

### Leistungsbeurteilung im öffentlichen Dienst: zur Validierung des Leistungsbeurteilungs- und -bewertungssystems LBB-SYS

Holtmann, Doris; Matiaske, Wenzel; Möllenhoff, Dieter; Weller, Ingo

Erstveröffentlichung / Primary Publication

Arbeitspapier / working paper

#### Empfohlene Zitierung / Suggested Citation:

Holtmann, D., Matiaske, W., Möllenhoff, D., & Weller, I. (2001). *Leistungsbeurteilung im öffentlichen Dienst: zur Validierung des Leistungsbeurteilungs- und -bewertungssystems LBB-SYS*. (Berichte der Werkstatt für Organisations- und Personalforschung e.V., 4). Berlin: Werkstatt für Organisations- und Personalforschung e.V.. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-410360>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:

<https://creativecommons.org/licenses/by-nc-nd/4.0>

**Werkstatt**

für Organisations- und  
Personalforschung e.V.



## Leistungsbeurteilung im öffentlichen Dienst

---

Zur Validierung des Leistungsbeurteilungs-  
und -bewertungssystems LBB-SYS

**Doris Holtmann, Wenzel Matiaske, Dieter Möllenhoff und Ingo Weller**

---

Berichte der Werkstatt für Organisations- und Personalforschung e.V., ISSN 1615-8261

## Die Autoren:

**Dipl.-Betriebswirtin Doris Holtmann** ist Mitarbeiterin im Projekt „Leistungsbeurteilung und leistungsorientierte Vergütung in öffentlichen Organisationen“ am Lehrstuhl für ABWL, insbesondere Personal und Organisation, am Institut für Internationales Management der Universität Flensburg und Mitglied der Werkstatt für Organisations- und Personalforschung e.V. Berlin.

Email: Doris.Holtmann@werkstatt-opf.de

**Prof. Dr. Wenzel Matiaske** ist Professor für ABWL, insbesondere Personal und Organisation, am Institut für Internationales Management der Universität Flensburg und Mitglied der Werkstatt für Organisations- und Personalforschung e.V. Berlin.

Email: Wenzel.Matiaske@werkstatt-opf.de

**Dieter Möllenhoff**, D. Möllenhoff GmbH, Auf der Au 16, 69502 Hemsbach und Mitglied der Werkstatt für Organisations- und Personalforschung e.V. Berlin.

Email: Dieter.Moellenhoff@werkstatt-opf.de

**Dipl.-Kfm. Ingo Weller** ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Personalwirtschaft der Universität Paderborn und Mitglied der Werkstatt für Organisations- und Personalforschung e.V. Berlin.

Email: Ingo.Weller@werkstatt-opf.de

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung der Werkstatt für Organisations- und Personalforschung e.V. unzulässig. Dies gilt insbesondere für Vervielfältigungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

© **Werkstatt für Organisations- und Personalforschung e.V.**  
Berlin 2001

## Berichte der Werkstatt für Organisations- und Personalforschung e.V.

### Bericht Nr. 4, Berlin 2001

ISSN 1615-8261

### Kontakt zur Werkstatt für Organisations- und Personalforschung e.V.:

Dr. Renate Ortlieb  
Südwestkorso 76  
12161 Berlin  
Tel.: 030/89739919  
Fax: 030/89739919

email: Kontakt@Werkstatt-opf.de  
Internet: www.werkstatt-opf.de

### Vorstandsmitglieder und wissenschaftlicher Beirat der Werkstatt für Organisations- und Personalforschung e.V.:

Prof. Dr. Albert Martin  
Prof. Dr. Wenzel Matiaske  
Prof. Dr. Thomas Mellewig  
Prof. Dr. Eckart Minx  
Prof. Dr. Werner Nienhüser  
Prof. Dr. Florian Schramm

# **Leistungsbeurteilung im öffentlichen Dienst**

## **Zur Validierung des Leistungsbeurteilungs- und -bewertungssystems LBB-SYS**

Doris Holtmann      Wenzel Matiaske      Dieter Möllenhoff      Ingo Weller

### **Zusammenfassung**

Ein zentrales personalwirtschaftliches Steuerungselement im Reformprozess öffentlicher Organisationen ist der verstärkte Einsatz materieller Anreizsysteme. Die Vergabe von Leistungszulagen hat jedoch Leistungsbeurteilungen zur Voraussetzung. Der Bericht diskutiert Anforderungen an Instrumente zur Leistungsbeurteilung und stellt mit dem Leistungsbeurteilungs- und -bewertungssystem LBB-SYS ein solches Instrument vor. Empirisch wird die Verlässlichkeit (zeitpunktbezogene interne Konsistenz, Inter-Rater-Reliabilität und Retest-Reliabilität) und die faktorielle Gültigkeit des mehrdimensionalen Beurteilungssystems geprüft.

### **Abstract**

Monetary compensation packages are central HR-elements in the reformative efforts of public organizations. But performance oriented payments are bound to prior performance appraisals. Therefore this report discusses basic elements and requirements of performance appraisal instruments and introduces the LBB-SYS (Leistungsbeurteilungs- und -bewertungssystem) as an example. The empirical part of the report features reliability measures (internal scale consistencies, inter-rater-reliabilities and retest-reliabilities) as well as the results of a confirmator factor analysis to examine the validity of the multi-dimensional appraisal system.

---

## Inhaltsverzeichnis

1	Einführende Bemerkungen . . . . .	3
2	Serviceorientierung und Personalmanagement in öffentlichen Organisationen: Das Neue Steuerungsmodell . . . . .	4
3	Leistungsbeurteilung und -vergütung in öffentlichen Organisationen	5
3.1	(Tarif-)rechtliche Rahmenbedingungen . . . . .	5
3.2	Erwartungen an Leistungsbeurteilung und -vergütung . . . . .	6
3.3	Anforderungen an Leistungsbeurteilungssysteme . . . . .	8
4	Das Leistungsbeurteilungs- und -bewertungssystem LBB-SYS . .	11
4.1	Aufbau des Beurteilungs- und Bewertungsinstrumentes . . . . .	11
4.2	Durchführung von Beurteilungen . . . . .	14
4.3	Einführung des Leistungsbeurteilungs- und -bewertungssystems .	15
4.4	Personalpolitische Leitlinien und Prinzipien der Personalführung im LBB-SYS . . . . .	16
5	Verlässlichkeit und Gültigkeit des LBB-SYS . . . . .	16
5.1	Zeitpunktbezogene Verlässlichkeit und Gültigkeit . . . . .	17
5.2	Übereinstimmungs-Verlässlichkeit . . . . .	21
5.3	Retest-Verlässlichkeit . . . . .	22
6	Resümee . . . . .	23
	Literatur . . . . .	24
A	Statistiken der Subskalen des LBB-SYS . . . . .	25
B	Ergebnisse der Explorativen Hauptkomponentenanalyse . . . . .	27

## 1 Einführende Bemerkungen

Das deutsche Verwaltungswesen befindet sich im Umbruch. Akute Finanzprobleme, gestiegene Ansprüche der neuerdings gern als „Kunden“ bezeichneten Bürger sowie die oft zitierten Leistungslücken der Verwaltungen und ihres Personals bilden den Forderungshintergrund der Reformbemühungen (Klages 1995). Auf Ebene der Kommunalverwaltungen bündelt das sogenannte „Neue Steuerungsmodell“ die Leitideen zu ihrer Reform (KGSt 1993). Dieses Modell beinhaltet neben allgemein verwaltungstechnischen, betriebswirtschaftlichen und organisatorischen Überlegungen vor allem auch neue Aufgaben für das Personalmanagement in öffentlichen Organisationen. Eine zentrale Forderung, die auch in der öffentlichen Diskussion erhoben wird, ist die Einführung leistungsbezogener Entgeltsysteme. Diese haben wiederum Systeme der Leistungsbeurteilung zur Voraussetzung.

Unser Bericht stellt mit dem LBB-SYS<sup>1</sup> ein solches Leistungsbeurteilungssystem vor. Vorrangiges Ziel des Beitrages ist es, in mehrjähriger Praxis mit dem Instrument gewonnene Daten zur Prüfung der testtheoretischen Gültigkeit des LBB-SYS zu nutzen und auf Grundlage der statistischen Analysen Verbesserungsvorschläge zur Weiterentwicklung des Systems auszuarbeiten. Gleichzeitig soll diese Arbeit exemplarisch zeigen, welchen Ansprüchen Leistungsbeurteilungssysteme genügen sollten und wie diese aus testtheoretischer Sicht zu validieren sind.

Praktikable Leistungsbeurteilungssysteme müssen jedoch nicht nur testtheoretischen Kriterien genügen. Sie sollten darüber hinaus in den organisatorischen Kontext passen und von Beurteilten sowie Beurteilern akzeptiert werden. Wir diskutieren daher im Folgenden zunächst Rahmenbedingungen und Möglichkeiten von Leistungsbeurteilungen in öffentlichen Organisationen, insbesondere Kommunalverwaltungen. Ausgangspunkt ist eine Skizze der personalwirtschaftlichen Implikationen des Neuen Steuerungsmodells. Darauf aufbauend legen wir im dritten Abschnitt des Berichtes Grundlagen, Zielsetzungen und Nutzen von Verfahren der Leistungsbeurteilung im öffentlichen Dienst dar. Das vierte Kapitel stellt das LBB-SYS näher vor. Wir beschreiben den Aufbau des Instrumentes und gehen auf Details wie enthaltene Skalen und Bewertungskriterien näher ein. Darüber hinaus erläutern wir auch Aspekte der Durchführung und Handhabung des Systems. Das fünfte Kapitel bildet den eigentlichen Analyseteil. Basis unserer Analyse sind die Daten von drei Jahresbeurteilungen der Mitarbeiter eines kommunalen Betriebshofes mittels des LBB-SYS. Wir können daher neben der zeitpunktbezogenen Verlässlichkeit und Gültigkeit des Instrumentariums auch die Verlässlichkeit bei wiederholtem Einsatz analysieren. Ein Ausblick sowie auf Basis dieser Analysen entwickelte Verbesserungsvorschläge zum Instrument runden die Arbeit ab.

