

WESTFÄLISCHE  
WILHELMS-UNIVERSITÄT  
MÜNSTER

# Rezeption und Nutzung von Ergebnissen aus Schulinspektionen

Umgang von Schulen und Schulaufsicht mit Ergebnissen  
der Qualitätsanalyse NRW

Jana Kleine



# **Rezeption und Nutzung von Ergebnissen aus Schulinspektionen**

-

**Umgang von Schulen und Schulaufsicht mit  
Ergebnissen der Qualitätsanalyse NRW**

Inaugural-Dissertation  
zur Erlangung des Doktorgrades  
der  
Philosophischen Fakultät  
der  
Westfälischen Wilhelms-Universität  
zu  
Münster (Westf.)

Jana Kleine, geb. Sticka

2012

Tag der mündlichen Prüfung: 13.12.2012

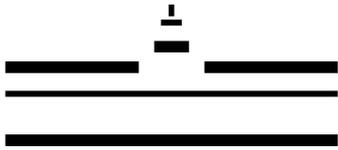
Dekan der Philosophischen Fakultät: Prof. Dr. Christian Pietsch

Erstgutachter: Prof. Dr. Ewald Terhart

Zweitgutachterin: Prof. Dr. Sabine Gruehn

**Jana Kleine**

## **Rezeption und Nutzung von Ergebnissen aus Schulinspektionen**



**WESTFÄLISCHE  
WILHELMS-UNIVERSITÄT  
MÜNSTER**

Wissenschaftliche Schriften der WWU Münster

## **Reihe VI**

**Band 14**

**Jana Kleine**

# **Rezeption und Nutzung von Ergebnissen aus Schulinspektionen**

Umgang von Schulen und Schulaufsicht mit Ergebnissen  
der Qualitätsanalyse NRW



MV WISSENSCHAFT

## Wissenschaftliche Schriften der WWU Münster

herausgegeben von der Universitäts- und Landesbibliothek Münster

<http://www.ulb.uni-muenster.de>

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Dieses Buch steht gleichzeitig in einer elektronischen Version über den Publikations- und Archivierungsserver der WWU Münster zur Verfügung.

<http://www.ulb.uni-muenster.de/wissenschaftliche-schriften>

Jana Kleine

„Rezeption und Nutzung von Ergebnissen aus Schulinspektionen. Umgang von Schulen und Schulaufsicht mit Ergebnissen der Qualitätsanalyse NRW“

Wissenschaftliche Schriften der WWU Münster, Reihe VI, Band 14

© 2015 der vorliegenden Ausgabe:

Die Reihe „Wissenschaftliche Schriften der WWU Münster“ erscheint im Verlagshaus Monsenstein und Vannerdat OHG Münster

[www.mv-wissenschaft.com](http://www.mv-wissenschaft.com)

Dieses Werk ist unter der Creative-Commons-Lizenz vom Typ 'CC BY 3.0 DE'

lizenziert: <http://creativecommons.org/licenses/by/3.0/de/>

Von dieser Lizenz ausgenommen sind Abbildungen, welche sich nicht im Besitz der Autorin oder der ULB Münster befinden.



ISBN 978-3-8405-0129-6

(Druckausgabe)

URN urn:nbn:de:hbz:6-60329449488

(elektronische Version)

direkt zur Online-Version:

© 2015 Jana Kleine

Alle Rechte vorbehalten

Satz:

Jana Kleine

Titelbild:

*Magnifying glass on green background*

Stock Foto, Bildnr. 146941343, Urheber: Laborant

[http://www.shutterstock.com/pic-146941343/stock-](http://www.shutterstock.com/pic-146941343/stock-photo-magnifying-glass-on-green-background.html)

[photo-magnifying-glass-on-green-background.html](http://www.shutterstock.com/pic-146941343/stock-photo-magnifying-glass-on-green-background.html)

Umschlag:

MV-Verlag

Druck und Bindung:

MV-Verlag



Die vorliegende Arbeit ist im Jahr 2012 fertiggestellt worden. Bis zum Erscheinungszeitpunkt wurde das Forschungsfeld um wichtige Befunde angereichert und weiterentwickelt (s. z. B. van Ackeren/Heinrich/Thiel 2013; Keune 2014). Diese neueren Arbeiten werden in dieser Publikation nicht berücksichtigt.

Zudem ist zu beachten, dass die Arbeit auf der Grundlage des „alten“ Instrumentariums und Verfahrens der *Qualitätsanalyse* Nordrhein-Westfalen vor deren Neuausrichtung im Jahr 2013 entstanden ist.

Ausdrücklich danken möchte ich dem Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen und insbesondere Sabine Müller und Wulf Homeier, mit den vorhandenen Daten der Qualitätsanalyse eine eigenverantwortliche Analyse vornehmen zu dürfen. Zudem danke ich den Schulabteilungsleitern und schulfachlichen Mitarbeitern der nordrhein-westfälischen Schulaufsicht für ihre Bereitschaft, an einer Befragung teilzunehmen und insbesondere Walter Ruhwinkel, Rainer Hofmann, Wolfgang Weber und Brunhilde Jacobi für ihre freundliche Unterstützung.

Vor allem Herrn Prof. Dr. Ewald Terhart und Frau Prof. Dr. Sabine Gruehn danke ich für ihre kontinuierliche Unterstützung und wissenschaftliche Beratung bei der Erstellung meiner Dissertation.



# INHALTSVERZEICHNIS

---

<b>I</b>	<b>Einleitung</b> .....	<b>1</b>
1.	<b>Ausgangspunkt, Zielsetzung und Aufbau der Arbeit</b> .....	<b>1</b>
<b>II</b>	<b>Theoretischer Teil</b> .....	<b>5</b>
2.	<b>Schulqualität und Schulreform</b> .....	<b>5</b>
2.1.	Qualität von Schule und Unterricht .....	5
2.1.1.	Qualitätsbegriff im Bildungsbereich.....	5
2.1.2.	Merkmale guter Schulen und guten Unterrichts .....	7
2.1.3.	Modell zur Schulqualität.....	9
2.2.	Steuerung und Evaluation .....	11
2.2.1.	Neue Steuerung im Bildungswesen .....	11
2.2.2.	Evaluation und ihre Typisierung.....	14
2.2.3.	Externe Evaluationsinstrumente .....	17
3.	<b>Schulinspektion</b> .....	<b>20</b>
3.1.	Schulinspektion im europäischen Ausland.....	20
3.1.1.	Inspectie van het Onderwijs in den Niederlanden .....	22
3.1.2.	Inspection of schools in England .....	25
3.1.3.	Vergleichende Gegenüberstellung der niederländischen und englischen Inspektionsverfahren.....	30
3.2.	Schulinspektion in Deutschland.....	32
3.2.1.	Definition .....	32
3.2.2.	Ziele und Funktionen .....	33
3.2.3.	Qualitätsverständnis der Schulinspektion.....	36
3.2.4.	Schulinspektion in den deutschen Bundesländern.....	40
3.2.5.	Kritik an der Schulinspektion .....	46
3.3.	Die Qualitätsanalyse an Schulen in Nordrhein-Westfalen .....	50
3.3.1.	Zielsetzungen und normativer Rahmen .....	50
3.3.2.	Institutionalisierung und Arbeitsstand .....	52
3.3.3.	Schulauswahl und (zeitlicher) Verlauf .....	54

---

3.3.4.	Verfahren und Instrumente.....	55
3.3.5.	Rückmeldung und Ergebnismutzung .....	57
3.3.6.	Zentrale Befunde der bisherigen Qualitätsanalysen.....	58
3.3.7.	Zusammenfassung .....	60
3.4.	Forschungsstand zur Schulinspektion.....	61
3.4.1.	Systematisierungen der Schulinspektionsforschung .....	62
3.4.2.	Forschungstyp I: Einstellungen und Akzeptanz der Schulinspektion .....	65
3.4.3.	Forschungstyp II: Auswirkungen der Schulinspektion auf die einzelschulische Schulentwicklung.....	69
3.4.4.	Forschungstyp III: Einfluss der Schulinspektion auf die Schülerleistungen .....	77
3.4.5.	Forschungstyp IV: Messgüte der Schulinspektions- instrumente .....	80
3.4.6.	Wirkungsmodell zur Schulinspektion .....	81
3.4.7.	Zusammenfassung.....	84
<b>4.</b>	<b>Rezeption und Nutzung von Ergebnissen der Schulinspektion .....</b>	<b>87</b>
4.1.	Rolle der Schulaufsicht im Zusammenhang mit der Schulinspektion ....	89
4.2.	Zielvereinbarungen im Anschluss an die Schulinspektion .....	95
4.2.1.	Definition von Ziel und Zielvereinbarung.....	95
4.2.2.	Die ökonomische Perspektive: Management by Objectives.....	97
4.2.3.	Die psychologische Perspektive: Rubikon-Modell und Zielsetzungstheorie.....	97
4.2.4.	Zielmerkmale.....	100
4.2.5.	Zielvereinbarungen im Bildungsbereich .....	102
4.2.6.	Zielvereinbarungen nach der Schulinspektion – am Beispiel der Qualitätsanalyse NRW .....	103
4.3.	Rezeptionsforschung .....	108
4.3.1.	Modelle zur Rezeption und Nutzung von Ergebnisrückmeldungen.....	112
4.3.2.	Überblick über ausgewählte Rezeptionsstudien.....	116
4.3.3.	Zusammenfassung.....	123

