbaren Verbesserungen im Serviceniveau gekommen ist (vgl. Schuster 2001: 216 ff.), die partiell auf die Anreizwirkung (Wettbewerbssurrogat) der Vergleichsprojekte zurückzuführen sind. So waren in den fünf Städten des sog. „Kernrings“ der Einwohnermeldeämter im Bertelsmann-Projekt (siehe Tabelle 1) sichtbare Niveauverbesserungen in den Leistungsparametern „Öffnungszeiten“ und „Wartezeiten“ zu verzeichnen. Zwischen 1995 und 1997 waren die Wartezeiten in den beteiligten Kommunen dieses „Kernringes“ durchweg rückläufig, und 1997, also kurz vor Abschluss des Projekts, gab es in keiner Kommune mehr Wartezeiten von über 10 Minuten²². Andererseits zeigte sich, dass Städte mit mittlerer bis guter Platzierung eher geneigt sind, bei ihrem bisherigen Serviceniveau zu bleiben (Bsp. Castrop-Rauxel, Gütersloh), so dass mithin

Consulting) koordinierte interkommunale Vergleichsprojekt mangels Interesses auf Seiten der Städte und Gemeinden 1996 eingestellt (Interviewprotokoll, 17.12.2003).

²⁰ Als Beispiele aus dem „Bertelsmann-Vergleichsprojekt“ wären etwa zu nennen: Mehreinnahmen von jährlich 5.600 DM in der Stadt Pforzheim durch Aufstellung eines Passbildautomaten im Bürgerzentrum; Einsparung von zwei Stellen durch Auflösung des Schreibbüros im Landkreis Osnabrück; Reduzierung des Unterhaltungsaufwandes bei Spielplätzen in der Stadt Gütersloh durch Festsetzung einer Mindestgröße bei Spielplätzen und Verkauf der kleinen Grundstücke von „Minispielplätzen“ (vgl. Bertelsmann Stiftung <http://www.leistungsvergleich.de/info/2.htm>; Recherche vom 9.7.2002).

²¹ Opportunitätskosten fallen dadurch an, dass die Verwaltungsbediensteten für den Leistungsvergleich Arbeitszeit aufwenden müssen, während der sie nicht für andere Aufgaben zur Verfügung stehen (vgl. Pollitt 1995: 139).

²² Siehe auch die vielfältigen Städtebeispiele bei Adamaschek (1998) und unter <http://www.leistungsvergleich.de/info/2.htm>.

eine beachtliche „Unbeweglichkeit im Mittelfeld“ (ebd.: 220) festzustellen war. Im Bertelsmann-Grundlagen-Projekt wurde phasenweise – etwa im Bereich der Wartezeiten – sogar eine Verschlechterung des Leistungsniveaus beobachtet und zudem konstatiert, dass die „objektiven“ Serviceverbesserungen zum Teil nicht mit der subjektiven Kundenzufriedenheit korrelierten²³. Darüber hinaus gibt es Hinweise darauf, dass Wirtschaftlichkeitsgewinne und Einsparungen (etwa durch Stellenabbau) häufig auf Kosten der Leistungsqualität „erkauft“ werden²⁴. Dies dürfte nicht zuletzt damit zusammenhängen, dass sich der Leistungsvergleich aufgrund des Bestrebens der lokalen Akteure, die aktuelle Verwaltungsmodernisierung für Haushaltseinsparungen zu nutzen, überwiegend auf quantifizierbare und monetarisierbare Größen, wie z.B. Bearbeitungsdauer, bezieht, wohingegen inhaltliche, policy-spezifische Qualitätsaspekte ausgeblendet bleiben. Mit Blick auf die rechtsstaatlich noch wenig geübten ostdeutschen Verwaltungen scheint zudem der Umstand problematisch, dass die rechtliche Qualität des Verwaltungshandelns als Performanzkriterium (rechtliche Korrektheit der Entscheidungen, „Prozessanfälligkeit“ usw.) im Zusammenhang mit interkommunalen Leistungsvergleichen kaum thematisiert wird (vgl. Kuhlmann 2003b; Kuhlmann 2004c).

3.3 Outcomes und System-Veränderungen

Die „Erfolgsbilanz“ des interkommunalen Leistungsvergleichs im Hinblick auf die Forderung, Kosten-, Leistungs- und Wirkungsinformationen für eine verbesserte (evidenzbasierte) politisch-strategische Steuerung nutzbar zu machen, muss in mehrfacher Hinsicht besonders kritisch ausfallen. Zum einen fällt auf, dass Leistungsvergleiche fast ausschließlich aus der Verwaltung heraus initiiert und von der Verwaltung durchgeführt werden. Auch die Nutzung von Vergleichsinformationen erfolgt überwiegend binnenadministrativ. Nennenswerte Versuche, neben den Verwaltungsakteuren auch die Kommunalpolitiker oder gar die lokale Bürgerschaft und gesellschaftliche Akteure in Vergleichsprojekte einzubeziehen, sind lediglich in Anfängen zu verzeichnen und zudem mit erheblichen Mobilisierungsproblemen verbunden²⁵. Zum andern ist festzustellen, dass sich kommunale

²³ So verschlechterte sich in Pforzheim im Zeitraum 1994 bis 1996 die Kundenzufriedenheit ungeachtet dessen, dass in dieser Zeit die Öffnungszeiten von 26,5 auf 28,5 Wochenstunden erweitert worden waren (Schuster 2003).

²⁴ Beispielhaft hierfür sei die Stadt Kamenz erwähnt, die zur Verringerung des Zeitaufwands bei Eheschließungen den personellen Betreuungsaufwand verringerte und auf die Nutzung von Musterreden bei Trauungen zurückgegriffen hat (vgl. Adamaschek 1998: 102).