---

<sup>1</sup> Das LBB-SYS ist ein eingetragenes Warenzeichen der D. Möllenhoff GmbH (Auf der Au 16, 69502 Hemsbach, Tel. 06201/7780, Email: info@moellenhoff-gmbh.de).

## **2 Serviceorientierung und Personalmanagement in öffentlichen Organisationen: Das Neue Steuerungsmodell**

Das 1993 von der Kommunalen Gemeinschaftsstelle für Verwaltungsvereinfachung (KGSt) erarbeitete Neue Steuerungsmodell zur Modernisierung der deutschen Kommunalverwaltung ist zentraler Bezugspunkt der Reformbestrebungen auf kommunaler Ebene. Das Modell hat die Aufgabe, die Verwaltung als „primär behördlich geprägte Eingriffs- und Betreuungsapparatur zu einer kostenbewußten, marktnahen, mit ihren Bürgern zusammenarbeitenden ‚Problemlöseeinrichtung‘ umzuwandeln.“ (KGSt 1993). Impulse zu dieser Verwaltungsreform stammen unter anderem aus dem Ausland, wo bereits Anfang der 80er Jahre Konzepte zum „New Public Management“ umgesetzt wurden. Diese waren durch Stichworte wie Privatisierung, drastische Kostensenkung sowie neue Führungskonzeptionen gekennzeichnet, welche letztendlich auch als Grundlage für die Leitlinie der KGSt dienten.

Die KGSt definiert folgende Kernelemente des Neuen Steuerungsmodells: klare Verantwortungsabgrenzung zwischen Politik und Verwaltung, Führung durch Leistungsabsprache statt durch Einzeleingriff (Kontraktmanagement), dezentrale Gesamtverantwortung in den Fachbereichen und Einsatz von Instrumenten zur Steuerung der Verwaltung von der Leistungsseite (Outputsteuerung) (KGSt 1993). Diese Elemente sind als abstrakte organisatorische Bausteine für die Schaffung einer serviceorientierten Verwaltung anzusehen. Aus diesen Bausteinen wiederum lassen sich Gestaltungshinweise für das Personalmanagement im öffentlichen Sektor ableiten. „Die Umsetzung der Reformüberlegungen hängt wesentlich von der Mitwirkung der Mitarbeiterinnen und Mitarbeiter der Verwaltung ab. Ihre Akzeptanz und aktive Beteiligung sind genauso wie ihre Motivation und Qualifikation neben der Organisation der Zusammenarbeit entscheidend für den Erfolg der Verwaltungsreform“ (KGSt 1996). Um dieses Mitarbeiterpotential zu nutzen, setzt das Neue Steuerungsmodell vor allem auf Maßnahmen der Personalentwicklung: So sollen lernförderliche Arbeitsstrukturen (Teams) geschaffen werden, Mitarbeitergespräche sollen die Kommunikations- und Kooperationsfähigkeiten schulen und der Leistungs- und Zielvereinbarung dienen. Zusätzlich werden kooperative Führungsformen propagiert sowie ständige Qualifizierungsmaßnahmen angeregt.

Neben immateriellen Anreizen sollen jedoch auch materielle Anreize zum Einsatz kommen (Tondorf 1995). Hier müssen Leistungszulagen von Leistungsprämien unterschieden werden. Leistungsprämien können auf der Anspruchsgrundlage eines einmalig erbrachten Arbeitsergebnisses gewährt werden (Uttlinger et al. 2000). Demgegenüber muß die „Gewährung von Leistungszulagen an einzelne Arbeitnehmer ... auf einem leistungsbezogenen System zur Leistungs-

feststellung und Leistungsbewertung beruhen. Es ist ein Beurteilungssystem erforderlich, das verständliche Kriterien vermittelt“ (Uttlinger et al. 2000).

### **3 Leistungsbeurteilung und -vergütung in öffentlichen Organisationen**

#### **3.1 (Tarif-)rechtliche Rahmenbedingungen**

Die seit Jahrzehnten geführte Diskussion zur Leistungsentlohnung in der öffentlichen Verwaltung mündete im Dienstrechtsreformgesetz aus dem Jahr 1997. Das Leitmotiv des Gesetzes ist die Stärkung der Leistungsorientierung im öffentlichen Dienst und es ermöglicht die Einführung von leistungsbezogenen Zulagen für Beamte. Bereits in den Jahren 1984/85 wurde die Zahlung leistungsorientierter Zulagen in den Manteltarifvertrag für Arbeiter aufgenommen. Ebenfalls 1997 wurde für Angestellte im öffentlichen Dienst auf Bezirksvertragsebene ein Zulagentarif eingeführt.

Am Beispiel des Bundesangestelltentarifs (BAT) aus dem Jahr 1961 lässt sich die zentrale Problematik der Entgeltdifferenzierung im öffentlichen Dienst veranschaulichen, die für Arbeiter und Beamte in gleicher Weise gilt. Die Gesamtvergütung im BAT besteht aus den Komponenten Grundvergütung, Steigerungsbeträge nach Lebensaltersstufen sowie Ortszuschlag. Sie richtet sich in ihrer Höhe nach einer Eingruppierung der Arbeit auf Grund von Tätigkeitsmerkmalen, welche sich am Kriterium der Anforderungsgerechtigkeit orientiert. Innerhalb der Vergütungsgruppen sind nach zeitlichen Kriterien bemessene Bewährungsaufstiege vorgesehen, welche in den genannten Steigerungsbeträgen ihren Niederschlag finden

Im Vergütungssystem des BAT besteht also grundsätzlich nicht die Möglichkeit der besonderen Entlohnung individueller Leistungen. Durch die gegebene Objektivität und Transparenz des Systems kann dies zu Motivationsproblemen führen, da intersubjektive Vergleiche der Inputs (Arbeitsleistung) und Outputs (Gehalt) ermöglicht werden (z. B. Adams 1963). Die Einführung einer leistungsorientierten Entlohnung läßt eine Verminderung der Motivationsproblematik erwarten und könnte somit positive motivationale Zielsetzungen erfüllen. Leistungsorientierte Entlohnungsformen stellen jedoch nicht nur ein Instrument der Leistungssteuerung in der Verwaltung dar, „viel bedeutsamer ist gegenwärtig die Chance, dieses Instrumentarium für eine Reform der öffentlichen Verwaltung einzusetzen“ (Tondorf 1995, S. 13). Die Einführung eines solchen Reforminstrumentes ermöglicht eine individuelle Gestaltung der Umsetzung in den Kommunen, da die zentrale Tarifebene die Differenziertheit der Tätigkeiten und die Anwendungsvielfalt von Leistungsentlohnung im Einzelnen nicht regeln kann (Ton-



dorf 1995, S. 9). Die konkrete Ausgestaltung von Leistungsbewertung- und beurteilung erfolgt durch Dienstvereinbarungen vor Ort.

Die Denzentralisierung der Lohnpolitik und die damit verbundene Verlagerung von Tarifkompetenzen, die mit dem Einsatz dieses lohnpolitischen Instrumentes einhergeht, wird von Seiten der Gewerkschaften kritisch gesehen. „Auch im Hinblick auf die Schutzfunktion, die Gewerkschaften gegenüber ihren Mitgliedern wahrnehmen, ist Leistungsentlohnung ein heikles tarifpolitisches Thema“ (Tondorf 1995, S. 9).

### **3.2 Erwartungen an Leistungsbeurteilung und -vergütung**

Trotz aller Befürchtungen stellt die Leistungsbeurteilung und -bewertung ein für die Praxis unverzichtbares Instrument dar. „Individuelle Leistung bleibt offenbar das zentrale Kriterium, nach dem knappe Mittel wie Arbeitsentgelt und berufliche Aufstieg verteilt bzw. gesteuert werden können. Auch wenn der Leistungsbezug hierfür nicht das einzig akzeptierte und legitime Mittel in unserem Wirtschafts- und Gesellschaftssystem ist, scheint es zur Leistung keinen effektiven alternativen Verteilungsmechanismus zu geben“ (Liebel/Oechsler 1992, S. 13). Im Fall öffentlicher Verwaltungen ist darüber hinaus zu berücksichtigen, dass diese dem Druck zur Modernisierung mittels der Steuerungsmöglichkeit von Leistung und Personal in Form von leistungsgerechter Entlohnung als einem Baustein des Personalmanagements begegnen wollen.

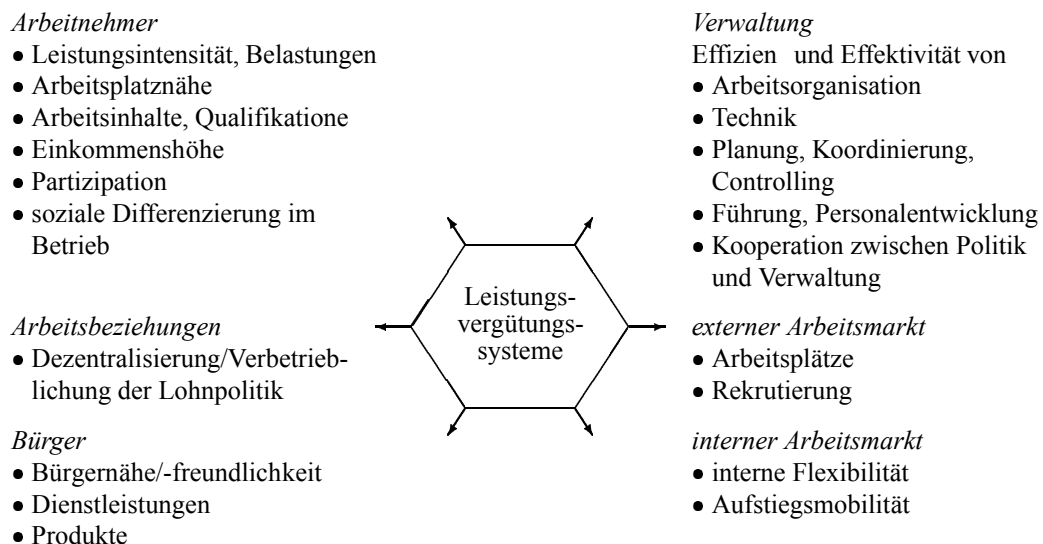
Entsprechend hoch sind die mit diesem Instrumentarium in der Wirtschafts- und Verwaltungspraxis verbundenen Erwartungen und Ziele. Sie lassen sich allgemein wie folgt skizzieren (Liebel/Oechsler 1992, S. 15):