---

<b>5. Entwicklung eines theoretischen Rahmenmodells zur Erfassung der Rezeption und Nutzung von Schulinspektionsergebnissen .....</b>	<b>125</b>
5.1. Verwendung des Helmke-Modells (2004) in Untersuchungen zum Umgang mit Ergebnisrückmeldungen aus Lernstandserhebungen .....	125
5.2. Übertragung des Helmke-Modells (2004) auf das Instrument <i>Schulinspektion</i> .....	128
<b>6. Ableitung von Forschungsfragen.....</b>	<b>133</b>
<b>III Empirischer Teil.....</b>	<b>135</b>
<b>7. Methodisches Vorgehen: Triangulation .....</b>	<b>135</b>
<b>8. Studie I: Sekundäranalyse der <i>Evaluationsbögen</i> der Qualitätsanalyse .....</b>	<b>141</b>
8.1. Datengrundlage und methodisches Vorgehen .....	141
8.1.1. Beschreibung und Charakterisierung der Datengrundlage der Sekundäranalyse.....	142
8.1.2. Fragestellungen .....	146
8.1.3. Methodisches Vorgehen und Analyseverfahren .....	146
8.2. Ergebnisse .....	148
8.2.1. Ergebnisse der explorativen Faktorenanalyse (EFA) .....	148
8.2.2. Ergebnisse der konfirmatorischen Faktorenanalyse (CFA).....	150
8.2.3. Skalenbeschreibungen.....	152
8.2.4. Deskriptive Skalenkennwerte .....	155
8.2.5. Diskussion der Skalen.....	159
8.2.6. Zusammenhänge zwischen den Phasen und Rahmenbedingungen.....	161
8.3. Gewinn und Grenzen der Sekundäranalyse: eine kritische Reflexion ..	165
<b>9. Studie II: Inhaltsanalyse der Zielvereinbarungen.....</b>	<b>167</b>
9.1. Dokumentenanalyse der Zielvereinbarungen .....	167
9.1.1. Beschreibung des Ausgangsmaterials.....	168
9.1.2. Quellenkritik .....	169

9.1.3.	Fragestellungen.....	171
9.2.	Methodisches Vorgehen: Inhaltsanalyse.....	171
9.2.1.	Entwicklung des Kategoriensystems.....	173
9.2.2.	Darstellung des eingesetzten Kategoriensystems.....	178
9.2.3.	Codierung .....	181
9.3.	Ergebnisse.....	182
9.3.1.	Formale Merkmale der Zielvereinbarungen.....	182
9.3.2.	Inhaltliche Merkmale der Zielvereinbarungen.....	185
9.3.3.	Zusammenhang zwischen den Zielinhalten und den festgestellten Stärken und Schwächen der Schulen .....	190
9.3.4.	Zielqualität: Bestimmung der Zielspezifität.....	197
9.3.5.	Zusammenhang zwischen der Zielspezifität und formalen Merkmale der Zielvereinbarungen .....	199
9.4.	Bewertende Zusammenfassung und Schlussfolgerungen aus der Zielvereinbarungsanalyse.....	200
<b>10.</b>	<b>Studie III: Onlinebefragung der Schulaufsicht .....</b>	<b>205</b>
10.1.	Fragestellungen.....	205
10.2.	Methodisches Vorgehen: Onlinebefragung.....	206
10.3.	Erhebungsinstrument.....	207
10.3.1.	Fragebogenkonstruktion.....	207
10.3.2.	Pretest: Maßnahmen zur Qualitätssicherung des Fragebogens.....	210
10.4.	Durchführung der Befragung .....	211
10.4.1.	Rekrutierung .....	211
10.4.2.	Teilnahmeverhalten: Rücklauf und Ausschöpfungsquote .....	213
10.5.	Stichprobe.....	216
10.6.	Ergebnisse.....	217
10.6.1.	Explorative Faktorenanalyse .....	217
10.6.2.	Skalenbeschreibung und -ausprägung.....	217

---

10.6.3. Zusammenhänge zwischen den Phasen und Rahmenbedingungen.....	221
10.6.4. Auseinandersetzung der Schulaufsicht mit der Qualitätsanalyse NRW und ihren Ergebnissen.....	223
10.6.5. Zielvereinbarungsprozess .....	226
10.6.6. Zielinhalte .....	228
10.6.7. Beratungs- und Unterstützungsleistungen .....	230
10.6.8. Zusammenfassung der offenen Antworten im Kommentarfeld .....	231
10.7. Resümee .....	232
<b>11. Triangulation der Studien I, II und III.....</b>	<b>235</b>
11.1. Erweitertes Rezeptions-Nutzungs-Modell von Schulinspektionsergebnissen .....	236
11.2. Triangulation auf der Ebene der Ergebnisse.....	238
11.3. Triangulation auf der Ebene des Einzelfalls.....	242
<b>IV. Schluss .....</b>	<b>247</b>
<b>12. Zusammenfassende Bewertung der Ergebnisse und Ausblick.....</b>	<b>247</b>
12.1. Ausgangspunkt der Arbeit .....	247
12.2. Zusammenfassende Bewertung der Studien I, II und III und ihrer Triangulation.....	250
12.3. Ausblick auf zukünftige Forschung.....	254
<b>V. Literatur.....</b>	<b>257</b>



# ABBILDUNGSVERZEICHNIS

---

Abb. 1:	Modell zu Qualität und Qualitätsentwicklung im Bildungsbereich .....	10
Abb. 2:	Grundstruktur outputorientierter Steuerung.....	12
Abb. 3:	Externe Evaluationsformen.....	17
Abb. 4:	Wirkungsmodell zur Schulinspektion.....	82
Abb. 5:	Struktur der Schulaufsicht in Nordrhein-Westfalen .....	89
Abb. 6:	Das Rubikon-Modell der Handlungsphasen .....	98
Abb. 7:	Zielsetzungstheorie .....	99
Abb. 8:	Idealtypischer Ablauf des Zielvereinbarungsprozesses zwischen Schule und Schulaufsicht nach erfolgter Qualitätsanalyse .....	105
Abb. 9:	Rahmenmodell für die Rückmeldung von Leistungsdaten (Bonsen/von der Gathen 2004) .....	113
Abb. 10:	Der Zyklus Rezeption – Reflexion – Aktion – Evaluation im Kontext von Lernstandserhebungen (Helmke 2004).....	115
Abb. 11:	Prozessphasen der Schulinspektion .....	129
Abb. 12:	Eigenes Rezeption-Nutzungs-Modell von Schulinspektionsergebnissen .....	131
Abb. 13:	Untersuchungsdesign „Trichtermodell Triangulation“ zur Rezeption und Nutzung von Rückmeldungen der Qualitätsanalyse NRW .....	137
Abb. 14:	Ergebnisse der konfirmatorischen Faktorenanalyse der Evaluationsbögen (N = 578).....	151
Abb. 15:	Boxplots der Skalen der Sekundäranalyse.....	158
Abb. 16:	Prozentuale Häufigkeit der Zielinhalte (> 1 %) auf der Ebene der Qualitätsaspekte 1.1 bis 6.3 .....	187
Abb. 17:	Bewertung der Ziele (N = 323) als Teilmenge aller bewerteten Qualitätsaspekte (N = 2772) .....	190
Abb. 18:	Vergleichende Darstellung der Bewertungsstufen 1 bis 4 auf der Ebene der Ziele (N = 323) und aller bewerteten Qualitätsaspekte (N = 2772).....	191
Abb. 19:	Kombinierte Darstellung der Aspektbewertungen und der Häufigkeit der Zielinhalte auf der Basis der Qualitätsaspekte (N = 114 ZV) .....	193

---

Abb. 20: Gegenüberstellung aller Aspektbewertungen und der Bewertungen der Ziele auf Einzelschulebene (N = 111 Schulen) .....	195
Abb. 21: Durchschnittliche Zielbewertungen nach Schulform (Balken) und Schuljahr (Punkte); (N = 114 Zielvereinbarungen) .....	196
Abb. 22: Prozentuale Häufigkeit der Zielspezifität auf der Ebene der Ziele (N = 339) und der Zielvereinbarungen (N = 114) .....	197
Abb. 23: Häufigkeitsverteilung der Items zur schulaufsichtlichen Auseinandersetzung mit der Qualitätsanalyse und ihren Ergebnissen (N = 72) .....	224
Abb. 24: Häufigkeitsverteilung der Items zum Zielvereinbarungsprozess aus der Perspektive der Schulaufsicht .....	227
Abb. 25: Auf der Grundlage der Ergebnisse der Studie I, II und III erweitertes Rahmenmodell zum Rezeptions-Nutzungs-Prozess von Schulinspektionsergebnissen .....	237

# TABELLENVERZEICHNIS

---

Tab. 1:	Merkmale guten Unterrichts nach Helmke (2009) und Meyer (2005).....	8
Tab. 2:	Typisierung von Evaluationsformen.....	15
Tab. 3:	Vergleich des Umfangs der Unterrichtsbeobachtungsbögen in sechs Bundesländern.....	43
Tab. 4:	Übersicht der Qualitätsanalysen nach Schulform und Bezirksregierung bis zum Ende des Schuljahres 2008/09.....	53
Tab. 5:	Systematiken der Schulinspektionsforschung.....	63
Tab. 6:	Arten der Schulaufsicht nach Dubs.....	91
Tab. 7:	Aufgaben der zuständigen Schulaufsicht im Zusammenhang mit der Schulinspektion in den Bundesländern.....	94
Tab. 8:	SMART-Regel für Ziele.....	101
Tab. 9:	Verwendung des theoretischen Rahmenmodells zur Nutzung von Vergleichsarbeiten (Helmke 2004).....	127
Tab. 10:	Aspekte des Inspektionswirkungsmodells nach Sommer (2011).....	128
Tab. 11:	Verteilung der Schulformen in der Stichprobe der Sekundäranalyse und der Grundgesamtheit in der Amtlichen Schulstatistik NRW.....	144
Tab. 12:	Ergebnisse der explorativen Faktorenanalyse (PCA) der Evaluationsbögen.....	148
Tab. 13:	Itemformulierungen und statistisch relevante Kennwerte der Skalen der Sekundäranalyse (N = 1155).....	155
Tab. 14:	Übersicht über die Skalen der Sekundäranalyse (N = 1155).....	157
Tab. 15:	Skalen-Interkorrelationen der Sekundäranalyse.....	162
Tab. 16:	Modellzusammenfassungen der schrittweisen Regressionsanalyse für die Skala <i>Rezeption</i> als AV.....	163
Tab. 17:	Modellzusammenfassungen der schrittweisen Regressionsanalyse für die Skala <i>Reflexion</i> als AV.....	164
Tab. 18:	Anzahl der analysierten Zielvereinbarungen nach Schulform und -jahr.....	169
Tab. 19:	Vorgaben zu den Bestandteilen von Schulprogrammen und Zielvereinbarungen im Rahmen der Qualitätsanalyse NRW.....	175