²⁵ Das von der Bertelsmann Stiftung im Jahr 2000 initiierte Projekt Kompass (vgl. Tabelle 1), in dem es darum ging, den Erfolg und die Auswirkungen kommunalpolitischen Handelns auf die „Lebensqualität“ in der Kommune zu bestimmen, bildete hier ein Ausnahmebeispiel. In diesem Projekt werden die Kommunalpolitik und lokale Bürgerschaft - sei es durch Befragungen, sei es auch durch institutionelle Einbindung - in den Prozess der (strategischen) Zielbildung und Indikatorenentwicklung einbezogen.

Leistungsmessungen und -vergleiche derzeit fast ausschließlich auf der Ebene von „Produkten“ (Outputs) bzw. der für deren Erstellung erforderlichen Ressourcen (Inputs) bewegen, wohingegen die aus Sicht der Kommunalpolitiker interessanten Wirkungen des Verwaltungshandelns (Outcomes, z.B. Nachhaltigkeit, Verteilungsgerechtigkeit, Lebensqualität usw.) kaum eine Rolle spielen (vgl. Brüggemeier 2004; Reichard 2004). Vor diesem Hintergrund bieten Vergleichsdaten bislang nur wenig Anknüpfungspunkte für Entscheidungen im Rahmen eines strategischen Controllings oder einer wirkungsorientierten Steuerung, da diese die Messung von Wirkungen anhand von Outcome-Indikatoren voraussetzen. Wie aus der Evaluationsforschung bekannt, liegt dieses Defizit nicht zuletzt darin begründet, dass Wirkungsanalysen auf besondere methodische Konzipierungs-, Zuordnungs- und Kausalitätsprobleme stoßen (vgl. Wollmann 2000), da Outcome-Werte zumeist „multikausale“ Größen sind, bei deren Zustandekommen der kommunale Beitrag nicht eindeutig festzustellen ist. Da der Ausbau der institutionellen Architektur von Berichtswesen und Controlling damit vor allem der Stärkung der inneradministrativen Informations- und Steuerungskapazität gedient hat, während die Kommunalvertretungen davon weitgehend abgeschnitten blieben (Wollmann 2001: 33), hat sich im Ergebnis die Handlungsschwäche der Kommunalvertretung gegenüber der Verwaltung eher verschärft und haben die (politisch-parlamentarischen) Steuerungsdefizite noch zugenommen.

Nicht-intendierte Effekte und „Performanzverschlechterungen“ hat der interkommunale Leistungsvergleich auch im Bewertungskriterium der Mitarbeiterzufriedenheit hervorgebracht. Zunehmend werden kritische Stimmen laut, die insbesondere auf gravierende Motivationseinbrüche bei den Mitarbeitern aufmerksam machen²⁶. Da mit den Kennzahlenmessungen und Fortschreibungen von Vergleichsdaten eine erhebliche, zusätzlich zum Tagesgeschäft zu leistende Mehrarbeit anfällt, ist die Motivation der Beschäftigten und ihre Bereitschaft, entsprechende Vorhaben zu unterstützen, vor allem in denjenigen kommunalen Fachämtern markant rückläufig, deren Arbeitsvolumen ohnedies schon die verfügbaren knappen Ressourcen übersteigt (etwa Sozialämter). Dieser Motivationsverlust wird dadurch noch zusätzlich befördert, dass sich die einzelnen Mitarbeiter durch die Aufschreibungen kontrolliert fühlen (Schuster 2001: 217) und mithin von „formalem Hierarchie-Druck“ die Rede ist, da die Arbeit in Vergleichsprojekten vielfach intern top-down gesteuert und auf eine breite Einbindung und Informierung der Mitarbeiter, die die

²⁶ Siehe zur Problematik der Beschäftigtenbeteiligung auch Herbing 2004.

Leistungserhebungen durchführen müssen, verzichtet wird (Anke u.a. 1999: 43)²⁷. Vor diesem Hintergrund verwundert es wenig, dass vielerorts eher eine rückläufige denn steigende Mitarbeiterzufriedenheit festzustellen ist²⁸, zumal auch eine Reihe von Maßnahmen, die auf Effizienzsteigerung (z.B. durch Personaleinsparungen) und Erhöhung der Kundenzufriedenheit (z.B. durch Verlängerung der Öffnungszeiten) gerichtet sind, fast naturgemäß in einem Spannungsverhältnis zum Ziel der Mitarbeiterorientierung steht.

3.4 Implementations- und Opportunitätskosten interkommunaler Leistungsvergleiche

Angesichts dessen, dass interkommunale Leistungsvergleiche darauf zielen (sollten), die Leistungsfähigkeit (*performance*) lokaler Politik und Verwaltung zu verbessern, muss zunächst der Befund überraschen, dass sich nur wenige der an entsprechenden Projekten beteiligten Kommunen tatsächlich aus dem Leistungsvergleich „Performanzverbesserungen“ versprechen²⁹. Insbesondere diejenigen Verwaltungen, die mittlerweile aus den Vergleichsprojekten ausgestiegen sind, monieren, dass bislang signifikante Veränderungen und ein sichtbarer Nutzen der Kennzahlenvergleiche auf der Ebene des praktischen Verwaltungshandelns kaum zu erkennen seien, so dass der damit verbundene Aufwand als nicht gerechtfertigt erscheint.

Aus einer institutionenökonomischen Sicht dürfte die mancherorts begrenzte Motivation zur Teilnahme am interkommunalen Leistungsvergleich, aber vor allem auch die im Verlauf der Vergleichsprojekte häufig schwindende Bereitschaft zur dauerhaften Fortführung der Kennzahlenmessungen daraus zu erklären sein, dass die mit der Konzipierung, Implementierung und längerfristigen Fortschreibung der Indikatorensysteme und Vergleichsdaten verbundenen Kosten den – noch weitgehend als ungesichert erscheinenden – Gewinn dieser Unternehmungen übersteigen. Kurz gesagt, ist es aus Sicht vieler kommunaler Akteure zu einer deutlichen Schieflage im Kosten-Nutzen-Verhältnis interkommunaler

²⁷ In der Stadtverwaltung Vauberg wurden im Ergebnis einer zweijährigen Projektlaufzeit (Bertelsmann-Grundlagen-Projekt) „tiefe Enttäuschung“, „Burn out-Symptome“ und durchweg negative Erwartungen der Mitarbeiter bezüglich einer Weiterführung des Leistungsvergleichs festgestellt (Anke u.a. 1999: 43).