- Individuelle Lohn- und Gehaltsfindung (Einkommensgerechtigkeit)
- Steigerung der Arbeitszufriedenheit
- Leistungsverbesserung durch Verhaltens- und Motivationssteuerung (Mitarbeitergespräche)
- Entscheidungsgrundlage für personelle Einzelmaßnahmen (Potenzialnutzung bei Stellenbesetzung durch Beförderung, Versetzung, Übernahme, Kündigung)
- Entscheidungsgrundlage auf kollektiver Ebene (Personalplanung)
- Förderung der Mitarbeiterqualifikation (Planung, Auswahl und Gestaltung von Maßnahmen der Personalentwicklung)
- Einführung einer Personalentwicklungsplanung (kollektive Ebene)

- Gestaltung von Arbeitsbedingungen (Arbeitsplatz und Arbeitsumgebung), Ausgangspunkt von Organisationsdiagnose und Organisationsentwicklung
- Individuelle Beratung und Förderung von Mitarbeitern
- Verbesserung der Führungskompetenz der Vorgesetzten
- Evaluation von Selektionskonzepten, personellen Entscheidungen, Maßnahmen der Personalentwicklung, Programmen der Organisationsentwicklung, Anreiz- und Verstärkungssystemen

Leistungsbeurteilung und -vergütung ist demnach Ausgangspunkt einer Vielzahl weiterer personalwirtschaftlicher Maßnahmen. Dies bedeutet auch, dass sich die weitreichenden Erwartungen an dieses Instrumentarium nur erfüllen lassen, wenn es systematisch mit anderen Aufgabenbereichen des Personalmanagements verzahnt wird. Richtet man den Fokus auf die erwarteten Wirkungsdimensionen leistungsabhängiger Entgeltsysteme öffentlicher Verwaltungen, lassen sich diese — unterteilt nach ökonomischen, politischen und sozialen Bedeutungen — wie in Abbildung 1 zusammenfassen.

Bezieht man diese Erwartungen auf die Instrumente zur Personalbeurteilung wird deutlich, dass mittels eines Beurteilungsverfahrens nicht alle Erwartungen mit der gleichen Priorität bedient werden können. Darüber hinaus sind einige der üblichen Instrumente — z. B. Verfahren der freien Eindruckschilderung —



**Abb. 1:** Wirkungsdimensionen von Leistungsvergütungssystemen in der Verwaltung (Tondorf 1995, 30)

grundsätzlich kaum geeignet, den Erwartungen gerecht zu werden. Abbildung 2 vermittelt einen Überblick gebräuchlicher Verfahren der Personalbeurteilung.

		Differenzierungsgrad	
		summarische Beurteilung	analytische Beurteilung
Skalen- niveau	nominal	—	Checklisten-Verfahren
	ordinal	Rangfolgeverfahren	Rangreihenverfahren
	intervall	Leistungsgruppenverfahren	Skalenverfahren

Abb. 2: Verfahren der Personalbeurteilung

### 3.3 Anforderungen an Leistungsbeurteilungssysteme

Die in der personalwirtschaftlichen Literatur diskutierten Anforderungen, denen Leistungsbeurteilungsverfahren genügen sollen, sind vielfältig. „So besteht bis heute keine einheitliche Auffassung (abgesehen von der klassischen Testtheorie) über die Zusammenstellung und den Inhalt von Gütekriterien“ (Mungenast 1990, S. 11). Eine zweiteilige Klassifizierung genereller (z. B. Objektivität, Praktikabilität, Akzeptanz, Schutz der Persönlichkeitsrechte) und zweckspezifische (z. B. Verwertbarkeit) Anforderungen an Beurteilungssystemen überwiegen in der Diskussion (Gaugler et al. 1978, Lessmann 1980, Becker 1992).

Alternativ unterscheidet man die Wirksamkeit von Beurteilungssystemen anhand von Funktionsvoraussetzungen (Boerger 1982):

- methodische Voraussetzungen (z. B. Einheitlichkeit)
- funktionale Voraussetzungen (z. B. Vorhandensein von Stellenanforderungen)
- personelle Voraussetzungen (z. B. Akzeptanz des Verfahrens bei den Beteiligten)
- situative Voraussetzungen (z. B. Angemessenheit der Organisationsstruktur, Angemessenheit der Führungsphilosophie, Bereitstellung von Kapazitäten und Mitteln)

Damit sind wichtige Voraussetzungen für die Durchführung und damit das Gelingen von Leistungsbeurteilungen genannt. Die Beurteilungssysteme sollten wie erwähnt auch den Gütekriterien der (klassischen) Testtheorie entsprechen.

chen. Diese Kriterien verlangen Verlässlichkeit (Reliabilität) und darüber hinaus Gültigkeit (Validität) der Beurteilungsinstrumente (Matiaske 1996 oder Nienhüser/Becker 2000). Unter Verlässlichkeit versteht man, dass bei mehrfachem Einsatz eines Beurteilungsinstrumentes möglichst gleiche Ergebnisse erzielt werden (Retest-Verlässlichkeit). Im Fall von Leistungsbeurteilungen schwanken die Ergebnisse auch bei an sich verlässlichen Beurteilungsinstrumenten auf Grund der Veränderungen im Leistungsverhalten der Beurteilten. Daher wird zur Prüfung der Verlässlichkeit eines Instrumentes zusätzlich die zeitpunktbezogene Reliabilität bestimmt, d. h. es werden mehrere vergleichbare Beobachtungskriterien auf ihre Übereinstimmung hin geprüft. Im Falle mehrerer Beurteiler bedeutet dies, dass diese mittels des Beurteilungsinstrumentes zu möglichst übereinstimmenden Urteilen bezogen auf einem Beurteilungszeitpunkt gelangen sollten (Inter-Rater-Reliabilität). Gültigkeit meint darüber hinaus, dass die Maßstäbe der Beurteilung auch das erfassen, was gemessen werden soll. Soll z. B. die individuelle Güte bei der Bewältigung der Aufgaben eines Mitarbeiters bestimmt werden, so sollte die kollegiale Unterstützung bei diesem Beobachtungskriterium keine Rolle spielen. Im Falle mehrdimensionaler sogenannter analytischer Beurteilungsinstrumente beinhaltet dies, dass sich die Beurteilungsdimensionen statistisch voneinander trennen lassen.

Anzumerken ist, dass die Verwendung dieser Gütekriterien bei betrieblich eingesetzten Beurteilungsverfahren gelegentlich kritisch betrachtet oder zumindest als überflüssig angesehen wird. So weist beispielsweise Mungenast (1990, S. 12) hinsichtlich der Kriterien Reliabilität und Validität darauf hin, „dass es sich hierbei in erster Linie um Maße handelt, denen wissenschaftliche Tests zu genügen haben. Die bei Experimenten zu erreichenden Formalierungs- und Standardisierungsbedingungen (z. B. Willkürlichkeit, Wiederholbarkeit, Variierbarkeit) bestehen in der betrieblichen Beurteilungsrealität weitgehend nicht.“ Es ist jedoch zu bedenken, dass mit Leistungsbeurteilungen, wie zuvor besprochen, weitreichende personelle Entscheidungen verknüpft sind. Wie z. B. Eltern nicht wünschen, dass die Entscheidung über die schulische Karriere ihrer Kinder von einem wenig verlässlichen Intelligenztest oder subjektiven Urteilen der Lehrer abhängt, so wenig wünscht ein Mitarbeiter, dass seine Leistungszulagen oder seine Karriere auf unzuverlässigen Leistungsbeurteilungen basieren. Die Güteprüfung des LBB-SYS ist daher Bestandteil des fünften Kapitels dieses Berichts und wird unter den Aspekten zeitpunktbezogener Verlässlichkeit und Gültigkeit sowie Inter-Rater-Reliabilität und Test-Retest-Reliabilität behandelt. Letzlich dient die Bewertung von Beurteilungsverfahren anhand der Gütekriterien der klassischen Testtheorie der Konstruktion von Instrumenten, die möglichst wenige Beurteilungsfehler produzieren.

Außerhalb der wissenschaftlichen Klassifizierung von Verlässlichkeit und Gültigkeit ist anzumerken, dass der öffentliche Dienst zur Verbesserung bishe-

riger Beurteilungsverfahren im Zuge des Aktionsprogramms zur Dienstrechtsreform folgende Anforderungen an ein Beurteilungswesen formuliert (Kroppenstedt 1981):

- Stärkere Differenzierung zwischen Leistung und Befähigung
- Stärkere Differenzierung zwischen den Mitarbeitern
- Versachlichung der Leistungsbeurteilung
- Erhöhung des Aussagegehalts
- Erhöhung der Transparenz
- Minimierung von Beurteilungsfehlern
- Stärkere Einbeziehung des Mitarbeiters in den Beurteilungsablauf
- Anwendung gleicher Beurteilungsmaßstäbe in den Dienststellen

Diese Anforderungen lassen sich wohl nicht allein, jedoch sicherlich nur auf Grundlage verlässlicher und gültiger Beurteilungsinstrumente erreichen.

Aus Perspektive eines Personalmanagements im Sinne des Neuen Steuerungsmodells sind vor allem die mit den Leistungsvergütungen verbundenen Erwartungen bezüglich der Arbeits- und Reformbereitschaft der Beschäftigten hervorzuheben. Die notwendige Zielsetzung, die individuelle Leistungsbereitschaft der Beschäftigten im öffentlichen Dienst zu sichern und zu steigern, ist unumstritten. Notwendig, so zeigen die geführten Diskussionen, ist ebenfalls die systematische Verknüpfung mit anderen Reformansätzen, wie beispielsweise Dezentralisierung oder Qualitätsmanagement.