Tab. 20:	Eingesetztes Kategoriensystem zur inhaltsanalytischen Auswertung von Zielvereinbarungen .....	179
Tab. 21:	Ausprägungen der Zielspezifität zur Bestimmung der Zielqualität.....	180
Tab. 22:	Umfang der Zielvereinbarungen in Seiten und Wörtern im Schulformvergleich.....	183
Tab. 23:	Anzahl der Ziele und Teilziele in den Zielvereinbarungen im Schulformvergleich.....	184
Tab. 24:	Prozentuale Häufigkeiten der Zielinhalte in den Qualitätsbereichen 1 bis 6 im Schulformvergleich.....	186
Tab. 25:	Prozentuale Häufigkeiten der Ziele auf Ebene der Kriterien der Qualitätsaspekte 2.4, 2.6 und 2.2 (absteigend sortiert).....	189
Tab. 26:	Kreuztabelle Zielspezifität und Seitenanzahl .....	199
Tab. 27:	Kreuztabelle Zielspezifität und Zielanzahl .....	200
Tab. 28:	Inhaltliche Themen des Online-Fragebogens der Schulaufsicht .....	208
Tab. 29:	Itemformulierungen und statistisch relevante Skalen-Kennwerte der Befragung der Schulaufsicht (N = 72).....	219
Tab. 30:	Übersicht über die Skalen der Befragung der Schulaufsicht.....	220
Tab. 31:	Skalen-Interkorrelationen der Befragung der Schulaufsicht .....	222
Tab. 32:	Mittelwertunterschiede der Skalen in Abhängigkeit der Intensität der Auseinandersetzung mit der Qualitätsanalyse und ihren Ergebnissen .....	226
Tab. 33:	Auswertung der offenen Angaben zu den Zielinhalten aus der Perspektive der Schulaufsicht (N = 159 Nennungen; N = 53 Fälle) .....	229
Tab. 34:	Bereitgestellte Unterstützungsleistungen durch die Schulaufsicht.....	230
Tab. 35:	Gegenüberstellung der drei häufigsten Zielinhalte in den Studien II und III.....	241
Tab. 36:	Verteilung der Triangulationsstichprobe nach Schulform und -jahr .....	243
Tab. 37:	Skalen-Interkorrelationen auf Einzelfallebene .....	245

# ABKÜRZUNGSVERZEICHNIS

---

<b>Abkürzung</b>	<b>Beschreibung</b>
$\alpha$	Cronbachs Alpha
Abb.	Abbildung
AV	abhängige Variable
BeLesen	Berliner Längsschnittstudie zur Lesekompetenzentwicklung von Grundschulkindern mit Migrationshintergrund
bspw.	beispielsweise
bzgl.	bezüglich
bzw.	beziehungsweise
BLK	Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung
BMBF	Bundesministerium für Bildung und Forschung
ca.	circa
CDU	Christlich Demokratische Union Deutschlands
CFA	Confirmatory Factor Analysis (Konfirmatorische Faktorenanalyse)
DDR	Deutsche Demokratische Republik
d. h.	das heißt
DIPF	Deutsches Institut für Internationale Pädagogische Forschung
ebd.	ebenda
EBIS	Entwicklungsbilanzen im Schulamt
EFA	Exploratory Factor Analysis (Explorative Faktorenanalyse)
EFQM	European Foundation for Quality Management
etc.	et cetera
EU	Europäische Union
evtl.	eventuell
FDP	Freie Demokratische Partei

GE	Gesamtschule
ggf.	gegebenenfalls
GS	Grundschule
GY	Gymnasium
HS	Hauptschule
i. d. R.	in der Regel
i. e. S.	im engeren Sinne
IFS	Institut für Schulentwicklungsforschung an der Technischen Universität Dortmund
IGLU	Internationale Grundschul-Lese-Untersuchung
IQB	Institut für Qualitätsentwicklung im Bildungswesen, Berlin
KMK	Kultusministerkonferenz
KMO	Kaiser-Maier-Olkin-Koeffizient
MARKUS	Mathematik-Gesamterhebung Rheinland-Pfalz: Kompe- tenzen, Unterricht, Schulkontext
MbO	Management by Objectives
MSW	Ministerium für Schule und Weiterbildung Nordrhein- Westfalen
MW	Mittelwert
N	Population (alle Elemente der Grundgesamtheit)
n	Stichprobengröße
NPM	New Public Management
NRW	Nordrhein-Westfalen
OECD	Organisation for Economic Cooperation and Development (Organisation für ökonomische Zusammenarbeit und Entwicklung)
Ofsted	Office for Standards in Education, Children's Services and Skills
o. g.	oben genannt
PASW	Predictive Analytics Software
PC	Personal Computer

---

PCA	Principal Component Analysis (Hauptkomponentenanalyse)
PISA	Programme for International Student Assessment (Programm zur internationalen Überprüfung der Schülerleistungen)
QA	Qualitätsanalyse (NRW)
QB	Qualitätsbereich
QuaSUM	Qualitätsuntersuchungen an Schulen zum Unterricht in Mathematik (Brandenburg)
resp.	respektive
RS	Realschule
S.	Seite
s.	siehe
s. o.	siehe oben
s. u.	siehe unten
SchulG	Schulgesetz
SD	Standardabweichung
SE	Schulentwicklung
Sek I	Sekundarstufe I
Sek II	Sekundarstufe II
SICI	The Standing International Conference of Central and General Inspectorates of Education (Organisation nationaler und regionaler Inspektionen des Bildungswesens)
sog.	so genannt
SPD	Sozialdemokratische Partei Deutschlands
SPSS	Software Statistical Package for the Social Sciences
Tab.	Tabelle
TIMSS	Trends in International Mathematics and Science Study (Internationale Mathematik- und Naturwissenschaftsstudie)
u. a.	und andere
TQM	Total Quality Management

USA	United States of America (Vereinigte Staaten von Amerika)
u. U.	unter Umständen
UV	unabhängige Variable
VERA	Vergleichsarbeiten (auch als Lernstandserhebungen bezeichnet)
vgl.	vergleiche
VO	Verordnung
vs.	versus
WALZER	Wirkungsanalyse der Leistungsevaluation: Zielerreichung, Ertrag für die Bildungsqualität der Schule und die Rückmeldung von Evaluationsergebnissen
z. B.	zum Beispiel
z. T.	zum Teil
ZV	Zielvereinbarung

# I EINLEITUNG

---

## 1. Ausgangspunkt, Zielsetzung und Aufbau der Arbeit

*Externe Evaluation ist die Pflicht,  
die Nutzung der Ergebnisse für die Schulentwicklung  
ist bisher die Kür.*  
(Hovestadt 2009, S. 40)

Als Anlass der vorliegenden Arbeit können die Reformen im Bildungsbereich angeführt werden, welche durch die Veröffentlichung der ersten PISA-Ergebnisse im Jahr 2001 ausgelöst worden sind. Im Zuge der sog. Neuen Steuerung wurde eine Reihe von internen und externen Evaluationsinstrumenten implementiert, von denen sich die Schuladministration umfassende Impulse für die Schul- und Unterrichtsentwicklung in der Einzelschule und die Generierung evaluationsbasierten Steuerungswissens verspricht. Die *Schulinspektion*<sup>1</sup> als externe Evaluationsform gilt als ein solches Reforminstrument, dessen Potential als Impulsgeber für einzelschulische Qualitätsentwicklung jedoch weitgehend unbekannt und keineswegs empirisch untersucht und abgesichert ist. Die Frage nach den Wirkungen und Konsequenzen von Evaluationsinstrumenten gerät seit einigen Jahren immer stärker in den Fokus der Bildungsforschung:

- „Leistungsvergleiche – und was dann?“ (Terhart 2002, S. 91),
- „Evaluation und dann?“ (Ernst/Döbert 2006, S. 54),
- „Testen und dann?“ (Maier 2009, S. 191),
- „Testen, prüfen, vergleichen – und dann?“ (Heinrich 2010, S. 116),
- „Schulinspektion – und was dann?“ (Klippert 2010, S. 28).

---

<sup>1</sup> Im Folgenden wird der Begriff Schulinspektion stellvertretend für die unterschiedlichen Bezeichnungen in Europa und den Bundesländern (z. B. Externe Evaluation, Schulvisitation etc.) sowohl für die Institution als auch für das Verfahren in Anlehnung an Maritzen (2006, S. 7) verwendet. Ist nur eine bestimmte Schulinspektion gemeint, so wird die entsprechende Bezeichnung verwendet.

In diesen Fragen spiegelt sich die Unwissenheit über die Folgen der aktuellen externen Evaluationsinstrumente wider, während gleichzeitig implizit die Kritik ihres strukturellen Mangels und ihrer Anschlusslosigkeit mitschwingt. Bei der flächendeckenden Einführung der verschiedenen Evaluationsinstrumente wurde häufiger das Hauptaugenmerk auf die Implementierung gerichtet, während die Phase der Anschlussfähigkeit der rückgemeldeten Ergebnisse nachgeordnet behandelt wurde, obwohl gerade diese Phase entscheidend für die Wirksamkeit der jeweiligen Instrumente ist. Zwar gilt die Datenrückmeldung der gemessenen Schüler<sup>2</sup>- und Schulleistung bereits als Unterstützungssystem, jedoch werden den Schulen damit keine direkten Unterstützungsleistungen im Hinblick auf die Initiierung von Entwicklungsvorhaben angeboten (vgl. Berkemeyer 2011). Das von Terhart (2002, S. 108) definierte Problem

„Das Wissen über Leistungsergebnisse und Leistungsvoraussetzungen von Schulsystemen und Schulen wächst schneller als das Wissen darüber, was mit den Ergebnissen angefangen werden kann und soll“

gilt auch für die Schulinspektion und lässt sich um die Frage ergänzen, wie mit den Ergebnissen tatsächlich umgegangen wird.

Vor diesem Hintergrund bildet das einleitende o. g. Zitat von Hovestadt (2009) den roten Faden in der vorliegenden Arbeit. Die Nutzung von Evaluationsergebnissen stellt sich nicht automatisch als Folge externer Evaluation ein. Dieser Prozess ist von der Schuladministration intendiert, steht jedoch weitestgehend dem Unwissen gegenüber, ob – und wenn ja, wie – diese Ergebnisse von den verschiedenen Akteuren im schulischen Umfeld genutzt werden. An dieses Forschungsdefizit knüpft die vorliegende Arbeit an, indem sie anhand verschiedener empirischer Quellen untersucht, wie die Ergebnismeldungen der Schulinspektion in den betroffenen Schulen aufgenommen, verarbeitet und für die Weiterentwicklung der einzelnen Schule genutzt werden. Thematisch bewegt sich diese Arbeit somit an der Nahtstelle zwischen der Schulinspektions- und Rezeptionsforschung.