²⁸ Beispielhaft hier die „Bertelsmann-Kommune“ Gütersloh, in der sich die anfänglich hohe Mitarbeiterzufriedenheit (Schulnote 2,08 im Jahr 1993) im Laufe des Vergleichsprojekts drastisch verschlechterte (Schulnote 4,6 im Jahr 1997, vgl. Schuster 2003).

²⁹ Ausweislich einer IKO-Netz-Kundenbefragung aus dem Jahr 2003, in der 697 teilnehmende Kommunen angeschrieben wurden (Rücklaufquote 51,5%) versprechen sich nur 20% der „IKO-Netz-Kunden“ Verbesserungen in der Wirtschaftlichkeit, 8% eine Steigerung der Bürgerzufriedenheit, 12% eine Verbesserung der Organisation und gerade 3% eine Erhöhung der Mitarbeiterzufriedenheit (vgl. Korte 2004).

Leistungsvergleiche gekommen, die vielfach den Anlass dazu gegeben hat, sich von den Vergleichsprojekten zu verabschieden³⁰.

Der „Löwenanteil“ an den Gesamtaufwendungen, die für Leistungsmessungen und -vergleiche erforderlich sind, entfällt ohne Zweifel auf die Personalkosten³¹, die sich aus den arbeitszeitlichen Aufwendungen der Verwaltungsbeschäftigten für die Konzipierung und Implementierung der Kennzahlensysteme, für arbeitstägliche Aufschreibungen zu Bearbeitungs- und Durchlaufzeiten sowie für das Führen von (häufig manuellen) Strichlisten zur Zahl bearbeiteter Anträge oder antragsbezogener Verrichtungen usw. ergeben³². Aus Sicht der Kommunen erscheinen dabei, abgesehen von den zunächst einmalig anfallenden Einführungs- und Implementationskosten, vor allem diejenigen Aufwendungen als problematisch, die dauerhaft, d.h. bei einer längerfristigen Unterhaltung und Aktualisierung der Datenbestände und laufenden (teils unterjährigen) Berichterstattung entstehen und die auf längere Sicht personelle und zeitliche Kapazitäten binden. Zwar wird in diesem Zusammenhang von einigen Projektträgern argumentiert, dass sich der zu leistende Aufwand durch technische Unterstützung und „Digitalisierung“ beträchtlich vermindern werde und ohnedies als Bestandteil eines später mit der Einführung von Kosten- und Leistungsrechnung, Berichtswesen und Controlling zu erbringenden Arbeitsaufwandes zu sehen sei (Korte/Pook 2002: 17). Diese Argumentation lässt aber zum einen außer Acht, dass die technische Unterstützung von Kennzahlenmessungen in den Verwaltungen noch bei Weitem keine Selbstverständlichkeit ist und von einer automatisierten Fortschreibung wohl strukturierter Datenbestände „per Mausclick“ in vielen Kommunen noch keine Rede sein kann. Zum andern wird übersehen, dass der interkommunale Leistungsvergleich eben nur zum Teil – und nach Einschätzung vieler Kommunalakteure in einem viel zu geringen Ausmaße – für das interne Controlling der einzelnen Verwaltungen genutzt werden kann, also in der Regel einen beachtlichen *zusätzlichen* Aufwand darstellt³³. Gerade den ostdeutschen Kommunen, die

³⁰ Ein Projektteilnehmer des IKO-Netzes formulierte dieses Problem zugespitzt mit den Worten: „*Es war viel Aufwand und wir haben nichts geschafft*“ (Interviewprotokoll, 15.1.2003).

³¹ Außerdem sind die Beraterhonorare (Bsp. IKO-Netz: 1.500 bis 5.000 EUR für bis zu 2 Jahre Projektlaufzeit) sowie die Kosten für neue IT-Anschaffungen (Bsp. IKO-Netz: jährlich 200 bis 1.300 EUR) und Kommunikationslizenzen (Bsp. IKO-Netz: jährlich 200 bis 1.300 EUR) zu berücksichtigen, wobei Letztere im IKO-Netz den KGSt-Mitgliedskommunen mittlerweile erlassen werden (vgl. Kuhlmann 2003a m.w.N.)

³² Beispielhaft geschätzt durch eine befragte Fachbereichsleiterin einer mittelgroßen Stadt (27.000 Einwohner) mit einem zehn Mitarbeiter umfassenden Sozialamt, das sich über zwei Jahre am Leistungsvergleich der Sozialämter im KGSt IKO-Netz beteiligte, liegen diese „Opportunitätskosten“ (berechnet für die zweijährige Projektdauer) bei ca. 45.500 EUR, d.h. bei zwei Jahresgehältern für eine Vollzeitstelle im mittleren bis gehobenen Verwaltungsdienst (Interviewprotokoll, 10.4.2002).

³³ So die ausgesprochen kritische Einschätzung von Talkenberg (Interviewprotokoll, 19.3.1996), der seinerzeit Projektverantwortlicher im Bertelsmann-Grundlagen-Projekt in der Stadtverwaltung Potsdam war; ähnlich aber auch die Einschätzungen der Fachbereichsleiterin Soziales in der Stadt Hennigsdorf mit Blick auf den

zusätzliche Ressourcen und Handlungsspielräume (*organizational slack*) für die Vergleichsarbeit kaum mobilisieren können, erscheint der interkommunale Leistungsvergleich vor diesem Hintergrund häufig eher als ein exotisches Instrument, für das im Verwaltungsalltag kein Platz ist.