Andererseits wird der Einsatz von Leistungsvergütung unter dem Gesichtspunkt kurzfristiger Kosteneinsparungen mit im Vorfeld festgelegten Verteilungsvorgaben befürchtet (Steinort 1996). So wird von vielen Angestellten im öffentlichen Dienst die Einführung leistungsorientierter Entlohnungsformen mit Skepsis betrachtet. „Leistungsvergütung wird oft gleichgesetzt mit mehr Arbeitshetze, mit stärkerer Kontrolle durch den Vorgesetzten und erhöhter Konkurrenz innerhalb der Belegschaft. Manche Angestellte fürchten um ihre Autonomiespielräume in der Arbeit und sehen sich auf eine Stufe gestellt mit Akkordarbeitern, die — so die Vorstellung — eingepreßt sind in ein System von Vorgabezeiten und technischen Leistungskontrollen. Bei anderen Angestellten wecken Leistungsvergütungen Assoziationen an ‚Nasenprämien‘, die vom Vorgesetzten nach unkontrollierbaren Kriterien vergeben werden“ (Tondorf 1995, S. 13).

Dieses Bild bestätigte sich bei Gesprächen mit Führungskräften im Rahmen eines Begleitforschungsprojektes zur Einführung von Leistungsbeurteilungen in

einer baden-württembergischen Stadtverwaltung. Das Thema der Leistungsbeurteilung wurde von den verantwortlichen Führungskräften im Zuge einer Beratung aufgegriffen. Nach ihren Aussagen basierte die Gewährung von Leistungszulagen zum betreffenden Zeitpunkt weniger auf objektiven Kriterien, als vielmehr auf den persönlichen Präferenzen der Vorgesetzten. Demzufolge erhielten einige wenige bevorzugte Mitarbeiter relativ hohe Leistungszulagen, während der Mehrheit nur eine geringe oder gar keine Leistungszulage gewährt wurde. Die daraus resultierende Unzufriedenheit, gerade bei den jüngeren und neu hinzugekommenen Mitarbeitern, führte zu einer nur geringen Leistungsmotivation der Angestellten.

Diese Erfahrungen belegen exemplarisch, dass ein Instrument zur Leistungsbeurteilung nicht nur den Gütekriterien der klassischen Testtheorie genügen sollte, sondern darüber hinaus auf Akzeptanz der Beteiligten stoßen muss. Die Erwartungen an die Leistungsbeurteilung und -vergütung im öffentlichen Dienst werden nur dann nicht durch die Befürchtungen der Mitarbeiter unterlaufen, wenn die eingesetzten Instrumente zu transparenten, d. h. den Beteiligten verständlichen Beurteilungen führen, die Bewertungen verlässlich und gültig sind und die Konsequenzen der Beurteilungen nicht durch festgelegte Verteilungsvorgaben unter der Maßgabe von Kosteneinsparungen unterlaufen werden.

## **4 Das Leistungsbeurteilungs- und -bewertungssystem LBB-SYS**

### **4.1 Aufbau des Beurteilungs- und Bewertungsinstrumentes**

Das von der D. Möllenhoff GmbH entwickelte Leistungsbeurteilungsinstrument LBB-SYS ist den analytischen Verfahren zuzuordnen und stellt ein vergangenheitsbezogenes System der Leistungsbewertung dar. Bei analytischen Verfahren werden nicht die Gesamtleistung sondern einzelne Dimensionen der Leistung bewertet. Im hier vorgestellten Instrument werden diese Dimensionen Bewertungsgruppen genannt. Das modular aufgebaute LBB-SYS unterscheidet i. d. R. neun Bewertungsgruppen, denen zwischen drei und sechs Beobachtungskriterien zugeordnet sind und anhand derer die Leistung bzw. das Leistungsverhalten der Mitarbeiter beurteilt wird. Anzumerken ist, dass die Modularität des Systems eine aufgabenbezogene Erweiterung (z. B. zur Beurteilung des Führungsverhaltens) bzw. Kürzung der Bewertungsgruppen impliziert.

Auf den vorliegenden Datensatz bezogen wurden neun Bewertungsgruppen mit insgesamt 49 Kriterien beobachtet. Die Beobachtungen werden in Bezug auf die Merkmalserfüllung vorgegebenen Bewertungsstufen zugeordnet. Zur Ermittlung der Gesamtleistung werden die Gesamtpunktzahlen der einzelnen Bewertungsgruppen addiert und entweder in einem linearen oder progressiven Sy-

stem als prozentuale Leistungszulage ausgewiesen. Die Wahl der Zuordnung in ein lineares oder progressiv gestaltetes System obliegt dem Auftraggeber und ist abhängig von der Grundzielsetzung bei Einführung des Instruments.

Die neun Bewertungsgruppen beschreiben in Form von Oberkriterien aufgaben- und verhaltensbezogene Merkmale. Sie werden im Folgenden kurz charakterisiert:

1. Beginnend werden Arbeitseinsatz, Arbeitsintensität und Arbeitsergebnisse beurteilt, die anhand von sechs Beobachtungskriterien differenziert betrachtet werden.
2. Der zweite Katalog umfasst die Aspekte Verantwortung, Kostenbewußtsein und Selbständigkeit. In dieser Bewertungsgruppe beziehen sich die Beobachtungskriterien auf die Bereitschaft zur Übernahme von Verantwortung, auf das Ausschöpfen der Befugnisse durch den Mitarbeiter sowie die sachgerechte Umsetzung von Arbeitsrichtlinien. Eigenverantwortliches und selbständiges Handeln und Kriterien bezüglich des Umgangs mit anvertrauten Material und Geräten stehen im Vordergrund.
3. Verhaltensmerkmale zur Zusammenarbeit, zum Gruppenverhalten und zur Kritikverträglichkeit bilden die Oberkriterien in der dritten Gruppe. Hier wird die Fähigkeit zur Teamarbeit und Kooperation beurteilt, aber auch die Akzeptanz der Organisation wird eingestuft.
4. Die vierte Bewertungsgruppe befasst sich vorwiegend mit der Arbeitssystematik, der rationellen Arbeitsweise und der Ordnung am Arbeitsplatz. Die Gruppe enthält aber auch Fragen zur Eigeninitiative und zur Reaktion bei neuen Arbeitssituationen.
5. Im fünften Kriterienkatalog stehen Sorgfalt, Zuverlässigkeit und Arbeitsgüte im Vordergrund. Zu bewerten sind Arbeitsausführung und Arbeitsergebnis. Die Fähigkeit eigenverantwortlich zu arbeiten und andere anzuleiten sind weitere Bestandteile des Fragenkomplexes.
6. Belastbarkeit und Arbeitsmenge werden anschließend anhand von Kriterien zur Bereitschaft zur Erbringung von notwendiger Mehrarbeit, Übernahme von Arbeitsaufgaben, die nicht zur täglichen Arbeitserledigung zählen und Verhalten bei hohem Arbeitsdruck gemessen.
7. In der siebten Bewertungsgruppe richten sich die Beobachtungen auf das Informationsverhalten des zu Beurteilenden. Hier stehen sowohl die Fähigkeit zur Weitergabe von Informationen an Kollegen und andere Stellen als auch die Fähigkeit zur Informationsbeschaffung als Beobachtungskriterien

÷	Erfüllt <i>nicht</i> die an ihn/sie gestellten Anforderungen, im Durchschnitt des Jahres sind die Leistungen < 80%	Leistung und Verhalten entsprechen in den einzelnen Beobachtungspunkten der Bewertungsgruppen oder insgesamt nicht den Erwartungen. Der Mitarbeiter/die Mitarbeiterin ist trotz Unterstützung und Hilfestellung nicht für die Stelle geeignet.
⊙	Erfüllt die gestellten Anforderungen aus dem Arbeitsvertrag.	Leistung und Verhalten entsprechen den Erwartungen.
A	Übertrifft <i>teilweise</i> die an ihn/sie gestellten Anforderungen.	Leistung und Verhalten erfüllen die Ansprüche der Stelle und liegen teilweise über den Erwartungen (durchschnittlich ca. 20% über den Normalanforderungen aus dem Arbeitsvertrag).
B	Übertrifft <i>häufig</i> die an ihn/sie gestellten Anforderungen.	Leistung und Verhalten liegen häufig über den Erwartungen (durchschnittlich ca. 30% über den Normalanforderungen aus dem Arbeitsvertrag).
C	Übertrifft <i>meistens</i> die an ihn/sie gestellten Anforderungen.	Leistung und Verhalten liegen meistens über den Erwartungen (durchschnittlich ca. 40% über den Normalanforderungen aus dem Arbeitsvertrag).
D	Übertrifft <i>überwiegend</i> die an ihn/sie gestellten Anforderungen.	Leistung und Verhalten liegen deutlich über den Erwartungen (durchschnittlich ca. 50% über den Normalanforderungen aus dem Arbeitsvertrag).
E	Übertrifft <i>ständig</i> die an ihn/sie gestellten Anforderungen.	Leistungen und Verhalten liegen dauerhaft deutlich über den Erwartungen. Auch stellenübergreifende und besonders schwierige Anforderungen werden erfüllt. Der Mitarbeiter/die Mitarbeiterin ist grundsätzlich unterfordert (durchschnittlich mehr als 60% über den Normalanforderungen aus dem Arbeitsvertrag).

### Abb. 3: Bewertungsdimensionen

zur Verfügung. Auch die Bereitschaft, sich fachspezifisch in eine Diskussion einzubringen, ist dieser Bewertungsgruppe zugeordnet.

8. Im Fokus der vorletzten Bewertungsgruppe stehen die Einsetzbarkeit und die Bandbreite der Verwendungsmöglichkeiten. Die Bereitschaft zur Fortbildung ist hier als Beobachtungskriterium mit einbezogen.
9. Abschließend werden soziale Aspekte und Fragen zur Berufserfahrung, wie beispielsweise das Einsetzen der erworbenen Erfahrung im Fachgebiet beurteilt.



Die Beurteilung der einzelnen Beobachtungskriterien erfolgt über eine siebenstufige Einteilung. Anzumerken ist, dass sich die Leistungsbeurteilung im LBB-SYS auf einen durchschnittlichen Jahresarbeitsleistungswert von 80% bezieht. Die sieben Stufen werden im LBB-SYS wie in Abbildung 3 zusammengestellt definiert

Da den Bewertungsgruppen, wie eingangs erwähnt, eine unterschiedliche Anzahl von Beobachtungskriterien zugeordnet ist, erfolgt eine Normierung der Punktzahl. Damit finde der Grundgedanke Berücksichtigung, dass alle Bewertungsgruppen von gleicher Wichtigkeit sind und keine Bevorzugung von Stärken in bestimmten Bereichen erfolgen sollte.