---

<sup>2</sup> In der vorliegenden Arbeit wird auf die Verwendung weiblicher Personenbezeichnungen („Schülerin“, „Lehrerin“, „Qualitätsprüferin“, „Schulaufsichtsbeamtin“ etc.) aus Gründen der besseren Lesbarkeit verzichtet. Die männliche Form gilt, wenn nicht anders vermerkt, für beide Geschlechter.

Ziel dieser Arbeit ist es, die Rezeption und Nutzung von Schulinspektionsergebnissen unter besonderer Berücksichtigung der im Anschluss an die Schulinspektion von den Schulen und der *Schulaufsicht* getroffenen *Zielvereinbarungen* empirisch zu analysieren. Die in den Zielvereinbarungsdokumenten verschriftlichten Maßnahmenplanungen eignen sich dazu, Anhaltspunkte zur (geplanten) Nutzung der Ergebnisse und zu den intendierten Folgen der Schulinspektion im Sinne einer evidenzbasierten Nutzung der Rückmeldungen zu liefern. Die Rolle der Schulaufsicht im Kontext der Schulinspektion ist bisher kaum Gegenstand wissenschaftlicher Studien gewesen, obwohl ihr dabei eine nicht ganz unwichtige Funktion zukommt.

Durch die Nutzung unterschiedlicher Datenquellen der *Qualitätsanalyse* ist die vorliegende Arbeit im Bundesland Nordrhein-Westfalen (NRW) verortet. Den empirischen Untersuchungen wird ein Rahmenmodell zur pädagogischen Nutzung von Vergleichsarbeiten (vgl. Helmke 2004) zugrunde gelegt, welches einen hohen heuristischen Wert besitzt und in den letzten Jahren Forschungsaktivitäten angeregt hat. Für den Einsatz des Modells im Kontext der Schulinspektion wird es entsprechend angepasst.

Im Rahmen der theoretischen Einbettung dieser Arbeit werden in Kapitel 2 zunächst die zentralen Themenbereiche *Schulqualität* und *Schulreform* skizziert. Es folgt eine Darstellung der Ergebnisse der Forschung zu guten Schulen und gutem Unterricht, die Dittons Modell zur Schulqualität komprimiert (vgl. Ditton 2000), um dann die Schulinspektion als ein Instrument externer Evaluation in die neueren Entwicklungen zu schulischer Steuerung und Evaluation einzuordnen. Das Evaluationsinstrument *Schulinspektion* wird differenziert auf der Ebene des europäischen Auslands, der deutschen Bundesländer und speziell der Qualitätsanalyse NRW dargestellt. Kapitel 3 und 4 widmen sich den beiden Forschungssträngen, die für die vorliegende Arbeit von Bedeutung sind: der Schulinspektions- und der Rezeptionsforschung. Der empirische Forschungsstand zur Schulinspektion wird fokussiert auf deutsche Forschungsaktivitäten und die Schulentwicklungsfunktion systematisch aufgearbeitet. In Kapitel 4 wird ein Überblick über die bislang vorliegenden Befunde des relativ neuen Forschungstyps Rezeptionsforschung und Modelle zur Rezeption und Nutzung von Ergebnissrückmeldungen gegeben. In jeweils eigenen Teilkapiteln werden die Funktionen und Aufgaben der Schulaufsicht im Rahmen der Schulinspekti-

on und die Zielvereinbarungen beleuchtet. Den Abschluss des Theorieteils bilden die Entwicklung eines theoretischen Rahmenmodells zur Erfassung der Rezeption und Nutzung von Schulinspektionsergebnissen auf der Grundlage eines bestehenden Modells zu Vergleichsarbeiten (vgl. Helmke 2004) und die Überleitung zu den Fragestellungen.

Der empirische Teil leitet mit der Erläuterung der Methodik der Triangulation und des Untersuchungsdesigns ein. Um den Fragestellungen multimethodisch und -perspektivisch zu begegnen, werden drei empirische Studien konzipiert. In jeweils eigenen Kapiteln werden Datengrundlage, methodisches Vorgehen sowie Ergebnisse und Interpretation der drei Datenquellen präsentiert. Eine zusammenfassende Analyse der Ergebnisse erfolgt im letzten Kapitel des empirischen Teils, in dem die drei Einzelstudien auf Ergebnis- und Einzelfallebene trianguliert werden. Im Kern geht es in den drei Studien um die Analyse folgender Elemente:

In Studie I „*Sekundäranalyse der Evaluationsbögen der Qualitätsanalyse*“ werden ministeriell administrierte Evaluationsbögen für Schulen im Anschluss an die Qualitätsanalyse NRW sekundärstatistisch ausgewertet, um Erkenntnisse über die Rezeption und Reflexion der Ergebnisse aus der Perspektive der Schulen zu erhalten. Andererseits wird die Gültigkeit des modifizierten Rahmenmodells von Helmke einer Prüfung unterzogen. Studie II „*Inhaltsanalyse der Zielvereinbarungen*“ untersucht die zwischen Schulen und Schulaufsicht geschlossenen Zielvereinbarungsdokumente inhaltsanalytisch im Hinblick auf ihre Form, ihren Inhalt und ihre Qualität. Schließlich wird mit der dritten Studie „*Onlinebefragung der Schulaufsicht*“ die Perspektive der Schulaufsicht auf den Rezeptions-Nutzungs-Prozess einbezogen. Am Ende des empirischen Teils werden die aus den unterschiedlichen Datenquellen gewonnenen Informationen trianguliert, um auf diese Weise ein mehrperspektivisches Bild der Rezeption und Nutzung von Schulinspektionsergebnissen zeichnen zu können und den Forschungsgegenstand zu vertiefen. Dabei wird das modifizierte Helmkesche Rahmenmodell mit den erzielten Befunden ergänzt und weiter ausdifferenziert.

In Kapitel 12 werden in einer Zusammenfassung die forschungsleitenden Fragen beantwortet. Abschließend wird ein Ausblick auf verschiedene Forschungsszenarien gegeben.

## II THEORETISCHER TEIL

---

### 2. Schulqualität und Schulreform

#### 2.1. Qualität von Schule und Unterricht

Die Auseinandersetzung mit der Thematik, was eine *gute* Schule ist und diese ausmacht, gehört zu den Grundfragen der Bildungsforschung und muss auch im Rahmen der Implementierung von Schulinspektion geleistet werden, da die Schulqualität sowohl einen Ausgangspunkt als auch eine Zielgröße der Schulinspektion darstellt (vgl. Dederling/Müller/Bos 2008, S. 10). Dabei stützt sich die Schulinspektion auf bereits vorliegende Befunde empirischer Bildungsforschung zur Schulqualität (vgl. Bensen/Bos/Rolff 2008, S. 15).

In diesem Kapitel erfolgt zunächst eine Annäherung an den Qualitätsbegriff im Bildungsbereich, gefolgt von der Darstellung empirisch fundierter Merkmale guter Schulen und guten Unterrichts, die Dittons Modell zur Qualität und Qualitätsentwicklung im Bildungsbereich komprimiert (vgl. Ditton 2000; 2007). Auf diese Grundlagen wird im weiteren Verlauf zurückgegriffen, um zu klären, welches Qualitätsverständnis der Schulinspektion zugrunde liegt.

##### 2.1.1. Qualitätsbegriff im Bildungsbereich

Seit den 1990er Jahren findet der Qualitätsbegriff, welcher ursprünglich im betriebswirtschaftlichen Bereich eine Normierung und Standardisierung von Produkten und Produktionsprozessen definierte, verstärkt im Bildungsbereich Verwendung (vgl. Terhart 2000, S. 809). Betriebswirtschaftliche Managementinstrumente, z. B. die Qualitätsmanagementnorm ISO 9000 ff. und das Total Quality Management-Konzept (TQM) wurden von der *European Foundation for Quality Management* (EFQM) auf den Bildungsbereich, unter anderem auf die Schulevaluation, übertragen<sup>3</sup>.

Dabei lässt der Qualitätsdiskurs im Bildungsbereich folgende Probleme erkennen (vgl. Ditton 2000, S. 74f.): Zum einen besteht über den Qualitätsbegriff

---

<sup>3</sup> Zur Übertragungsproblematik privat-wirtschaftlichen Denkens auf den Bildungsbereich wird auf Terhart (2000) und Böttcher (2002) verwiesen.

und seine Verwendung kein Konsens, zum anderen sind Indifferenzen der Qualitätsindikatoren erkennbar und schließlich liegen langfristige Wirkungen des Schulbesuchs noch im Verborgenen. Insgesamt ist ungewiss – so Ditton (2000, S. 75) –,

auf welcher Ebene (Schüler und Lehrer, Unterricht, Schule, Schulsystem), mit welchen Verfahren (qualitativ vs. quantitativ; intern vs. extern), anhand welcher Bewertungsmaßstäbe und Erwartungen (ökonomische vs. pädagogische) reliable und valide Aussagen über die schulische Qualität (...) getroffen werden können.

Die erstgenannte Problematik bezüglich der Konsenslosigkeit des Qualitätsbegriffs aufgreifend wird zunächst ein Definitionsversuch vorgenommen. Der Qualitätsbegriff stellt zunächst eine relative Größe dar, indem er die „Unterscheidung von etwas Höherwertigem gegenüber etwas Geringerwertigem“ (Fend 2000, S. 56) ermöglicht (vgl. auch Harvey/Green 2000, S. 17). Im Bildungsbereich gilt Qualität als ein

verabredeter Gütemaßstab, der auf den offiziellen Zwecksetzungen des Bildungssystems insgesamt bzw. einzelner Bildungs- und Sozialeinrichtungen basiert und der insofern auch als Bezugspunkt für die Ermittlung und ggf. den Vergleich der faktischen Wirkungen dieser Einrichtungen herangezogen werden kann (Terhart 2000, S. 815).