Unverhältnismäßig hohe Opportunitätskosten (vgl. FN 21) hängen nicht zuletzt auch mit der beachtlichen Menge von Kennzahlen und Subkennzahlen, dem hohen Detaillierungsgrad und der tiefgegliederten Struktur der Indikatorensysteme zusammen, die charakteristisch für die „frühen“ Kennzahlensets der Bertelsmann Stiftung und der KGSt waren (vgl. oben Tabelle 1). Spiegelt sich darin einerseits das Bemühen der Akteure wider, ein möglichst vollständiges Bild kommunalen Handelns mittels Kennzahlen wiederzugeben, erwiesen sich diese Kennzahlensets andererseits für Steuerungszwecke in der Verwaltung kaum als geeignet, geschweige denn taugten sie als Informationsbasis für politische Entscheidungen. So ist man mittlerweile in einigen Vergleichsprojekten auch dazu übergegangen, anstelle einer Myriade von steuerungsungeeigneten Kennzahlen und Subkennzahlen nur noch einige aussagefähige und steuerungsrelevante „Kernkennzahlen“ zu definieren und eher einfache „abgespeckte“ Kennzahlensets zu verwenden (z. B. kik-Projekt; vgl. Tebbe 2004; Reichard 2004)³⁴, um damit den Erhebungsaufwand zu reduzieren und den Steuerungsnutzen zu erhöhen.

Als letzter Punkt sei noch erwähnt, dass auch interne und externe Qualitätskontrollen von Kennzahlenmessungen und Vergleichsdaten die Transaktionskosten (hier im Sinne von Kontrollkosten) der Vergleichsaktivitäten in die Höhe treiben. Aufgrund dessen, dass der Prozess der Kennzahlenmessung und -überprüfung bisweilen für nicht gesichert gehalten und Zweifel an der Datenqualität geäußert wurden (Adamaschek 1997), da die Daten oft in Eile von Einzelpersonen zusammengetragen, scheinbar gleiche Kennzahlen vollkommen unterschiedlich erhoben und auffällige Nachlässigkeiten bei der Datenerhebung festgestellt wurden (Schuster 2003), ging man etwa im Bertelsmann-Grundlagen-Projekt zeitweise dazu über, die Qualität der Daten und Datenerhebung – als eine Art Meta-Performance Measurement – zu prüfen. Jedoch scheint diese Strategie der Daten(qualitäts)prüfung mit Blick auf die sich dadurch noch einmal potenzierenden Transaktions-, Opportunitäts- und Kontrollkosten als ausgesprochen prekär.

IKO-Netz-Vergleichsring „Sozialämter in Brandenburg“ (Interviewprotokoll, 10.4.2002) und des Projektverantwortlichen der Stadt Düsseldorf im IKO-Netz Vergleichsring „Stadtplanung Großstädte“ (Interviewprotokoll, 15.1.2003).

³⁴ Während im Bertelsmann-Grundlagen-Projekt im Bereich des Sozialwesens noch mit 221 Kennzahlen operiert worden ist, stützte sich das Bertelsmann-kik-Projekt zur Leistungsmessung in der Sozialhilfe nur noch auf 14 Schlüsselkennzahlen (siehe Tabelle 1).

4 Zusammenfassung und Ausblick: zur politischen Rationalität von Selbstevaluation

Der interkommunale Leistungsvergleich in Deutschland hat in einer Vielzahl von Kommunen den Anstoß dafür gegeben, Verfahren und Instrumente zur laufenden Beobachtung und Selbstevaluation des Verwaltungshandelns zu etablieren, durch die die Entwicklung, Institutionalisierung und Verbreitung von Leistungsindikatoren wesentlich gefördert worden ist³⁵. Zahlreiche kommunale Akteure haben sich im Verlauf der Vergleichsprojekte mit Methoden und Verfahren strukturierter Datenerhebung und -auswertung vertraut machen können, die nunmehr einen Platz im „institutionellen Gedächtnis“ der Verwaltungen haben. Vor dem Hintergrund der vielfach, nicht zuletzt von Seiten der Wissenschaft, geäußerten Kritik am „Evaluationsdefizit“ in der öffentlichen Verwaltung können diese Aktivitäten der Kommunen durchaus als ein bemerkenswerter Schritt in Richtung auf die Institutionalisierung einer „Evaluationsfunktion“ angesehen werden (Wollmann 2000: 212ff.).

Andererseits deuten die wachsenden Umsetzungsprobleme und „Ermüdungserscheinungen“ bei Vergleichsprojekten³⁶, die wesentlich mit den hohen Implementations- und Opportunitätskosten bei weithin ungesichertem Nutzen zusammenhängen, auch auf die Grenzen dieses Reformansatzes hin. Hinzu kommt, dass neben methodisch-konzeptionellen und durchführungsbezogenen „Fallstricken“ von Leistungsmessungen und -vergleichen die Handlungsrationalität in konkurrenzdemokratisch verfassten Systemen berücksichtigt werden muss, die für Leistungs- und Wirkungsmessungen in Politik und Verwaltung einige Probleme aufwirft. Selbstevaluationen im öffentlichen Sektor müssen sich der politischen Frage stellen, welche Akteure, unter welchen Bedingungen und aus welchen Motiven heraus ein Interesse an Erfolgskontrollen und Wirkungstransparenz haben könnten. Bei der Beantwortung dieser Frage stehen sich bis heute zwei widerstreitende Positionen gegenüber:

Aus einer (eher „pessimistischen“) Perspektive, die – den Prämissen der ökonomischen Theorie der Politik folgend – das Streben der politischen Akteure nach *Stimmenmaximierung* und *Machterhalt* in den Mittelpunkt rückt, kann es im politischen Raum kaum ein Interesse an „objektiven“ Erfolgskontrollen geben, da diese das Risiko mit sich bringen, dass nach außen hin sichtbar wird, um wieviel und aus welchen Gründen ein gesetztes Ziel verfehlt worden ist

³⁵ Im IKO-Netz der KGSt lagen Anfang 2003 über 50 Kennzahlensysteme zu den verschiedensten kommunalen Aufgabenbereichen vor. Vergleichswerte auf der Basis der Kennzahlensysteme waren zu Beginn des Jahres 2002 für mehr als 10 IKO-Netz-Vergleichsringe über die IKON-Vergleichsdatenbank abrufbar.