## 4.2 Durchführung von Beurteilungen

Um das Verfahren zu objektivieren, wird die Leistungsbeurteilung von zwei Beurteilern unabhängig voneinander durchgeführt. Im vorgestellten Anwendungsfall waren dies zumeist der unmittelbare Vorgesetzte sowie dessen Vorgesetzter. Eine gültige Beurteilung eines Mitarbeiters setzt die Übereinstimmung beider Ergebnisse voraus. Bei einer bestimmten Abweichung der erreichten Punktzahl in einzelnen Bewertungsgruppen oder bei höheren Abweichungen in der Gesamtsumme wird ein Moderationsgespräch durchgeführt. Im Rahmen der Moderation finde eine Auseinandersetzung mit den Urteilen beider Beurteiler statt. Hierbei sind die Beurteiler gehalten, ihre Gründe für die Beurteilung offenzulegen, um so zu einem Konsens zu kommen. Gelingt dies nicht, wird eine unabhängige Leistungsbeurteilung von einem Drittbeurteiler durchgeführt.

Da wie oben diskutiert mit der Personalbeurteilung weitere Funktionen im personalwirtschaftlichen Kontext verbunden sind, also eine Differenzierung in der Entgeltfindung den zuvor geschilderten Problemdruck allein nicht lindern hilft, skizzieren wir kurz die personalwirtschaftlichen Anknüpfungspunkte des Instrumentariums.

Das LBB-SYS stellt nicht nur ein Bewertungsinstrument im engeren Sinne dar, sondern ist vielmehr eingebettet in einen Personalbeurteilungs- und Entwicklungsprozess. Es ist ein wichtiges Instrument der Personalführung und damit integraler Bestandteil des Personalmanagements, dessen Zielsetzung sowohl

- die Motivation der Mitarbeiter,
- die Förderung und Entwicklung der Mitarbeiter,
- eine anforderungsgerechte Besetzung von Stellen
- und organisationale Ziele

umfaßt.

### 4.3 Einführung des Leistungsbeurteilungs- und -bewertungssystems

Das Einführungskonzept des LBB-SYS beinhaltet die Bildung einer Projekthierarchie, um die Akzeptanz und gegebenenfalls eine Anpassung auf die organisatorischen Rahmenbedingungen und Zielsetzungen der jeweiligen Verwaltung zu sichern. Diese gliedert sich in eine Lenkungs- und Arbeitsgruppe. Die Lenkungsgruppe übernimmt alle vorbereitenden und steuernden Aufgaben, die zur Implementierung des Beurteilungssystems erforderlich sind. In der Arbeitsgruppe stehen die Durchführung und Analyse der von der Lenkungsgruppe übergebenen Problemstellung im Vordergrund. Die Projektgruppen bilden die einzelnen Funktionen in der Verwaltung ab und werden durch den externen Berater in allen Einzelschritten unterstützt. Die in den Projektgruppen konkretisierten Schritte zur Durchführung des Projektes lassen sich wie folgt gliedern:

- Information der zu beurteilenden Mitarbeiter (z. B. im Rahmen eines Informationstages)
- Festlegung der Beurteiler (jedem Mitarbeiter wird ein Erst- und Zweitbeurteiler zugeordnet)
- Einweisung und Schulung der Beurteiler
- Durchführung der Beurteilung
- Durchführung der Moderationsgespräche und gegebenenfalls eine Drittbeurteilung
- Mitteilung und Abstimmung der Beurteilungsergebnisse mit der Projektleitung/Betriebsleitung
- Führen der Mitarbeitergespräche

Für Vorgesetzte bedeutet die Personalbeurteilung einen ständigen Prozess der Beobachtung und Beurteilung von Mitarbeiterverhalten, das sich durch die schriftliche Personalbeurteilung am Ende einer Beurteilungsperiode konkretisiert. Dies erfordert einen permanenten steuernden und motivierenden Kontakt zwischen Mitarbeiter und Vorgesetztem, um organisationale Zielsetzungen für die einzelne Stelle transparent und nachvollziehbar zu gestalten. Die Mitarbeiterbeurteilung sollte daher lediglich das Schlussglied einer Vielzahl von regelmässig durchgeführten Beratungs- und Fördergesprächen zwischen Mitarbeiter und dem beurteilenden Vorgesetzten sein.

#### **4.4 Personalpolitische Leitlinien und Prinzipien der Personalführung im LBB-SYS**

Das LBB-SYS wurde als Beurteilungssystem mit Blick auf das Leitbild der Serviceorientierung öffentlicher Organisationen entworfen. Aus dieser Sicht ist die Basis der (Dienst-)Leistungserstellung in serviceorientierten Organisationen durch drei Komponenten geprägt:

- das Engagement der Mitarbeiter,
- wechselseitiges Vertrauen und individuelle Wertschätzung der Mitarbeiter sowie
- partnerschaftliche Zusammenarbeit und intensive Kommunikation.

Diese Leitlinien des LBB-SYS implizieren, dass die Leistungsbeurteilung nicht als isoliertes Instrumentarium zur Förderung der Motivation durch materielle Leistungsanreize eingesetzt werden kann und sollte. Von entscheidender Bedeutung für den Erfolg einer serviceorientierten Verwaltung ist der Umfang, in dem die Mitarbeiter ihre Leistungsfähigkeit entfalten und Engagement zeigen können. Die Organisation hat entsprechend die Aufgabe, die Leistungsfähigkeit und Leistungsbereitschaft ihrer Mitarbeiter zu unterstützen und zu fördern. Diese personalpolitische Leitlinie fordert die Einbettung des Leistungsbeurteilungs- und -bewertungssystems in organisatorische Gestaltungsmaßnahmen, in die Personalentwicklung und insbesondere in die Personalführung.

In der direkten Personalführung manifestieren sich Vertrauen, Wertschätzung, Partnerschaftlichkeit und Kommunikation. Die systematische Förderung der Kompetenzen und Fähigkeiten der Mitarbeiter ist aus Sicht des LBB-SYS eine zentrale Aufgabe der Führungskräfte. Partnerschaftliche Zusammenarbeit in Verbindung mit der Leistungsbeurteilung beinhaltet, dass Führungskräfte einerseits anspruchsvolle sowie realistische Ziele gemeinsam mit ihren Mitarbeitern vereinbaren und andererseits Rahmenbedingungen und Mittel der Leistungserstellung beeinflusse bzw. bereitstellen. Die Beurteilung der erreichten Leistungen soll im Rahmen des LBB-SYS von Partnerschaft und Fairness charakterisiert sein. Dazu gehört die kritische Betrachtung des Führungsverhaltens im Rahmen der Beurteilungsgespräche. Die Anerkennung von Leistung gilt als besonders wichtige Aufgabe der Personalführung.

### **5 Verlässlichkeit und Gültigkeit des LBB-SYS**

Die statistischen Auswertungen basieren auf drei Regelbeurteilungen (1997, 1999, 2000) der Mitarbeiter eines Baubetriebshofes einer kleinen Gemeinde in Baden-Württemberg mittels des LBB-SYS. Beurteilt wurden insgesamt 394 Arbeiter und

Vorarbeiter (1997:  $n = 126$ , 1999:  $n = 134$ , 2000:  $n = 134$ ). Das Durchschnittsalter der weit überwiegend männlichen Beurteilten (der Frauenanteil liegt zwischen 3,2 und 4,5%) liegt bei  $\bar{x} = 44$  Jahren ( $s = 10,9$ ) bei einer durchschnittlichen Betriebszugehörigkeit von  $\bar{x} = 11,5$  Jahren ( $s = 7,6$ ). Wichtige Aufgabenbereiche sind, gemessen in Anteilen der Beschäftigten, Grünpflanze (13,2%), interner Betrieb (12,7%), Straßenreinigung (11,7%) und das Führen von Kraftfahrzeugen (5,6%). Die Beurteiler ( $n = 50$ ) sind die Vorarbeiter und deren Vorgesetzte in der Hierarchie des Baubetriebshofes.

Auf dieser Datenbasis untersuchen wir das beschriebene Instrument hinsichtlich seiner Verlässlichkeit und, mit Bezug auf das Kriterium der Gültigkeit, seiner faktoriellen Struktur. Wir verwenden im Folgenden häufige die Fachterminologie der Sozialforschung und sprechen von Items (Beobachtungskriterien) und Skalen oder Faktoren (Bewertungsgruppen). Zur Analyse der Verlässlichkeit oder Reliabilität betrachten wir zunächst die zeitpunktbezogene interne Konsistenz der Beobachtungskriterien der neun Bewertungsgruppen (BG) des LBB-SYS. Ferner berichten wir Item- und Skalenkennwerte. Zur Analyse der Gültigkeit des Instrumentes untersuchen wir die faktorielle Struktur der Beobachtungskriterien. Mittels der explorativen und konfirmatorische Faktorenanalyse prüfen wir, inwieweit sich die Beobachtungskriterien auf Basis der gesammelten Daten den vorgegebenen Beobachtungsgruppen zuordnen lassen. Im dann folgenden Abschnitt fokussieren wir nochmals die Reliabilität. Wir analysieren die Inter-Rater-Reliabilität — also die Übereinstimmung der Beurteilungen zweier Beurteiler — und die Stabilität der Beurteilungen im Zeitverlauf, die sogenannte Retest-Reliabilität.