Es sind nach Terhart (2000) vier unterschiedliche Herangehensweisen zur näheren Bestimmung der Schulqualität zu differenzieren:

- *Normative Bestimmungsversuche*: Über Merkmale guter Schulen existieren normative Vorstellungen und inhaltliche Entscheidungen, welche z. B. auf Traditionen oder politischen Überzeugungen beruhen.
- *Analytische Bestimmungsversuche*: Analytische Bestimmungsversuche können keine inhaltliche Bestimmung von Qualität leisten, sondern nehmen eine formale, begriffsklärende Aufgabe ein.
- *Empirische Bestimmungsversuche*: Damit wird eine Erfassung der tatsächlichen Leistungsfähigkeit bzw. Wirkung von Schulen beschrieben, die zu ihrem Auftrag in Beziehung gesetzt wird. Im Vordergrund steht eine Wirkungskontrolle.

- *Qualität als operatives Problem*: Im Zentrum steht die Fragestellung, mit welchen Strategien und Instrumenten eine möglichst hohe und homogene Qualität erreicht werden kann.

Bevor das Qualitätsverständnis der Schulinspektion im weiteren Verlauf dargelegt wird, werden zunächst Merkmale guter Schulen und guten Unterrichts vorgestellt und in ein Modell zur Schulqualität eingebettet.

### **2.1.2. Merkmale guter Schulen und guten Unterrichts**

Trotz der von Ditton (2000, S. 74f.) angebrachten Problematik bezüglich der Differenzen der Qualitätsindikatoren liegt der internationalen Diskussion zur Schulqualität mittlerweile ein weitgehender Konsens über bestimmte Merkmale guter – im Sinne von effektiver – Schulen zugrunde, die als bedeutsam für Schülerleistungen erachtet werden. Die zentralen, von der internationalen Schuleffektivitätsforschung extrahierten Indikatoren der Qualität einer Schule bezogen auf den Lernerfolg der Schüler (vgl. zur Übersicht Bos/Bonsen/Berke-meyer 2010, S. 64f.; Scheerens/Glas/Thomas 2003, S. 241-243) sind:

- *Leistungsorientierung*: hohe und zugleich angemessene Erwartung an Schüler als auch Lehrkräfte.
- *Professionelle Kooperation im Kollegium*: konsensorientiertes Verständnis der Schul- und Unterrichtsqualität im Lehrerkollegium.
- *Pädagogische Führung*: besondere Stellung der Schulleitung in der Qualitätsentwicklung der Schule.
- *Qualität des Curriculums*: Reflexion der eigenen pädagogischen Arbeit; Abgleich des intendierten und implementierten Curriculums.
- *Geordnete Lernatmosphäre*: geordnete Arbeitsbedingungen; positives Sozialklima zwischen Schülern, Schülern und Lehrern sowie zwischen Lehrern.
- *Evaluation*: Einsatz von Evaluationen auf unterschiedlichen Ebenen (Individual-, Unterrichts-, Organisationsebene), z. B. Selbst- und Fremdevaluationen.

Trotz der Übereinstimmungen liegen den einzelnen empirischen Untersuchungen, die diese Indikatoren isoliert haben, z. T. beträchtliche Unterschiede bzgl. ihrer Operationalisierung zugrunde. Zudem fehlen empirische Nachweise, wie

sich Schulqualität auf die Schülerleistungen auswirkt, denn auch wenn die o. g. zentralen Indikatoren guter Schulqualität von Schulen berücksichtigt und umgesetzt werden, münden sie nicht automatisch in erfolgreichen Schülerleistungen. Ein Grund hierfür ist, dass der durch die Einzelschule aufklärbare Anteil der Gesamtvarianz in den Schülerleistungen bei etwa 10 % liegt (vgl. Scheerens u.a. 2003, S. 251).

Die Qualität von Schule zeigt sich insbesondere auf der Ebene des Unterrichts (vgl. Terhart 2002, S. 99). Hierzu konnte die empirische Unterrichtsforschung Faktoren effizienten Unterrichts identifizieren. Tabelle 1 enthält die wesentlichen Merkmale des aktuellen empirisch-psychologischen und schulpädagogischen Diskurses zum Thema Unterrichtsqualität. Es werden zwei fächerübergreifende Zehnerkataloge von Merkmalen guten Unterrichts präsentiert (vgl. Helmke 2009; Meyer 2005), welche „zu dauerhaft hohen kognitiven, methodischen und sozialen Lernerfolgen beitragen“ (Meyer 2010, S. 166). Die Merkmale haben jedoch keine „Effektivitätsautomatik“, d. h. ihre Berücksichtigung führt nicht automatisch zu einer hohen Unterrichtsqualität (ebd., S. 169). Zudem ist der Wirkungsgrad einzelner Merkmale begrenzt, da starke Interdependenzen zwischen den Merkmalen vorliegen. Gerade das Zusammenspiel der Merkmale ist jedoch empirisch noch nicht vollständig erforscht.

**Tab. 1: Merkmale guten Unterrichts nach Helmke (2009) und Meyer (2005)**

	<b>Helmke (2009)</b>	<b>Meyer (2005)</b>
1.	Klassenführung	Klare Strukturierung
2.	Klarheit und Strukturiertheit	Hoher Anteil echter Lernzeit
3.	Konsolidierung und Sicherung	Lernförderliches Klima
4.	Aktivierung	Inhaltliche Klarheit
5.	Motivierung	Sinnstiftendes Kommunizieren
6.	Lernförderliches Klima	Methodenvielfalt
7.	Schülerorientierung	Individuelles Fördern
8.	Kompetenzorientierung	Intelligentes Üben
9.	Umgang mit Heterogenität	Transparente Leistungserwartungen
10.	Angebotsvariation	Vorbereitete Umgebung

### 2.1.3. Modell zur Schulqualität

Das Wissen über Schulqualität ist in den letzten Jahren sowohl durch internationale Schulleistungsvergleichsstudien als auch durch die Schul-, Unterrichts- und Bildungsforschung (vgl. Überblick z. B. bei Gröhlich 2012; Gruehn 2000)<sup>4</sup> gewachsen. Die Bestimmung von Schulqualität ist kein Selbstzweck, sondern ermöglicht insbesondere Schulentwicklung. Damit ist in Anlehnung an Dubs (2005, S. 10) Folgendes gemeint:

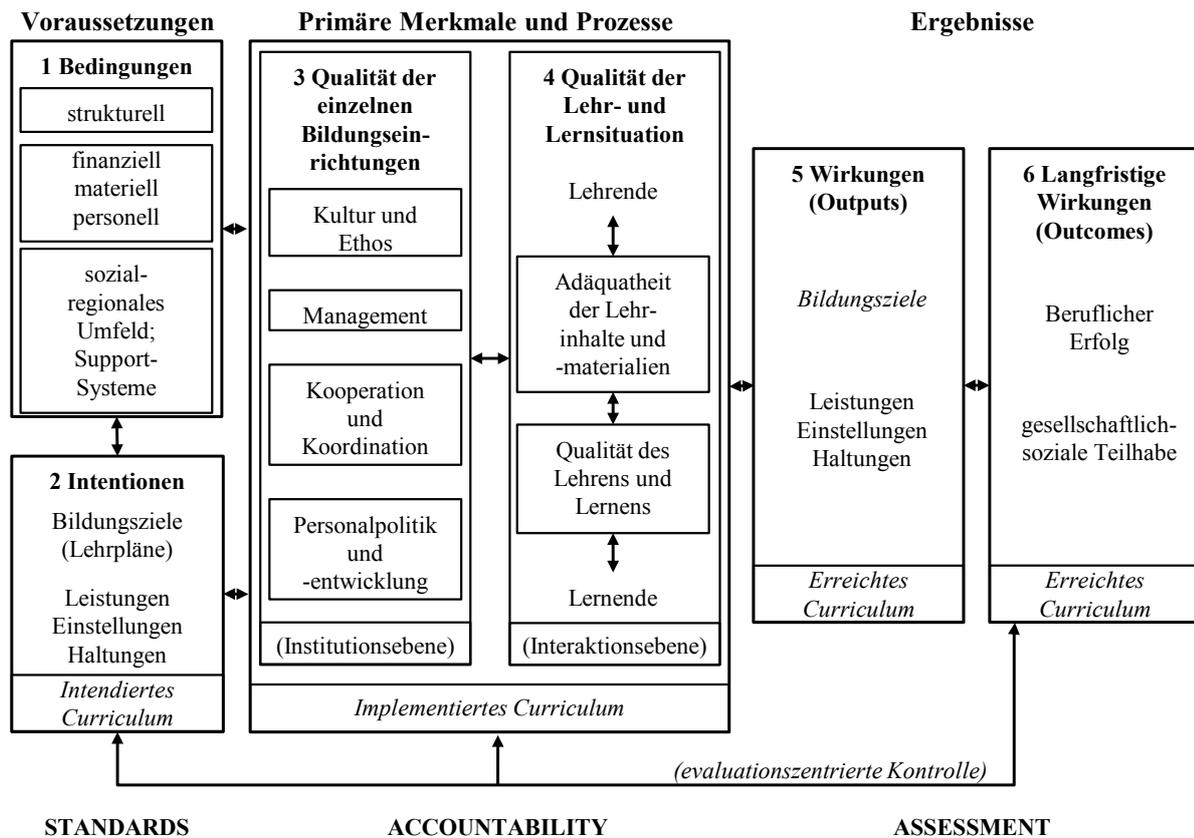
Ziel der Schulentwicklung ist es, mittels einer regelmäßigen Analyse einer Schule und gezielten Innovationen einen angestrebten, neuen Zustand in einer Schule zu erreichen. Schulentwicklung ist also ein langer, kontinuierlicher, dynamischer und planmäßiger Analyse-, Problemlöse-, Innovations- und Lernprozess (...). Die Schulentwicklung umfasst die Organisationsentwicklung, die Lehrplan- und Unterrichtsentwicklung sowie die Personalentwicklung.

Diesem Kerngedanken folgend stehen somit Veränderungen und Entwicklungsprozesse im Vordergrund der Schulentwicklung. Während sich der Kenntnisstand zu Schulprozessen zunehmend erweitert und über den schulischen Output durch die internationalen Vergleichsstudien ein solides Grundwissen vorliegt, bestehen zu schulischen Outcomes, d. h. langfristigen Wirkungen, noch wenige empirische Evidenzen.