³⁶ Vor diesem Hintergrund wurde bisweilen auch die Forderung erhoben, Leistungsvergleiche und die Veröffentlichung von Vergleichsergebnissen zukünftig rechtsverbindlich vorzuschreiben, etwa durch entsprechende Regelungen in den Kommunalverfassungen (vgl. Banner 2001: 63ff.).

(Bogumil 2003: 22). Dieser Position zufolge handeln Politiker, vor allem unter wettbewerbsdemokratischen Bedingungen, ähnlich wie Unternehmen und sind in erster Linie an „Marktsicherung“ interessiert, während inhaltliche Politikziele oder eine Orientierung am Gemeinwohl nachrangig sind. Einer solchen „ökonomischen“ Handlungsrationalität entspricht es auch, politische Ziele allgemein und unpräzise zu formulieren, um sich Spielräume für politische Manöver zu erhalten und dem politischen Gegner gegenüber möglichst wenig Angriffsfläche zu bieten. Die Chancen, vermittels Leistungs- und Wirkungsmessungen zu einem präzisen Bild des Zielerreichungs- und Wirkungsgrades politisch-administrativen Handelns zu gelangen, erscheinen vor diesem Hintergrund eher begrenzt, was sich in den beschriebenen Mobilisierungs- und Implementationsproblemen, auf die entsprechende Initiativen im kommunalen Raum treffen, widerspiegelt. Dass sich Kennzahlenvergleiche und Leistungsmessungen dennoch in vielen Bereichen als Reforminstrumente etabliert haben, erklärt sich dieser Argumentation zufolge vor allem aus dem Interesse der Politikakteure nach vorzeigbaren Politikerfolgen (Reichard 2001: 84) und daraus, dass sie als „Kennzahlenshow“ (vgl. Nullmeier 2004) genutzt werden können, um bestimmte Politikstrategien öffentlichkeitswirksam zu rechtfertigen und symbolische Politik zu betreiben. Soweit jedoch schlechte „Erfolgsbilanzen“ in Aussicht stehen, die die Wiederwahl gefährden könnten, ist den lokalen Mandatsträgern, insbesondere der Mehrheitsfraktionen, und leitenden Beamten aus nachvollziehbaren Gründen kaum an Leistungs- und Wirkungstransparenz gelegen (Banner 2001)³⁷.

In diese Argumentation fügt sich auch die Beobachtung ein, dass politisches Handeln nicht strikt ziel- und wirkungsorientiert, strategisch vorausschauend und nach einer rational-umfassenden Methode mit klaren Zielen, Bewertungskriterien und Ergebnistransparenz erfolgt, sondern eher einem iterativen, schrittweise probierenden Vorgehen des „Sich Durchwurstelns“ entspricht (Lindblom 1975: 162), das von den (kontingenten) Handlungskalkülen der Akteure abhängig ist. Anstelle einheitlicher Bewertungsmaßstäbe können die Auffassungen relevanter Akteure über Ziele und Bewertungskriterien stark divergieren³⁸; die Zielauswahl erfolgt oft erst im administrativen Entscheidungsprozess

³⁷ „(...) the politician’s prime goal is to be re-elected rather than to respect technical evidence“ (Cook zitiert nach Sanderson 2002).

³⁸ Es sei daran erinnert, dass die Bewertungskriterien öffentlichen Handelns, nach denen dieses „gemessen“ und evaluiert werden kann, vielfältig und widerstreitend sind, da im Bereich der öffentlichen Aufgabenerfüllung – im Gegensatz zum privaten Unternehmen – die negativen Folgekosten des Wirtschaftens nicht externalisiert, sondern abgefangen, verarbeitet und internalisiert werden. Für die öffentliche Leistungserbringung bedeutet dies, dass etwa das Bewertungskriterium hoher Effektivität (beispielsweise bei der Prävention von Jugendkriminalität) im Konflikt mit der Forderung nach einem möglichst wirtschaftlichen

(Stichwort: Behörde als „Clearingstelle“), das Bestreben, politischen Konsens zu finden, ist häufig wichtiger, als klare politische Ziele zu formulieren; ad hoc-Anpassungen an situative Problemlagen sind aus Sicht der Politikakteure in der Regel „rationaler“ als langfristige Planungen usw. usf. Das teilweise zu beobachtende Scheitern von Leistungs- und Wirkungsmessungen erklärt sich vor diesem Hintergrund aus den Umsetzungsproblemen, auf die das „rationale Politikmodell“ mit seinem Anspruch auf „Evidenzbasierung“, Planung und Steuerung stößt³⁹.

Dagegen lehnt sich eine andere – eher „optimistische“ – Position stärker an Argumente der vergleichenden Policy-Forschung, insbesondere der Parteiendifferenztheorie, an und erklärt ein grundsätzliches Interesse der Politikakteure an Wirkungskontrollen aus ihrem Streben nach inhaltlichen (Policy-)Problemlösungen und Programmumsetzung (Schmidt 1995: 580 ff.). Aus dieser Sicht haben die Politikakteure durchaus ein Interesse an Ergebnistransparenz vor allem deshalb, weil sie ihr Handeln (auch) mit programmbezogenen Verpflichtungen und der Durchsetzung *inhaltlicher Politikziele* verbinden, die ja gerade eine unerlässliche Voraussetzung für die Wiederwahl sind. Denn erst, wenn das Tun und Lassen eines Amtsinhabers den Erwartungen seiner Klientel näherungsweise gerecht wird, scheint seine Wiederwahl gesichert. Eine solche Orientierung an bestimmten „Policy-Werten“ und inhaltlichen Politikzielen vorausgesetzt, dürften die politischen Amtsträger ein ureigenes Interesse an Informationen über den Zielerreichungs- und Wirkungsgrad ihres Handelns hegen, nicht zuletzt, um Programmentscheidungen ggf. rechtzeitig korrigieren zu können. Diese Politik- und Programmorientierung vorausgesetzt, können sich Leistungsmessungen und -vergleiche gerade im Zuge von Cutback Management und Haushaltskonsolidierung als Alternativen zum „Rasenmäherprinzip“ und als geeignete Instrumente erweisen, um die Ressourcenallokation innerhalb des gegebenen Budgets zu rationalisieren und „evidenzbasiert“ (vgl. Sanderson 2002) zu legitimierbaren Budgetkürzungen zu gelangen (Derlien 1997: 7). Mit anderen Worten: „Programme, die gute Resultate hervorbringen, werden ausgedehnt; diejenigen, die sich als unfruchtbar erweisen, werden fallengelassen oder drastisch modifiziert“ (Weiss 1974: 20).