### 5.1 Zeitpunktbezogene Verlässlichkeit und Gültigkeit

Ein wichtiges Kriterium zur Beurteilung der zeitpunktbezogenen Verlässlichkeit einer Skala ist der zwischen 0 und 1 variierende Koeffizient Cronbach's  $\alpha$ . Hohe Koeffizienten drücken aus, dass die Beobachtungskriterien eine gemeinsame Beurteilungsskala bilden. Die internen Konsistenzen der angenommenen LBB-SYS Faktoren fallen ausnahmslos sehr hoch aus ( $\alpha \geq .90$ ). Diese Aussage trifft sowohl für den gesamten Datensatz als auch für die Untersuchung der einzelnen Wellen zu. Über alle Wellen betrachtet liegt der Koeffizient zwischen  $\alpha = .91$  (BG 7) und  $\alpha = .96$  (BG 9). Da die Skala BG 7 lediglich aus drei Items besteht, kann sogar dieser Wert als ausgezeichnet eingestuft werden. Betrachtet man die einzelnen Wellen, fällt auf, dass der errechnete Koeffizient mit wenigen Ausnahmen von Welle zu Welle abnimmt. Die sinkende interne Konsistenz lässt sich (sowohl isoliert als auch in Interaktion) auf zwei Ursachen zurückführen: erstens kann die wachsende Erfahrung der Beurteiler zur Folge haben, dass mit fortschreitender Anwendungsdauer des Instrumentariums differenzierter beurteilt wird. Zweitens

kann die Begründung im differenzierteren Leistungsverhalten der Beurteilten gesucht werden. Tabelle 1 gibt Auskunft über die internen Konsistenzen der einzelnen Skalen zu den unterschiedlichen Beurteilungszeitpunkten ( $t_1 \equiv 1997, t_2 \equiv 1999, t_3 \equiv 2000$ ) sowie im Mittel der Beurteilungsjahre.

Skala	$t_1$	$t_2$	$t_3$	Gesamt
BG 1	.95	.94	.93	.94
BG 2	.95	.94	.93	.95
BG 3	.94	.91	.92	.93
BG 4	.94	.93	.90	.93
BG 5	.94	.92	.91	.93
BG 6	.93	.93	.91	.92
BG 7	.91	.92	.90	.91
BG 8	.93	.93	.90	.93
BG 9	.97	.96	.95	.96

**Tab. 1:** Interne Konsistenzen der Skalen des LBB-SYS (Cronbachs  $\alpha$ )

Ein zweites Kriterium zur Beurteilung von Beurteilungsskalen ist die sogenannte Trennschärfe, d. h. die Korrelation eines Beobachtungskriteriums oder Items mit dem summierten Urteil einer Beobachtungsgruppe. Die in die neun Bewertungsgruppen einbezogenen Items weisen durchgängig sehr hohe Trennschärfen auf, welche sich zwischen  $r_{it} = .62$  und  $r_{it} = .93$  bewegen (vgl. Anhang A). Diese Zahlen bestätigen sich in homogenen Lage- und Streuungsparametern. Diese Analysen der Beobachtungsgruppen oder Subskalen deuten keinerlei Verbesserungsbedarf an, sondern weisen das LBB-SYS als sehr verlässliches Instrument aus.

Im Fall analytischer Beurteilungsverfahren, die aus mehreren Skalen zusammengesetzt sind, ist neben der Verlässlichkeit dieser Subskalen auch deren interne Struktur zu prüfen. Die auf empirischer Basis ermittelte Struktur der Subskalen des Systems sollte der postulierten Einteilung in neun Bewertungsgruppen entsprechen. Zur Analyse der faktoriellen Gültigkeit des LBB-SYS betrachten wir zunächst die Interkorrelationen der Bewertungsgruppen, skizzieren im zweiten Schritt die Ergebnisse einer explorativen Hauptkomponentenanalyse und stellen diesen Auswertungsschritt abschließend die Ergebnisse einer konfirmatorische Faktorenanalyse vor. Wir beschränken uns dabei, wegen der lediglich geringfügigen Unterschiede zwischen den einzelnen Beurteilungswellen, auf die Darstellung von Ergebnissen zum Gesamtdatensatz.

Tabelle 2 gibt Auskunft über die Interkorrelationen der Subskalen oder Bewertungsgruppen (BG) des LBB-SYS. Auffällig ist, dass diese hoch interkorrelieren.

Skala (Items)	BG1	BG2	BG3	BG4	BG5	BG6	BG7	BG8
BG1 (6)	–							
BG2 (6)	.92	–						
BG3 (6)	.77	.77	–					
BG4 (6)	.90	.93	.78	–				
BG5 (5)	.90	.94	.74	.94	–			
BG6 (6)	.85	.85	.80	.88	.84	–		
BG7 (3)	.79	.76	.66	.74	.74	.71	–	
BG8 (6)	.82	.81	.74	.83	.80	.88	.74	–
BG9 (5)	.21	.18	.27	.22	.21	.21	.13	.11

Anmerkungen:  $N = 394$ ; Produkt-Moment-Korrelationen  $r$ .

Alle Korrelationen signifikant ( $p \leq .05$ ).

Das Signifikanzniveau  $p \leq .01$  wird erreicht bei  $r \geq .12$ .

**Tab. 2:** Skalenkennwerte und Interkorrelationen der LBB-SYS Skalen

Ausnahmen bilden die Skalen BG 3, BG 7 und BG 9, die mit den anderen Skalen schwächer korrelieren. Als „Catch-all“-Bewertungsgruppe erweist sich BG 1. Inhaltlich bedeutet dies, dass Mitarbeiter die bezüglich einer Bewertungsgruppe gut bzw. weniger gut bewertet werden auch in (fast) allen anderen Bewertungsgruppen hohe bzw. niedrige Punktwerte erhalten. Auf Basis des vorliegenden Datensatzes ist eine faktorielle Trennung in die vorgegebenen Bewertungsgruppen also nur eingeschränkt möglich.

Diese Aussage stützt sich auch auf die Ergebnisse einer explorativen Hauptkomponentenanalyse über alle Items aller Wellen, deren Ergebnisse im Anhang B eingesehen werden können. Diese Befunde zeigen, dass sich die neun-faktorielle Struktur des LBB-SYS nur mit Einschränkungen rekonstruieren lässt. Die explorative Analyse basiert im Unterschied zur postulierten Struktur nur auf fünf Faktoren. Darüber hinaus weist eine Reihe von Items erhebliche Fehlloadungen, d. h. Korrelationen zu anderen Hauptkomponenten, auf. Die Items der Skalen BG 7, BG 9 und BG 3 lassen sich in der genannten Reihenfolge den Komponenten 3, 4 und 5 zuordnen, zumindest die Komponenten 4 und 5 sind eindeutig interpretierbar. Die restlichen Items entfallen auf die Komponenten 1 und 2. Die Hauptkomponente 1 bindet einen Varianzanteil von annähernd 30% bei einem gesamten prozentualen Anteil von 77,5% erklärter Varianz. Die Gültigkeit der postulierten 9-faktoriellen Struktur des LBB-SYS ist also zumindest vorerst in Frage zu stellen.

Zur weiteren Klärung der Faktorenstruktur führen wir eine konfirmatorische Faktorenanalyse mittels LISREL durch (Jöreskog/Sörbom 1994). Zur Mo-

dellschätzung finde Produkt-Moment-Korrelationen sowie die Methode der ungewichteten kleinsten Quadrate Anwendung. Zwischen den spezifizierten Faktoren werden Korrelationen zugelassen, die Messfehler werden als unkorreliert betrachtet. Wir untersuchen drei verschiedene Modellvarianten:

- 1-Faktor-Modell (alle 9 BG bilden zusammen einen Faktor)
- 4-Faktoren-Modell (Faktoren sind BG 1/2/4/5/6/8, BG 3, BG 7 und BG 9)
- 9-Faktor-Modell (alle BG bilden jeweils eigene Faktoren)

Das 1-Faktor-Modell trägt der starken ersten Hauptkomponente Rechnung, das 4-Faktoren-Modell repräsentiert das Resultat der explorativen Faktorenlösung. Das 9-Faktoren-Modell entspricht der angenommenen Struktur des LBB-SYS. Tabelle 3 zeigt die Passungsparameter der Modelle auf:

Modell				Passungskennwerte					$\Delta\chi^2$
	$df$	$\chi^2$	$\chi^2/df$	GFI	AGFI	RMR	NFI	PNFI	
1-F.-Modell	1127	4507.46	4.00	.98	.97	.10	.98	.93	
4-F.-Modell <sup>a</sup>	1121	1445.56	1.29	.99	.99	.06	.99	.95	3061.90***
9-F.-Modell	1091	1150.86	1.06	.99	.99	.05	.99	.92	294.70***
Anmerkungen:	<sup>a</sup> Faktoren sind: BG1/2/4/5/6/8, sowie BG3, BG7 und BG9								
	*** $p \leq .001$ .								

**Tab. 3:** Passungsparameter der Alternativmodelle

Die Tabelle weist verschiedene Kennwerte aus, welche die Anpassungsgüte der Modelle, d. h. den sogenannten Fit zwischen Modell und empirischer Korrelationsmatrix, ausdrücken. Abnehmende  $\chi^2$ -Werte signalisieren eine Modellverbesserung. Das Verhältnis von  $\chi^2$ -Werten zu den Freiheitsgraden ( $df$ ) sollte laut Konvention den Wert 5 nicht überschreiten. Das Verhältnis der Differenzen der  $\chi^2$ -Kennzahlen ( $\Delta\chi^2$ ) zu denen der Freiheitsgrade ( $\Delta df$ ) kann zur Bewertung der Modellverbesserung herangezogen werden. Für die ausgewiesenen Fit Indizes (GFI, AGFI, NFI und PNFI) gilt die Faustregel, dass steigende (gegen 1 strebende) Werte eine bessere Passung des Modells an die beobachteten Werte bedeuten. Das gleiche gilt für sinkende (gegen 0 tendierende) RMR-Werte. Fit-Indizes  $> .90$  signalisieren eine gute Modellanpassung (Bentler/Bonett 1980).

Die berichteten Parameter zeigen, dass das 9-Faktoren-Modell den vergleichsweise besten Fit aufweist. Es ergibt sich eine signifikant sukzessive Verbesserung der Modellpassung bis zum 9-Faktoren-Modell ( $p \leq .001$ ). Die ermittelten Fit-Indices und RMR-Werte können als Indizien für einen sehr guten Fit des 9-Faktoren-Modells an die vorliegenden Daten erachtet werden. Im Vergleich zum

4-Faktoren-Modell sinken allerdings beim postulierten 9-Faktoren-Modell Freiheitsgrade ( $df$ ) wie  $\chi^2$ -Werte nur geringfügig. Inhaltlich bedeutet dies, dass das Modell trotz höherer Komplexität nur einen relativ geringen Zugewinn an Erklärungskraft aufweist. Diese Entwicklung drückt auch der parsimonisch korrigierten Fit Index (PNFI) aus. Dieser Parameter fällt für das 9-Faktoren-Modell niedriger aus als für das 4-Faktoren-Modell.