Vor dem Hintergrund empirisch fundierter Untersuchungen zur Systematisierung von Schulqualität (vgl. z. B. Scheerens/Bosker 1997) hat Ditton (2000; 2007) ein *Modell zur Qualität und Qualitätsentwicklung im Bildungsbereich* entwickelt, welches sich als konzeptioneller Rahmen zur Bestimmung von Schulqualität eignet. Das Modell komprimiert die Forschung zu guten Schulen und gutem Unterricht und kommt im Rahmen empirischer Schulqualitätsforschung vielfach zum Einsatz. Es zeichnet sich zum einen durch seinen Prozesscharakter (Input – Prozess – Output) aus und ist durch eine Mehrebenensicht (individuelle, unterrichtliche, schulische und kontextuelle Ebenen) gekennzeichnet (vgl. Ditton 2007, S. 84).

---

<sup>4</sup> Ein Review zur Schuleffektivitätsforschung ist bei Scheerens/Glas/Thomas (2003, S. 235-299) zu finden. Zum Stand und zur Paradigmenbildung der Unterrichtsforschung wird auf Gruehn (2000, S. 3-92) verwiesen.



**Abb. 1: Modell zu Qualität und Qualitätsentwicklung im Bildungsbereich**

Zusätzlich wird die Perspektive der Systemsteuerung in Form von bspw. Bildungsstandards, Lernstandserhebungen und der Schulinspektion modelliert. Zwischen den einzelnen Elementen des Modells, welche durch die drei Steuerungskomponenten (Input, Prozess, Output) verknüpft sind, bestehen Interdependenzen. Die Schulqualität kann einen Einfluss auf die Unterrichtsqualität nehmen, aber auch umgekehrt (s. Box 3 und 4 in Abb. 1). Insgesamt impliziert das Modell, dass das Zusammenwirken der unterschiedlichen und zeitlich aufeinanderfolgenden Faktoren einen positiven Einfluss auf die Outputs, d. h. die Schülerleistungen, haben kann. Dabei zeigt sich die Qualität nicht nur in fachlichen Leistungen, sondern auch in überfachlichen Kompetenzen.

## **2.2. Steuerung und Evaluation**

Seit dem *PISA-Schock* vor rund zehn Jahren gehört es in Deutschland zum argumentativen bildungspolitischen Standard, in der Diskussion über Bildungsdefizite und mögliche Lösungswege auf die Steuerung und Evaluation von Schulen zu verweisen. Diese beiden Elemente haben in den letzten Jahren die Diskussion über Schul- und Unterrichtsentwicklung nachhaltig geprägt.

In diesem Kapitel wird zunächst eine Vorstellung des neuen Steuerungskonzepts im Bildungswesen sowie eine Klärung des Grundverständnisses des Begriffs *Evaluation* gegeben, woran Typisierungsvorschläge verschiedener Evaluationsformen anknüpfen. Anschließend werden interne und externe Steuerungsinstrumente aufgezeigt sowie die Einbettung der Schulinspektion in die expandierende Anzahl schulischer Steuerungsinstrumente vorgenommen.

### **2.2.1. Neue Steuerung im Bildungswesen**

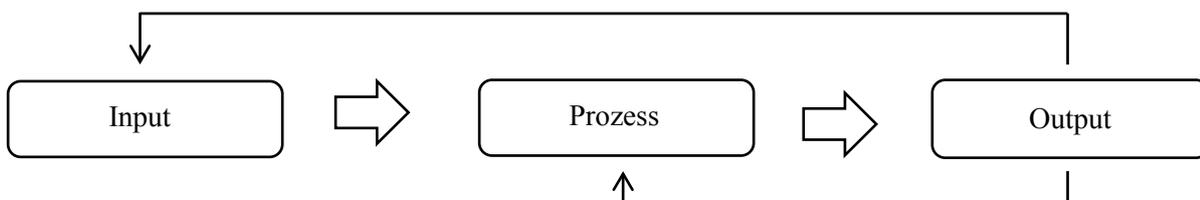
Die im Rahmen der TIMS- und PISA-Studien aufgedeckten Qualitätsdefizite im deutschen Schulwesen, gekennzeichnet durch unterdurchschnittliche Kompetenzen der Schüler, soziale Disparitäten sowie eine hohe Varianz zwischen Einzelschulen und Bundesländern, hatten einen Paradigmenwechsel in der Steuerungsphilosophie und die Implementierung neuer Steuerungsinstrumente zur Folge (vgl. Altrichter 2010, S. 171). Das Umdenken in der Schulsteuerung von der Input- zur Outputorientierung wurde maßgeblich dadurch gestützt, dass in den PISA-Studien diejenigen Länder erfolgreich waren, die eine im Vergleich zu Deutschland andere Steuerungsphilosophie vertraten. Daraus wurde abgeleitet, dass die Art der Systemsteuerung Einfluss auf die Leistungsergebnisse der Schulen und Schüler ausübe und somit von einer Steuerungs- und Schulreform eine Qualitätssteigerung im deutschen Schulwesen zu erwarten sei (vgl. Böttcher/Rürup 2010, S. 58).

Mit dem neuen Steuerungsparadigma nahm auch die Bedeutung externer Evaluationsformen, welche systematische Informationen über die Arbeits- und Wirkungsweise der Schulen bereitstellen, automatisch zu. Im Jahr 2006 wurde von der Kultusministerkonferenz (KMK) eine Gesamtstrategie zur Verbesserung der Schulqualität und Schülerleistungen beschlossen, welche die folgenden vier Instrumente umfasst (vgl. KMK 2006):

1. Teilnahme an internationalen Schulleistungsuntersuchungen
2. Zentrale Überprüfung von Bildungsstandards im Ländervergleich
3. Zentrale landesweite Lernstandserhebungen zur Überprüfung der Leistungsfähigkeit einzelner Schulen
4. Bildungsberichterstattung

In diesem Kontext wurde auch der Schulinspektion eine bedeutende Rolle zugedacht, allerdings nicht per KMK-Beschluss, sondern von den einzelnen Bundesländern konstituiert.

Die Umsetzung der Reformmaßnahmen wird unter den politikwissenschaftlichen Begriff der *Steuerung*, welcher die „Möglichkeiten der Einflussnahme des staatlichen Handelns in einem bestimmten gesellschaftlichen Feld“ (Zlatkin-Troitschanskaia 2006, S. 82) beschreibt, gefasst. Dabei bezeichnet die *Steuerung des Bildungswesens* die Möglichkeit der „vorausschauenden Planbarkeit und Beeinflussbarkeit von Entwicklungen im Schulwesen“ (Böttcher/Rürup 2010, S. 57; s. auch Recum 2003, S. 105) auf der Ebene von Bildungsstrukturen, -prozessen und -inhalten. Traditionell dominierte in Deutschland eine Input-Steuerung mit hierarchischem und zentralistischem Charakter, welche den Schulen detaillierte curriculare, Finanzierungs- und Gesetzgebungsvorgaben machte (vgl. Brauckmann u.a. 2010, S. 119). Das aktuelle Output-Steuerungskonzept ist hingegen durch eine stärkere Fokussierung auf die Erfassung und Bewertung von Bildungsprozessen und -ergebnissen gekennzeichnet und ergänzt die input- und prozessorientierte Steuerung (vgl. Altrichter/Maag Merki 2010, S. 37; Bellmann/Weiß 2009, S. 288; Heinrich 2008, S. 42). Auf der Grundlage von Informationen über den Output und Kausalannahmen über den Zusammenhang zwischen Input, Prozess und Output wird durch Interventionen des Inputs und Prozesses gesteuert (s. Abb. 2).



**Abb. 2: Grundstruktur outputorientierter Steuerung**

Für das neue Steuerungsverständnis ist neben der Outputorientierung das Konzept der Dezentralisierung kennzeichnend. Mit der Stärkung der Schulautonomie Anfang der 1990er Jahre wurde den Schulen mehr Freiraum bei den Entscheidungs- und Handlungsprozessen auf pädagogischer und personeller Ebene sowie in finanzieller Hinsicht zugestanden (vgl. zur Autonomiedebatte Altrichter/Rürup 2010; Heinrich 2007, S. 15-24; 59-70; Koch/Gräsel 2004, S. 15). Gleichzeitig sind die Schulen zur Rechenschaftslegung über ihre Arbeit, Leistung und Erreichung von Standards angehalten, welche durch verschiedene Evaluationsformen überprüft werden. Damit stellen einerseits die erweiterte Autonomie der Schulen und andererseits die systematische Überwachung der Zielerreichung von externer Seite zwei Seiten einer Medaille dar.

Vor diesem Hintergrund ergibt sich auch die Notwendigkeit der Governance-Perspektive, welche das komplexe Konstellationsgefüge der verschiedenen Akteure und ihre Handlungskoordination in einem Mehrebenensystem wie dem Schulwesen beleuchtet (vgl. Zlatkin-Troitschanskaia 2006; Benz 2004, S. 127; Altrichter 2011, S. 128; Kussau/Brüsemeister 2007, S. 26; Fuchs 2008, S. 22). Dieses erweiterte Steuerungsverständnis bettet die Schule in ein komplexes Gefüge des Zusammen- und Widerspiels von unterschiedlichen Akteuren aus der Bildungspolitik, Gesellschaft und der Privatwirtschaft ein.