Die Leistungsmessungen und -vergleiche in den Kommunen müssen sich grundsätzlich beiden „Rationalitätskriterien“ stellen und stehen damit vor der doppelten (und ambivalenten)

Aufgabenvollzug stehen kann, ohne dass dieser Konflikt (definitiv) aufgehoben werden kann (vgl. Bogumil 2003).

³⁹ „There is nothing a politician likes so little as to be well informed; it makes decision-making so complex and difficult“ (Keynes zitiert in Sanderson 2002: 5).

Herausforderung, einerseits Transparenz in den politischen Entscheidungsprozess zu bringen, an der aus Sicht einer inhaltlichen Politikorientierung auch ein grundsätzliches Interesse besteht. Andererseits müssen sie das Bestreben der politischen Akteure, sich unter den Bedingungen der Konkurrenzdemokratie zu behaupten und damit ihre Neigung im Blick behalten, das Ausmaß an Durchsichtigkeit im Politikprozess nach dessen Nützlichkeit für die eigene politische Profilierung und Wiederwahl auszurichten. Die Diskussionen um die Frage, inwieweit der interkommunale Leistungsvergleich stärker an die Öffentlichkeit gebracht und Vergleichsergebnisse auch nach außen hin transparent gemacht werden (sollten), ist hierfür symptomatisch (vgl. auch Schuster 2004). So besteht einerseits, insbesondere in den Fachbereichen und -ämtern ein fachlich-technisches (Policy-) Interesse an Leistungsinformationen und -vergleichen, während andererseits die Anonymisierung solcher Vergleichsdaten (Verzicht auf Namen und Einwohnerzahl der Kommune bei der Berichterstattung) nicht nur gegenüber der Öffentlichkeit, sondern auch gegenüber den lokalen politischen Vertretungsorganen vielfach zur Bedingung für eine Teilnahme am Leistungsvergleich gemacht wird (Bsp. Landkreise Niedersachsen, Saarland, Sachsen-Anhalt). Während man in den Verwaltungen noch (zähneknirschend) bereit ist, Leistungsdaten etwa über Intranet intern oder über ein interkommunales Berichtswesen sektoral, also nur für die beteiligten Fachämter zugänglich und bekanntzumachen⁴⁰, wird in der Befürchtung, die Vergleichswerte könnten „politisiert“ und in Form von Ranking-Listen medienwirksam „missbraucht“ werden, auf eine generelle Veröffentlichung verzichtet⁴¹. Gerade vor dem Hintergrund des politischen und budgetären Drucks, die Kennzahlenwerte zur Rechtfertigung weiterer Personalkürzungen zu nutzen, und der sich somit verstärkenden Angst der Beschäftigten vor dem Arbeitsplatzverlust, erweist sich die Frage der Veröffentlichung von Vergleichsdaten und ihre mögliche politische Funktion im Cutback Management als ein ausgesprochen „wunder Punkt“⁴² der Verwaltungsmodernisierung, der in Verwaltungspraxis und -politik weiterer Diskussion bedarf.

⁴⁰ So sprach der Vertreter des Saarländischen Landkreistages von einer wahren „Seelenmassage“, der es bedurft hatte, um die Verwaltungsmitarbeiter zu überzeugen, die Vergleichsdaten ins *Intranet* zu stellen (Interviewprotokoll, 17.12.2003).

⁴¹ Beispielsweise haben im Kennzahlenvergleich der Niedersächsischen Landkreise (siehe Tabelle 2) nur die Fachämter in den betreffenden Aufgabenbereichen einen Überblick über die landesweiten Vergleichsdaten, kennen also beispielsweise die Vergleichswerte ihres Nachbarkreises in dem für sie relevanten Aufgabengebiet. Dagegen bekommen Politik und Öffentlichkeit nur anonymisierte Datenbestände zur Verfügung gestellt (Interviewprotokoll, 18.12.2003).

⁴² So wörtlich der Vertreter des Saarländischen Landkreistages (Interviewprotokoll, 17.12.2003); ähnlich der Gesprächspartner aus dem Landkreistag Sachsen-Anhalt (Interviewprotokoll, 16.12.2003).