Die skizzierten Befunde zur faktoriellen Gültigkeit zeigen auf Basis der hier zu Grunde liegenden Daten an, dass das LBB-SYS vereinfacht werden könnte. Möglicherweise lassen sich die inhaltlich eng verwandten Bewertungsgruppen BG 1, BG 2, BG 4, BG 5, BG 6 und BG 8 zu weniger kohärenten Faktoren bündeln. Es ist allerdings anzumerken, dass die hier berichteten Befunde stichprobenabhängig sein können. Inhaltlich bedeutet dies, dass die Homogenität der Urteile in den genannten Bewertungsgruppen möglicherweise auch auf die Gleichartigkeit der Aufgaben bei den Mitarbeitern im untersuchten Baubetriebshof zurückzuführen sind. Die Entscheidung über die Revision des Beurteilungsinstrumentes in Richtung einer Zusammenfassung von Bewertungsgruppen ist von einer weitergehenden Analyse abhängig zu machen, die auf Daten über unterschiedliche Tätigkeitsbereiche in öffentlichen Organisationen basiert. Zumindest für das derzeitige Haupteinsatzgebiet des LBB-SYS, die Leistungsbeurteilung in kommunalen Betriebshöfen, ist allerdings eine Straffung der Bewertungsgruppen angezeigt.

## 5.2 Übereinstimmungs-Verlässlichkeit

Das LBB-SYS beinhaltet grundsätzlich die Beurteilung des leistungsbezogenen Arbeitsverhaltens durch zwei unabhängige Beurteiler. Dies bietet die Möglichkeit, die Übereinstimmung der Beurteilungen zu analysieren. Die Analyse der Übereinstimmungen in den Leistungsbeurteilungen ist mit Blick auf das System nicht nur von analytischem, sondern auch von praxeologischem Wert. Die Analyse indiziert, ob es häufig zu Moderationen im Fall nicht übereinstimmender Urteile kommt. Zu beachten ist dabei, dass nur einer der Beurteiler aus dem direkten Arbeitsumfeld des Beurteilten kommt.

Tabelle 4 zeigt die mittlere Korrelation der Übereinstimmungen beider Beurteiler (Inter-Rater-Reliabilität, I-R-R) zu den Beurteilungszeitpunkten sowie die Schwankungsbreite der Beurteilungen. Die mittlere Inter-Rater-Reliabilität liegt durchweg im positiven Bereich. Allerdings nimmt die Ähnlichkeit zwischen den Urteilen der einzelnen Beurteiler mit steigenden Anwendungsdauer des LBB-SYS ab. Die durchschnittliche Inter-Rater-Reliabilität liegt über alle Wellen betrachtet bei  $r = .55$  und nimmt von  $r = .60$  (Welle 1) über  $r = .59$  (Welle 2) bis zu  $r = .47$  (Welle 3) ab.

Die Abschwächung der Inter-Rater-Übereinstimmung im Zeitverlauf kann



	$t_1$	$t_2$	$t_3$	Gesamt
Mittlere I-R-R	.60	.59	.47	.55
Minimum	-.04	-.08	-.17	-.17
Maximum	.98	.97	.96	.98

**Tab. 4:** Kennzahlen zur Inter-Rater-Reliabilität (IRR)

verschiedene Ursachen haben. Neben der bereits angesprochenen Differenzierung im Urteilsvermögen der Beurteiler und im Leistungsverhalten der Beurteilten ist auch eine Kombination von Urteils- und Verhaltenseffekt möglich. Insofern nur einer der Beurteiler dauerhaft im Arbeitsgebiet des Beurteilten beschäftigt ist, können Veränderungen im Leistungsverhalten von diesem genauer registriert werden als vom Zweitbeurteiler. Entsprechend sollten die Beurteilungen, so unsere Hypothese, die Urteile des Zweitbeurteilers im Zeitverlauf eine größere Konstanz aufweisen als die des Erstbeurteilers. Auf Basis des vorliegenden Datensatzes ist eine weitergehende Prüfung dieser Hypothese, wegen der geringen Anzahl von Zweitbeurteilern, nicht möglich. Sie sollte daher auf größerer Datenbasis geprüft werden. Gegebenenfalls sind neben den jährlichen Regelbeurteilungen Zwischenbeobachtungen wie eine Teilnahme an den Beratungs- und Förderungsgesprächen für die Zweitbeurteiler einzuführen.

### 5.3 Retest-Verlässlichkeit

Während zwischen den beiden ersten Regelbeurteilungen rund 24 Monate liegen, liegen die Wellen 2 und 3 nur etwa 12 Monate auseinander. Entsprechend fällt die für alle Einzelitems gemittelte Retest-Reliabilität mit  $r = .50$  zwischen den Wellen 1 und 2 (Intervall  $t_1 - t_2$ ) niedriger aus als mit  $r = .67$  zwischen den Wellen 2 und 3 ( $t_2 - t_3$ ). Tabelle 5 gibt nähere Auskunft zu den ermittelten Retest-Reliabilitäten im Mittel der Einzelitems und der Skalensummen.

	Einzelitems			Skalensummen		
	$t_1 - t_2$	$t_2 - t_3$	$t_1 - t_3$	$t_1 - t_2$	$t_2 - t_3$	$t_1 - t_3$
Mittlere R-R	.50	.67	.48	.57	.79	.56
Minimum	.30	.52	.25	.50	.71	.46
Maximum	.65	.83	.69	.62	.84	.68

**Tab. 5:** Kennzahlen zur Retest-Reliabilität (RR)

Die Ergebnisse des Zeitpunktes  $t_0$  sind also wie erwartet nur bedingt dazu

geeignet, die Ergebnisse des Zeitpunktes  $t_2$  zu prognostizieren, die Ergebnisse des Zeitpunktes  $t_2$  jedoch stellen eine recht gute Grundlage dar, um diejenigen des Zeitpunktes  $t_3$  zu schätzen.

## 6 Resümee

Ein zentrales personalwirtschaftliches Steuerungselement der Reform öffentlicher Organisationen hin zu einer serviceorientierten Verwaltung ist der Einsatz materieller Anreizsysteme. Die Vergabe von Leistungszulagen hat jedoch Beurteilungssysteme zur Voraussetzung, die dem organisatorischen Kontext angepasst sind, seitens der Führungskräfte und Mitarbeiter auf Akzeptanz stoßen und darüber hinaus den Gütekriterien der empirischen Sozialforschung genügen. Das hier vorgestellte LBB-SYS ist ein Leistungsbeurteilungs- und -bewertungssystem, das auf die Ansprüche öffentlicher Organisationen zugeschnitten ist.

Im Kern handelt es sich um ein analytisches, neundimensionales Beurteilungsinstrument, das der Regelbeurteilung der Mitarbeiter dient. Die Beurteilungsergebnisse sind Grundlage der Gewährung von Leistungszulagen. Die Beurteilungen erfolgen im Anschluss an den Implementierungsprozess systematisch durch jeweils zwei Beurteiler. Das System sieht Moderationen im Fall von Beurteilerunstimmigkeiten vor und ist in eine Führungskonzeption, gekennzeichnet durch regelmäßige Beratungs- und Förderungsgespräche eingebettet.

Die statistische Analyse des Beurteilungsinstrumentes, basierend auf den Daten von drei aufeinander folgenden Regelbeurteilungen in einem kommunalen Baubetriebshof, zeigt, dass es sich um ein reliables, d. h. verlässliches Beurteilungsinstrument handelt. Das Instrument weist hohe zeitpunktbezogene Kennziffern der Reliabilität auf. Verfahrensbedingt schwächer sind die Gütekriterien der Übereinstimmung mehrerer Beurteiler bzw. der Urteile über mehrere Messzeitpunkte. Dies ist im ersten Fall unter Umständen darauf zurückzuführen, dass nur einer der Beurteiler als direkter Vorgesetzter im Arbeitsumfeld des Beurteilten tätig ist. Entsprechend sieht das System eine Abstimmungsrunde und gegebenenfalls eine Drittbewertung bei hoch differierenden Urteilen vor. Die Verlässlichkeit im Zeitverlauf ist niedriger als die zeitpunktbezogene Reliabilität, weil das Leistungsverhalten der Beurteilten variieren kann.

Trotz dieser positiven Befunde zur Verlässlichkeit weisen die Ergebnisse auf Verbesserungsmöglichkeiten hin. Die Analysen zur faktoriellen Gültigkeit zeigen, dass sich auf Basis der empirischen Daten die postulierten neun Beurteilungsgruppen nicht hinreichend trennen lassen. Dies ist möglicherweise auch darauf zurückzuführen, dass die Variation der Aufgaben im hier betrachteten Arbeitsgebiet recht gering ist. Zumindest für den Beurteilungsbereich der gewerblichen Arbeiten ist jedoch eine Straffung auf weniger Beurteilungsdimensionen angezeigt.