Dem durch den Paradigmenwechsel in der Steuerungsphilosophie erwarteten Ziel der Optimierung der Schulqualität im Sinne von Effizienz und Gerechtigkeit steht die kritische Sicht auf die *Neue Steuerung* gegenüber (vgl. Böttcher/Rürup 2010, S. 61): Eine empirisch gestützte Kritik zeigt sich in der Diskrepanz zwischen erwarteten und tatsächlich umgesetzten Handlungsschritten sowie den nicht-intendierten Effekten der Neuen Steuerung. Während die Wirkung der intendierten Effekte empirisch nicht eindeutig abgesichert ist, mehren sich die Hinweise auf Risiken und nicht-intendierte Nebenwirkungen, welche sich negativ auf die Schulqualität und die Schülerleistungen auswirken können (vgl. Bellmann/Weiß 2009; Maag Merki 2010, S. 158f.). In Ländern mit langjähriger Erfahrung bezüglich neuer Steuerungsinstrumente kann bspw. spezifisch für das Instrument Schulinspektion das *window-dressing*, d. h. eine Akzentuierung der Außenwirkung und nicht einer realistischen Darstellung der tatsächlichen Schulsituation, beobachtet werden (vgl. Bellmann/Weiß 2009, S. 296). Mit der Implementierung neuer Steuerungsinstrumente zeigt sich eine

weitere – insbesondere im Zusammenhang dieser Arbeit – relevante Problematik, welche die Nutzung des Outputs betrifft. Die einzelnen Reformmaßnahmen entwickeln sich erst dann zu Steuerungsinstrumenten, wenn sie von den beteiligten Akteuren im beabsichtigten Sinne genutzt werden: Es entspricht nicht der Intention neuer Steuerungsinstrumente, „bereits in der Output-Messung die Systemoptimierung“ (Leschinsky/Cortina 2008, 47) zu sehen. Aus den Ergebnisrückmeldungen der outputorientierten Steuerungsinstrumente sollen Handlungen abgeleitet und schulische Prozesse angepasst werden und sich so dem erwarteten Output annähern (vgl. Maag Merki 2010, S. 153; Maag Merki/Emmerich 2011, S. 155-160). Noch überwiegend ungeklärt ist die Frage, ob und wenn ja, wie die einzelnen Akteure die Output-Informationen nutzen.

### 2.2.2. Evaluation und ihre Typisierung

Der Begriff *Evaluation* wird inflationär für alle Formen der Bewertung verwendet (vgl. Bauer 2004, S. 164). In der Pädagogik und Psychologie kann aus einer Vielzahl an Definitionsversuchen<sup>5</sup> als Kernpunkt des Begriffs ein „Bewerten von Sachverhalten“ zugrunde gelegt werden, aus dem eine „Optimierung von Entscheidungen“ (Wottawa 2006, S. 162) resultiert. Damit fungiert Evaluation als eine Form der Steuerung, indem sie gezielt in Handlungsprozesse eingreift (vgl. Bauer 2004, S. 161; Scheerens u.a. 2003, S. 3). Auf den Bildungsbereich übertragen spielt Evaluation neben der Selbstvergewisserung für Schulen sowie der Rechenschaftslegung über erreichte Ergebnisse und die Einhaltung von Standards in erster Linie als Instrument der Schulentwicklung eine Rolle (vgl. Burkard/Eikenbusch 2000; Peek 2002, S. 324). Für das Verständnis von *Evaluation* im Kontext dieser Arbeit sind in Anlehnung an Helmke (2009, S. 269) folgende Bestandteile kennzeichnend:

- (1) Systematische Erfassung
- (2) der Durchführung oder der Ergebnisse
- (3) eines Programms oder einer Maßnahme,
- (4) verglichen mit vorgegebenen Standards, Kriterien, Erwartungen oder Hypothesen
- (5) mit dem Ziel der Verbesserung des Programms oder der Maßnahme.

---

<sup>5</sup> Eine intensive Auseinandersetzung mit dem Begriff *Evaluation* ist bei Ditton (2010) und Kromrey (2002) zu finden.

Die Helmkesche Definition zugrunde legend sind somit für Evaluationen die systematische Herangehensweise, Wirksamkeitsüberprüfung und Handlungsoptimierung kennzeichnend. Ergänzend können allgemeine Standards für Evaluationen angeführt werden: Neben ihrem Nützlichkeitspostulat sollten Standards der Durchführbarkeit, Fairness bzw. Korrektheit und Genauigkeit eingehalten werden (vgl. DeGEval 2012; Böttcher 2009, S. 213f.).

Es existieren unterschiedliche Formen und Einsatzbereiche von Evaluationen im Bildungsbereich, welche sich international in 15 verschiedene Evaluations- und Monitoringtypen differenzieren und folgenden vier übergeordneten Kategorien zuordnen lassen (vgl. Scheerens u.a. 2003, S. 7-9; 33-50; Übersetzung übernommen aus Berkemeyer/Bos 2009, S. 531):

1. Evaluationsformen, die auf der Messung von Schülerleistungen basieren  
(z. B. Lernstandserhebungen; zentrale Abschlussprüfungen; PISA)
2. Bildungsstatistisch basierte Evaluationsformen  
(z. B. nationaler Bildungsbericht)
3. Beobachtungsbasierte Evaluationsformen  
(z. B. Schulinspektion; Schulseelbstevaluation)
4. Programm- und Lehrerevaluation  
(z. B. Begleitforschung zu Modellversuchen)

Tabelle 2 enthält eine Typisierung von Evaluationsformen im schulischen Bereich anhand ihrer Fragestellungen und entsprechenden Varianten in Anlehnung an Helmke (2009, S. 270f.):

**Tab. 2: Typisierung von Evaluationsformen**

<b>Was?</b>	Aspekte des Kontextes, Inputs, Prozesses, Produktes
<b>Wozu?</b>	<i>Formative Evaluationen</i> : erfolgen nach einzelnen Lernphasen zur Förderung <i>Summative Evaluationen</i> : erfolgen nach Abschluss des Themas/Lehrgangs/ Schuljahres zur abschließenden Bewertung
<b>Wer?</b>	<i>Interne Evaluation</i> : Evaluation von Akteuren einer Schule durch andere Akteure derselben Schule <i>Selbstevaluation</i> : Schulische Akteure überprüfen ihre eigene Tätigkeit <i>Externe Evaluation</i> : Evaluatoren sind keine Mitglieder der Schulen
<b>Wann?</b>	<i>Singuläre Evaluationen</i> : umfassen eine einzige Erhebung <i>Zyklische Evaluationen</i> : umfassen mehrere Erhebungen, z. B. PISA
<b>Woran?/ an welchem Maßstab?</b>	<i>Soziale Norm</i> : Vergleich mit anderen Schülern, Schulen, Bundesländern etc. <i>Kriteriale Norm</i> : Vergleich mit einem absoluten, a priori vorgegebenen Kriterium <i>Individuelle Norm</i> : Vergleich mit einer früheren oder nachfolgenden Leistung der gleichen Person (Pre- und Posttest)

Im Zusammenhang mit der vorliegenden Arbeit interessiert insbesondere die *Wer-Frage*, anhand welcher Evaluationen im schulischen Bereich in interne oder externe Formen zu untergliedern sind (vgl. auch Stamm 2003a, S. 54f.). Eine Selbstevaluation liegt vor, wenn bestimmte schulische Akteure ihre eigene Arbeit überprüfen. Davon unterscheiden sich interne Evaluationen, indem die Arbeit bestimmter schulischer Akteure von anderen Akteuren derselben Organisation, d. h. bspw. von Kollegen überprüft wird. Insgesamt stellen die schulinternen Evaluationsinstrumente, z. B. Hospitationen, kollegiale Fallberatungen und Analysen statistischer Schuldaten den Schulen und externen Akteuren systematische Informationen bereit, welche zur schulischen Selbst- und Systemsteuerung in Hinblick auf die Schulentwicklung herangezogen werden können (vgl. Berkemeyer/Müller 2010, S. 205; Holtappels 2011). Hingegen erfolgt eine externe Evaluation durch Akteure, die keine Mitglieder der zu evaluierenden Organisation sind. Dabei kann sie den Schulen zum einen von außen durch Anordnungen der Systemebene auferlegt werden oder sie wird von den Schulen selbst in Auftrag gegeben (vgl. Peek 2002, S. 326f.; Peek 2007, S. 149; Böttcher 2002, S. 186f.). Entsprechende externe Evaluationsinstrumente, insbesondere auch die Schulinspektion, werden im nächsten Abschnitt präsentiert.

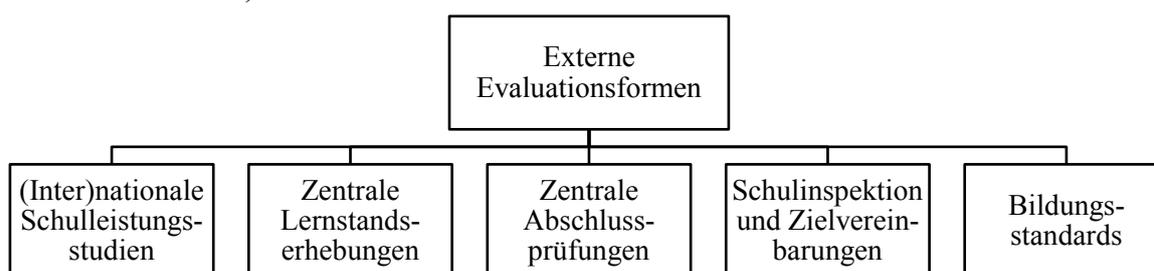
Die beiden Evaluationsformen (interne und externe) sollten nicht als dichotome und konkurrierende Elemente aufgefasst werden, bei denen die „moralische Sortierung intern = gut, pädagogisch vs. extern = böse, unpädagogisch“ (Terhart 2002, S. 104f.) vorgenommen wird. Grundsätzlich lassen sich drei Grundformen des Verhältnisses von interner und externer Evaluation unterscheiden: „Unverbundenheit, Passung/Verzahnung und Gegentakt“ (Rolff 2002, S. 342f.). Die optimale Wirkung wird bei der zweiten Variante erreicht, wenn interne und externe Evaluationsformen verknüpft werden, sich gegenseitig ergänzen und sich in einer Balance bewegen. Der ergänzende Einsatz beider Evaluationsformen erscheint zudem vor dem Hintergrund angebracht, dass der schulinternen Evaluation ein Glaubwürdigkeitsvorbehalt, d. h. der Vorwurf der Subjektivität und des „Verschleierns von Schwächen der eigenen Institution“ (van Ackeren/Klemm 2009, S. 161) anhaftet, genauso wie die externe Evaluation der Kritik eines nivellierenden Blicks auf die schulische Praxis ausgesetzt ist. Tatsächlich enthalten Evaluationsprozesse in der Praxis oftmals eine Mischform interner und externer Evaluationsformen und erschweren eine eindeutige

Trennung beider Evaluationsformen. Eine Mischform zeigt sich z. B. im Rahmen der Schulinspektion, indem sie als externe Evaluation auch den Stand der Selbst- und internen Evaluation der Schulen erhebt und durch die Ergebnismeldung wiederum die schulinterne Evaluation stimuliert. Im Folgenden werden Instrumente der externen Evaluation, darunter die Schulinspektion als ein Instrument externer Evaluation im schulischen Bereich skizziert.