Literatur

- Adamaschek, Bernd (1997): Leistungssteigerung durch Wettbewerb in deutschen Kommunen – der interkommunale Vergleich. In: Naschold, Frieder/Oppen, Maria/Wegener, Alexander (Hg.): Innovative Kommunen. Stuttgart, S. 107-124.
- Adamaschek, Bernd (Hg.) (1998): Interkommunaler Leistungsvergleich. 100 spürbare Erfolge. Gütersloh.
- Anke, Thomas/Grabowski, J./Wetzel, R., 1999: Interkommunaler Leistungsvergleich – Kritische Erfolgsfaktoren. In: Adamaschek, Bernd/Baitsch, C. (Hg.): Interkommunaler Leistungsvergleich – Kritische Erfolgsfaktoren. Gütersloh, S. 19-88.
- Askim, Jostein (2004): Benchmarking in Local Government Service Delivery: Window Dressing or a Potent Driver for Improvement? Evidence from Norway. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 277-301.
- Banner, Gerhard (2001): Politik und Transparenz: Benchmarking als Steuerungsinstrument. In: Hofmeister, Albert (Hg.): Bewerten – Vergleichen – Gewinnen. Self-Assessment – Benchmarking – Quality Awards. Schriftenreihe der Schweizerischen Gesellschaft für Verwaltungswissenschaften (SGVW), Band 41. Bern, S. 37-70.
- Benz, Arthur (2003): Einleitung: Governance – Modebegriff oder nützliches sozialwissenschaftliches Konzept? In: Benz, Arthur u.a.: Governance – Eine Einführung. Polis-Heft. Hagen, S. 13-33.
- Bogumil, Jörg (2003): Politische Rationalität im Modernisierungsprozess. In: Schedler, Kuno/Kettinger, D. (Hg.): Modernisieren mit der Politik. Bern u.a., S. 15-42.
- Bogumil, Jörg/Holtkamp, Lars/Schwarz, Gudrun (2003): Das Reformmodell Bürgerkommune. Leistungen – Grenzen – Perspektiven. Berlin.
- Bogumil, Jörg/Kuhlmann, Sabine (2004): Zehn Jahre kommunale Verwaltungsmodernisierung. Ansätze einer Wirkungsanalyse. In: Jann, Werner/Bogumil, Jörg/Bouckaert, Geert/Budäus, Dietrich/Holtkamp, Lars/KiBler, Leo/Kuhlmann, Sabine/Mezger, Erika/Reichard, Christoph/Wollmann, Hellmut: Statusreport Verwaltungsreform – eine Zwischenbilanz nach zehn Jahren. Berlin, S. 51-63.
- Bovaird, Tony/Löffler, Elke (2001): Moving from excellence models of local service delivery to benchmarking ‚good local governance‘. In: International Review of Administrative Science. Vol. 68 (2002), S. 9-24.

- Brüggemeier, Martin (2004): Von der Kunst, erfolgreich zu scheitern - Wirkungsorientiertes Controlling in öffentlichen Verwaltungen. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 374-391.
- Budäus, Dietrich (2004): Modernisierung des öffentlichen Haushalts- und Rechnungswesens. In: Jann, Werner/Bogumil, Jörg/Bouckaert, Geert/Budäus, Dietrich/Holtkamp, Lars/Kißler, Leo/Kuhlmann, Sabine/Mezger, Erika/Reichard, Christoph/Wollmann, Hellmut: Statusreport Verwaltungsreform – eine Zwischenbilanz nach zehn Jahren. Berlin, S. 75-86.
- Con_sens GmbH Hamburg (2002): Benchmarking der mittelgroßen Großstädte in Nordrhein-Westfalen. Kennzahlenvergleich 2000. Hilfe zur Arbeit (HzA). Bericht vom 31. Januar 2002.
- Derlien, Hans-Ulrich (1997): Die Entwicklung von Evaluation im internationalen Kontext. In: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter (Hg.): Einführung in die Politikevaluation. Basel/Frankfurt am Main, S. 4-12.
- Heinelt, Hubert (2003): Governance auf lokaler Ebene. In: Benz, Arthur u.a.: Governance – Eine Einführung. Polis-Heft. Hagen, S. 33-49.
- Herbing, Thomas (2004): Leistungsbemessung, Beteiligung und Motivation – Perspektiven für die gewerkschaftliche Politik. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hrsg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 361-373.
- Innenministerium Baden-Württemberg (2001): Produkte – Ziele – Kennzahlen. Kommunalen Produktplan Baden-Württemberg. Schriftenreihe des Innenministeriums Baden-Württemberg, Heft 7. Stuttgart.
- Jann, Werner (2002): Der Wandel verwaltungspolitischer Leitbilder: Von Management zu Governance? In: König, Klaus (Hg.): Deutsche Verwaltung an der Wende zum 21. Jahrhundert. Baden-Baden, S. 279-303.
- KGSt (1993): Das neue Steuerungsmodell. Begründungen. Konturen. Umsetzungen, KGSt-Bericht Nr. 5/93, Köln.
- Kienbaum Unternehmensberatung GmbH (1996): Projektbericht zum interkommunalen Kennzahlenvergleich für die Hilfe zum Lebensunterhalt.
- Klenk, Tanja/Nullmeier, Frank (2003): Public Governance als Reformstrategie. Düsseldorf.
- Korte, Rainer (2004): Ziele, Aufwand und lokaler Nutzen von Kennzahlen und Kennzahlenvergleichen für Kommunen aus Sicht interkommunaler Vergleichsringe –

- ein Praxisbericht. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 123-137.
- Korte, Rainer/Pook, M. (2002): Leistungsvergleiche haben die Bewährungsprobe bestanden. Austausch von Erfahrungen bietet den Kommunen Verbesserungspotentiale. In: Innovative Verwaltung 7-8/2002, S. 14-17.
- Kuhlmann, Sabine (2003a): Benchmarking auf dem Prüfstand: Kosten, Nutzen und Wirkungen interkommunaler Leistungsvergleiche in Deutschland. In: Verwaltungsarchiv, Heft 1, S. 99-126.
- Kuhlmann, Sabine (2003b): Rechtsstaatliches Verwaltungshandeln in Ostdeutschland. Eine Studie zum Gesetzesvollzug in der lokalen Bauverwaltung. Opladen.
- Kuhlmann, Sabine (2004a): Evaluation lokaler Verwaltungspolitik: Umsetzung und Wirksamkeit des Neuen Steuerungsmodells in den deutschen Kommunen. In: PVS 3/2004, S. 370-394.
- Kuhlmann, Sabine (2004b): Interkommunaler Leistungsvergleich in Deutschland: Zwischen Transparenzgebot und Politikprozess. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 94-121.
- Kuhlmann, Sabine (2004c): Evaluating the Transformation Process in East Germany: Institutional Performance of Local Administration. In: Garcia-Zamor, Jean-Claude (Ed.): Bureaucratic, Societal, and Ethical Transformation of the Former East Germany. Dallas etc., S. 27-65.
- Lindblom, Charles (1975): Inkrementalismus: die Lehre vom „Sich-Durchwursteln“. In: Narr, Wolf-Dieter/Offe, Claus (Hg.): Wohlfahrtsstaat und Massenloyalität. Köln, S. 161-177.
- Niedersächsischer Städte- und Gemeindebund (2002): Interkommunaler Kosten- und Leistungsvergleich 2002. Hannover.
- Niedersächsischer Landkreistag (2002): Allgemeine Grundsätze zu Kennzahlenvergleichen. Stand: Oktober 2002. (ohne Ort)
- Nordrhein-Westfälischer Städte- und Gemeindebund (1996): Interkommunaler Kosten- und Leistungsvergleich 1996. Berichtsband. (ohne Ort).
- Nullmeier, Frank (2004): Zwischen Performance und Performanz – Funktionen und Konzepte der Evaluierung in öffentlichen Verwaltungen. In: Kuhlmann, Sabine/Bogumil,

- Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 47-60.
- Pitarelli, Emilio/ Monnier, Eric (2000): Benchmarking: the missing link between evaluation and management? Paper presented at the 4th ESS Conference October 12-14th 2000, Lausanne.
- Pollitt, Christopher (1995): Justification by Works or by Faith? Evaluating the New Public Management. In: Evaluation (1995)2, S. 133-154.
- Pollitt, Christopher/Bouckaert, Geert (2000): Public Management Reform. A Comparative Analysis. Oxford.
- Reichard, Christoph (1998): Wettbewerbselemente in der öffentlichen Verwaltung – Kommentierung aus wissenschaftlicher Sicht. In: König, Klaus/Füchtner, Natascha (Hg.): „Schlanker Staat“ – Verwaltungsmodernisierung im Bund. Zwischenbericht, Praxisbeiträge, Kommentare. Speyerer Forschungsberichte 155, S. 305 ff.
- Reichard, Christoph (2001): Strategisches Management in der Kernverwaltung. In: Eichhorn, Peter/Wiechers, Matthias (Hg.): Strategisches Management für Kommunalverwaltungen. Baden-Baden, S. 80-91.
- Reichard, Christoph (2004): Ansätze zu Performance Measurement in deutschen Kommunen – eine Bewertung ihres Entwicklungsstandes und ihrer Wirksamkeit. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 341-356.
- Saatweber, Vera S. (2004): Das europäische Selbstbewertungsinstrument Common Assessment Framework (CAF). In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 227-249.
- Sächsischer Landkreistag/Sächsischer Städte- und Gemeindebund (2003): Kommunale Verwaltungsmodernisierung Sachsen. Abschluss der Projektphase II. (Ohne Ort).
- Sanderson, Ian (2002): Evaluation, Policy Learning and Evidence-Based Policy Making. In: Public Administration (2002)1, S. 1-22.
- Schmidt, Manfred G. (1995): Policy-Analyse. In: Mohr, Arno (Hg.): Grundzüge der Politikwissenschaft. München/Wien, S. 567-605.
- Schuster, Ferdinand (2001): Benchmarking als Ersatz für Wettbewerb. Können interkommunale Leistungsvergleiche ein Motor für Veränderungen sein? In: Edeling, Thomas/Jann, Werner/Wagner, Dieter (Hg.): Reorganisationsstrategien in Wirtschaft und Verwaltung. Opladen: Leske+Budrich, S. 201-228.

- Schuster, Ferdinand (2003): Der interkommunale Leistungsvergleich als Wettbewerbssurrogat. Berlin.
- Schuster, Ferdinand (2004): Nicht wie Feuer und Wasser: Leistungsvergleiche und Öffentlichkeit – ein Zwischenruf. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden,, S. 357-360.
- Strid, Lars (2004): Comparative Municipal Quality Networks in Sweden. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 267-276.
- Tebbe, Günter (2004): Arbeiten mit Kennzahlen: Anwendungserfahrungen aus den Projekten kik und kompass der Bertelsmann Stiftung. Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 138-147.
- Wegener, Alexander (2004): Benchmarking-Strategien im öffentlichen Sektor: Deutschland und Großbritannien im Vergleich. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 251-266.
- Weiss, Carol H. (1974): Einführung. In: Küchler, M. (Hrsg.): Evaluierungsforschung. Methoden zur Einschätzung von sozialen Reformprogrammen. Opladen, S. 19-30.
- Wollmann, Hellmut (2000): Evaluierung und Evaluierungsforschung von Verwaltungspolitik und -modernisierung – zwischen Analysepotenzial und –defizit. In: Stockmann, Reinhard (Hg.), Evaluationsforschung. Opladen, S. 195-233.
- Wollmann, Hellmut (2001): Zwischen Management- und Politiksystem: Die Kommunalverwaltung der 90er Jahre auf einer Modernisierungswelle. In: Edeling, Thomas/Jann, Werner/Wagner, Dieter (Hg.): Reorganisationsstrategien in Wirtschaft und Verwaltung. Opladen, S. 15-57.
- Wollmann, Hellmut (2004a): Leistungsmessung („performance measurement“) in Politik und Verwaltung: Phasen, Typen und Ansätze im internationalen Überblick. In: Kuhlmann, Sabine/Bogumil, Jörg/Wollmann, Hellmut (Hg.): Leistungsmessung und Leistungsvergleich in Politik und Verwaltung. Wiesbaden, S. 21-46.
- Wollmann, Hellmut (2004b): Reformen der kommunalen Politik- und Verwaltungsebene in Großbritannien, Schweden und Frankreich. Ansätze, Verläufe und Ergebnisse. In: Jann, Werner/Bogumil, Jörg/Bouckaert, Geert/Budäus, Dietrich/Holtkamp, Lars/Kißler, Leo/Kuhlmann, Sabine/Mezger, Erika/Reichard, Christoph/Wollmann,

Hellmut: Statusreport Verwaltungsreform – eine Zwischenbilanz nach zehn Jahren.
Berlin, S. 36-50.

Wollmann, Hellmut/Schröter, Eckhard (1997): Public Sector Reforms in Germany: Whence and Where? A Case of Ambivalence. In: Administrative Studies/ Hailinnon Tutmikus, 1997, no. 3, S. 184-200.