---

## Literatur

- Adams, J. S. (1963): Toward an understanding of inequity. In: *Journal of Abnormal and Social Psychology*, vol. 63, pp. 422–436.
- Becker, F. (1992): Grundlagen betrieblicher Leistungsbeurteilungen. Leistungsverständnis und -prinzip, Beurteilungsproblematik und Verfahrensprobleme. Stuttgart.
- Bentler, P. M./Bonett, D. G. (1980): Significant tests and goodness of fit in the analysis of covariance structures. In: *Psychological Bulletin*, vol. 88, pp. 588–606.
- Boerger, M. (1982): Mitarbeiterbeurteilung. In: Streutz, H. (Hg.): *Handwörterbuch der Verwaltung und Organisation*. Köln. S. 275–284.
- Gaugler, E./Kolvenbach, H./Lay, G./Ripke, M./Schilling, W. (1978): *Leistungsbeurteilung in der Wirtschaft: Verfahren und Anwendungen in der Praxis*. Baden-Baden.
- Jöreskog, K. G./Sörbom, D. (1994): *LISREL VIII. User's Reference Guide*. Mooresville, Ind.
- KGSt (1993): *Das Neue Steuerungsmodell: Begründung — Konturen — Umsetzung*. Arbeitspapier, Nr. 5, Köln, Kommunale Geschäftsstelle für Verwaltungsvereinfachung.
- KGSt (1996): *Personalentwicklung im Neuen Steuerungsmodell: Anforderungen an vorrangige Zielgruppen*. Arbeitspapier, Nr. 6, Köln, Kommunale Geschäftsstelle für Verwaltungsvereinfachung.
- Klages, H. (1995): Verwaltungsmodernisierung durch ‚Neue Steuerung‘? In: *Archiv für Kommunalwissenschaften*, 34. Jg., S. 203–228.
- Lessmann, K. G. (1980): *Personalbeurteilung als Instrument der Führungskräfteentwicklung*. Dissertation, Universität Essen.
- Liebel, H. J./Oechsler, W. A. (1992): *Personalbeurteilung: Neue Wege zur Bewertung von Leistung, Verhalten und Potential*. Wiesbaden.
- Lueger, G. (1992): *Die Bedeutung der Wahrnehmung bei der Personalbeurteilung: zur psychischen Konstruktion von Urteilen über Mitarbeiter*. München, Mering.
- Matiaska, W. (1996): *Statistische Datenanalyse mit Mikrocomputern*. 2. Aufl. München, Wien.
- Mungenast, M. (1990): *Grenzen merkmalsorientierter Einstufungsverfahren und ihre mögliche Überwindung durch zielorientierte Leistungsbeurteilungsverfahren*. München.
- Nienhüser, W./Becker, C. (2000): *Betriebliche Personalforschung: eine problemorientierte Einführung*. Berlin.
- Steinort, U. (1996): *Neue Wege bei der Leistungsbeurteilung im Öffentlichen Dienst*. In: Antoni, C. H./Eyer, E./Kutscher, J. (Hg.): *Das flexible Unternehmen: Arbeitszeit, Gruppenarbeit, Entgeltsysteme*. Wiesbaden.
- Tondorf, K. (1995): *Leistungszulagen als Reforminstrument? Neue Lohnpolitik zwischen Sparzwang und Modernisierung*. Berlin.
- Uttlinger, S./Breier, A./Kiefer, K.-H./Hoffmann, H. (2000): *Bundes-Angestelltentarifvertrag — BAT (Bund, Länder, Gemeinden) Kommentar*. 10. Aufl. München.

## A Statistiken der Subskalen des LBB-SYS

Bewertungs- gruppe	Item	$\bar{x}$	$s$	$\alpha$ ohne Item $i$	korrigiertes $r_{it}$
BG 1 ( $\alpha = .94$ )	1	2.764	1.265	.938	.784
	2	2.688	1.343	.928	.868
	3	2.503	1.356	.934	.818
	4	2.563	1.312	.934	.811
	5	2.355	1.320	.931	.843
	6	2.296	1.349	.931	.844
BG 2 ( $\alpha = .95$ )	1	2.366	1.351	.934	.833
	2	2.505	1.248	.932	.845
	3	2.135	1.253	.935	.824
	4	2.468	1.317	.928	.881
	5	2.876	1.375	.930	.864
	6	2.750	1.307	.945	.744
BG 3 ( $\alpha = .93$ )	1	2.766	1.205	.911	.804
	2	2.446	1.187	.916	.782
	3	2.348	1.229	.905	.853
	4	2.654	1.162	.907	.800
	5	2.972	1.186	.916	.771
	6	2.020	1.260	.924	.714
BG 4 ( $\alpha = .93$ )	1	2.611	1.203	.915	.811
	2	2.753	1.324	.918	.786
	3	2.496	1.268	.919	.776
	4	2.670	1.343	.913	.809
	5	2.223	1.226	.917	.778
	6	1.901	1.283	.916	.802
BG 5 ( $\alpha = .93$ )	1	2.746	1.232	.901	.841
	2	2.703	1.217	.901	.846
	3	2.490	1.224	.904	.830
	4	1.849	1.355	.936	.675
	5	2.802	1.367	.899	.852
BG 6 ( $\alpha = .92$ )	1	2.787	1.361	.915	.739
	2	2.565	1.282	.917	.728
	3	2.459	1.333	.917	.721
	4	2.490	1.429	.899	.856
	5	2.523	1.293	.906	.822
	6	2.076	1.311	.905	.820

Tab. 6: Statistiken der Subskalen

Bewertungs- gruppe	Item	$\bar{x}$	$s$	$\alpha$ ohne Item $i$	korrigiertes $r_{it}$
BG 7 ( $\alpha = .91$ )	1	2.365	1.305	.849	.848
	2	2.188	1.378	.894	.795
	3	2.424	1.310	.872	.820
BG 8 ( $\alpha = .93$ )	1	2.805	1.246	.910	.807
	2	2.642	1.230	.907	.830
	3	2.188	1.219	.904	.857
	4	2.355	1.226	.907	.828
	5	1.952	1.521	.940	.624
	6	2.241	1.234	.908	.825
BG 9 ( $\alpha = .96$ )	1	1.598	1.536	.950	.923
	2	1.558	1.541	.950	.924
	3	1.419	1.455	.953	.908
	4	1.343	1.611	.962	.854
	5	1.542	1.602	.958	.874

**Tab. 6:** Statistiken der Subskalen (Fortsetzung)

## B Ergebnisse der Explorativen Hauptkomponentenanalyse

Bewertungs- gruppe	Item	Hauptkomponenten				
		$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
BG 1	1	.470	.409	.423	.183	.313
	2	.594	.309	.506	.087	.236
	3	.617	.176	.567	-.002	.133
	4	.597	.470	.282	.050	.265
	5	.725	.301	.309	.099	.179
	6	.675	.246	.486	.105	.132
BG 2	1	.605	.363	.504	.056	.193
	2	.695	.308	.266	.085	.311
	3	.689	.228	.395	.068	.229
	4	.704	.330	.374	.096	.217
	5	.761	.252	.341	-.043	.210
	6	.728	.222	.127	.116	.235
BG 3	1	.276	.386	.136	.161	.651
	2	.346	.218	.188	.122	.726
	3	.381	.348	.226	.160	.658
	4	.362	.275	.225	.116	.679
	5	.400	.388	.238	.212	.491
	6	.536	.427	.277	.047	.354
BG 4	1	.750	.289	.152	.163	.265
	2	.726	.311	.149	.067	.180
	3	.696	.293	.146	.188	.202
	4	.724	.315	.367	.035	.152
	5	.535	.498	.325	.023	.279
	6	.565	.413	.502	.081	.147
BG 5	1	.792	.266	.097	.133	.311
	2	.796	.288	.120	.111	.312
	3	.786	.188	.324	.033	.116
	4	.491	.408	.526	.100	.075
	5	.758	.292	.322	.053	.127
BG 6	1	.287	.625	.238	.078	.281
	2	.266	.729	.209	.058	.280
	3	.728	.271	.227	.232	.187
	4	.593	.535	.237	.008	.242
	5	.654	.443	.259	.061	.220
	6	.474	.611	.270	.102	.239

**Tab. 7:** Varimax-rotierte Ladungsmatrix

Bewertungs- gruppe	Item	Hauptkomponenten				
		$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
BG 7	1	.344	.236	.704	.051	.296
	2	.354	.320	.696	.018	.192
	3	.439	.227	.638	.024	.241
BG 8	1	.508	.574	.210	-.053	.187
	2	.441	.685	.165	-.011	.252
	3	.530	.687	.244	-.030	.170
	4	.441	.640	.279	-.021	.220
	5	.192	.507	.562	.163	.053
	6	.410	.591	.386	.005	.246
BG 9	1	.144	.001	.062	.941	.038
	2	.105	.009	.057	.944	.032
	3	.103	.032	.070	.919	.067
	4	.006	.060	.023	.863	.134
	5	.062	.031	-.007	.884	.109
Eigenwerte		14.844	7.471	5.699	4.609	4.145
Varianz in %		30.3	15.2	11.6	9.4	8.5
kumuliert		30.3	45.5	57.2	66.6	75.0

**Tab. 7:** Varimax-rotierte Ladungsmatrix (Fortsetzung)

**Verzeichnis der  
Berichte der Werkstatt für Organisations- und Personalforschung e.V.:**

- 01. Weller, I./Steffen, E. 2000:** Ergebnisse einer Mitarbeiterbefragung bei der Lynx Consulting Group/Bielefeld. Berlin.
- 02. Bendel, K. 2000:** Zufriedenheit von Nutzerinnen und Nutzern mit ambulanten Pflegedienstleistungen. Forschungsbericht. Berlin.
- 03. Bendel, K./Matiaske, W./Schramm, F./Weller, I. 2000:** „Kundenzufriedenheit“ bei ambulanten Pflegedienstleistern. Bestandsaufnahme und Vorschläge für ein stress-theoretisch fundiertes Messinstrument. Berlin.
- 04. Holtmann, D./Matiaske, W./Möllenhoff, D./Weller, I. 2001:** Leistungsbeurteilung im öffentlichen Dienst. Zur Validierung des Leistungsbeurteilungs- und -bewertungssystems LBB-SYS. Berlin.
- 05. Martin, A./Purwin, J. 2001:** Soziale Fähigkeiten in Arbeitsgruppen. Eine empirische Studie zur Ermittlung der Kooperationsfähigkeit. Berlin.
- 06. Weller, I. 2001:** Fluktuationsneigung und Commitment. Eine empirische Betrachtung bei F&E-Mitarbeitern. Berlin.
- 07. Matiaske, W./Holtmann, D./Weller, I. 2002:** Anforderungen an Spitzenführungskräfte. Retrospektive und Perspektive: Eine empirische Untersuchung. Berlin.
- 08. Jütte, W./Matiaske, W. 2002:** Regionale Weiterbildungsnetzwerke. Eine exemplarische Analyse. Berlin.
- 09. Holtmann, D./Matiaske, W./Weller, I. 2002:** Transparenz und Kommunikation als Erfolgsfaktoren von Leistungsbeurteilungen im öffentlichen Dienst. Vorstellung eines Forschungsprojektes. Berlin.
- 10. Erbel, C. 2003:** Personalmanagement, Mitarbeiterverhalten und Kundenzufriedenheit im Dienstleistungskontakt. Eine empirische Analyse. Berlin.