### 2.2.3. Externe Evaluationsinstrumente

Von den externen Evaluationsinstrumenten wird neben der Messung der Schüler- und Schulleistungen eine Verbesserung der Schul- und Unterrichtsqualität erwartet, so dass sie sowohl die Funktion des Monitorings<sup>6</sup> als auch der Unterstützung erfüllen (vgl. Berkemeyer 2011). Die diesbezüglichen Wirkungserwartungen sind ambitioniert: „Externe Evaluation von Schulen ist derzeit in Deutschland dasjenige Instrument für die Steuerung im Schulsystem, von dem die stärkste Innovationswirkung erwartet wird“ (Döbert/Dedering 2008, S. 9).

Aktuell existieren im deutschen Schulwesen folgende schulexterne Steuerungsinstrumente: Internationale und nationale Schulleistungsvergleichsstudien (Large-Scale-Assessments), Bildungsstandards, zentrale Lernstandserhebungen, zentrale Abschlussprüfungen, Schulinspektionen und Zielvereinbarungen (s. Abb. 3; vgl. Maag Merki/Emmerich 2011, S. 152; Maag Merki 2010; Vock/Pant 2008).



**Abb. 3: Externe Evaluationsformen**

<sup>6</sup> Zur Einordnung des Begriffs wird folgende Definition zitiert: „Monitoring im Bildungswesen oder Bildungsmonitoring bezeichnet die kontinuierliche, datengestützte Information von Bildungspolitik und Öffentlichkeit über Rahmenbedingungen, Verlaufsmarkkmale, Ergebnisse und Erträge von Bildungsprozessen“ (Böttcher/Bos/Döbert/Holtappels 2008, S. 8).

*Internationale Schulleistungsvergleichsstudien* wie PISA, TIMSS und IGLU erheben Leistungen und Kompetenzen einzelner Schüler und -gruppen in verschiedenen Fächern, Jahrgangsstufen und Schulformen zum Zweck des Bildungsmonitorings (vgl. Hosenfeld/Zimmer-Müller 2010). Die *Bildungsstandards* beschreiben die Kompetenzanforderungen von Schülern zu einem bestimmten Zeitpunkt in ihrer Schullaufbahn. Ob und wie die Standards erreicht werden, wird anhand von *Lernstandserhebungen* kontrolliert und den Schulen mitsamt Referenzwerten übermittelt (vgl. Hartung-Beck 2009; van Ackeren/Bellenberg 2004). Als Messinstrument des Kompetenzstands einzelner Schüler, Klassen oder Jahrgänge anhand standardisierter Instrumente offenbaren Lernstandserhebungen bei genauerer Betrachtung über die einzelnen Bundesländer hinweg Unterschiede im Hinblick auf die Bezeichnungen (z. B. Kompetenztests, Vergleichsarbeiten, Orientierungsarbeiten, Diagnosearbeiten, VERA)<sup>7</sup>, die Testentwicklung, -durchführung, -auswertung und Rückmeldung der Ergebnisse. Die Lernstandserhebungen haben eine diagnostische Funktion, sollen Indizien für Veränderungen des pädagogischen Handelns von Lehrkräften liefern, finden schullaufbahnbegleitend in bestimmten Fächern statt und unterscheiden sich dadurch von *zentralen Abschlussprüfungen*, welche am Ende der Schullaufbahn (Sek I; Sek II) zu Zertifizierungs- und Selektionszwecken durchgeführt werden (vgl. van Ackeren/Klemm 2009, S. 163-168; Bellenberg 2008; Bosen/Büchter/Peek 2006; Maag Merki/Holmeier 2008; Klein/Kühn/van Ackeren/Block 2009). Während Bildungsstandards als Inputmerkmale die zu erzielenden Kompetenzen der Schüler in bestimmten Fächern und zu einem bestimmten Zeitpunkt festlegen, stellen zentrale Lernstandserhebungen, zentrale Abschlussprüfungen und Schulleistungsstudien eindeutig outputüberprüfende Verfahren dar. Hingegen erfasst die *Schulinspektion* neben dem schulischen Output unter Berücksichtigung der Inputs insbesondere die schulischen Prozesse und fungiert als Gelenkstück zwischen den Elementen der Input-, Prozess- und Outputsteuerung, indem sie z. B. auf Schulprogramme, zentrale Lernstandserhebungen und schulinterne Evaluationsergebnisse zurückgreift (vgl. Dederling/Müller/Bos 2008, S. 7f.).

<sup>7</sup> Zu den Zielsetzungen, Verfahren und zentralen Testergebnissen der Lernstandserhebungen 2004 in NRW wird auf Peek/Pallack/Dobbelstein/Fleischer/Leutner (2006) und Peek/Dobbelstein (2006) verwiesen.

Das Reforminstrument Schulinspektion lässt sich den kontrollbasierten Unterstützungssystemen zuordnen, indem ausgehend von der Rückmeldung der Schul- und Unterrichtsqualität Entwicklungsprozesse auf Schul- und Klassenebene angestoßen werden sollen, die zu einer Steigerung der Schulqualität führen sollen (vgl. Berkemeyer 2011). Als charakteristisch erweisen sich dabei die Impulsgebung von schulexterner Seite, die professionelle Durchführung der Schulinspektion sowie der anschließende Einsatz von Zielvereinbarungen. Mit der Implementierung von Schulinspektionen wird eine Kriseninduktion und Irritierung von Handlungsroutinen bezweckt:

Schulinspektionen stellen einen Eingriff in die Einzelschule dar, der in den seltensten Fällen völlig konfliktfrei ablaufen wird. Die Irritation des Einzelschulsystems stellt eine wesentliche Triebfeder für Entwicklungsprozesse dar (Lambrecht/Kotthoff/Maag Merki 2011, S. 188).

In den folgenden Kapiteln wird zunächst das Verfahren der Schulinspektion im europäischen Ausland dargestellt, dem eine eingehende Auseinandersetzung mit den Zielen, Funktionen und der Variantenvielfalt der Schulinspektion in den deutschen Bundesländern folgt. Daran schließt die Vorstellung der Qualitätsanalyse NRW an. Den Abschluss des theoretischen Teils zur Schulinspektion bildet die Systematisierung des diesbezüglichen Forschungsstands.

## 3. Schulinspektion

### 3.1. Schulinspektion im europäischen Ausland

Die Schulinspektion nimmt für bildungspolitische Entscheidungsträger und Praktiker international und insbesondere im europäischen Umfeld einen besonderen Stellenwert ein (vgl. Scheerens u.a. 2003, S. 321). Während die Schulinspektion in Deutschland relativ spät einsetzte und in den einzelnen Bundesländern bisher die systematische Implementierung der Schulinspektion im Vordergrund stand, können viele europäische Länder diesbezüglich auf eine längere Tradition und eine Vielzahl empirischer Untersuchungen zurückblicken (vgl. Scheerens u.a. 2003; Böttcher/Kotthoff 2007b; van Ackeren 2003). Einerseits zur Vermittlung impulsgebender Anregungen und andererseits durch die fehlende eigene empirische Befundlage zur Wirksamkeit der Schulinspektion nimmt ein Großteil der deutschsprachigen Publikationen zum Thema Schulinspektion Bezug auf europäische inspektionserprobte Länder und ihre diesbezüglichen wissenschaftlichen Erkenntnisse (vgl. z. B. Bos u.a. 2006, S. 85-91; Kotthoff/Böttcher 2010; Müller 2010).

Anhand der niederländischen und englischen Schulinspektionssysteme<sup>8</sup>, die sich sowohl in Bezug auf die Konzeption, Durchführung und ihren Stellenwert im Bildungssystem als auch ihre Wirkungen recht deutlich voneinander unterscheiden, soll ein selektiver Einblick in die europäischen Erfahrungen der Schulinspektion gegeben werden. Die Auswahl der beiden Länder wird folgendermaßen begründet: Die niederländische Schulinspektion gilt als Wegbereiter und Vorbild der Qualitätsanalyse NRW, welche in dieser Arbeit eine zentrale Stellung einnimmt. Hingegen weist die englische Schulinspektion z. T. erhebliche Unterschiede zum niederländischen und folglich auch zum nordrhein-westfälischen Verfahren auf und soll als kontrastierender Gegenpol fungieren. Beide Länder haben gemein, dass die Schulinspektion seit vielen Jahren praktiziert wird und somit bereits zahlreiche empirische Evidenzen vorliegen (vgl. z. B. Kotthoff 2003; van Ackeren 2003). Im Sinne des „policy borrowings“

---

<sup>8</sup> Zum Hintergrund des historisch-gesellschaftlichen und bildungspolitischen Kontextes beider Länder wird auf Sparka (2007, S. 64-92) sowie Döbrich, Schnell und Sroka (2008, S. 171-185; S. 192-200) verwiesen.

(Brauckmann u.a. 2010, S. 132) hat sich Deutschland bei der Implementierung der Schulinspektion durchaus an den inspektionserfahrenen Ländern orientiert, jedoch – wie Böttcher und Keune (2011, S. 123) kritisieren – ohne internationale Forschungsbefunde systematisch zu analysieren und kritische politische Debatten zu führen.

Dass sich im europäischen Raum und darüber hinaus eine Praxis der systematischen externen Evaluation<sup>9</sup> der Schulen durchgesetzt hat, ist durch die bildungspolitischen Reformen, insbesondere der schulischen Autonomiebewegung und Outputsteuerung des Schulwesens begründet. Diese wurden von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) und der Europäischen Union (EU) maßgeblich vorangetrieben. Für den europäischen Erfahrungsaustausch über die Entwicklung der Schulinspektionen wurde 1995 die SICI (Standing International Conference of Central and General Inspectorates of Education) mit mittlerweile 32 Mitgliedsinspektoraten gegründet. Zu den Mitgliedsstaaten zählt auch Deutschland mit den sechs Bundesländern Bayern, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz und Sachsen (vgl. Döbrich/Schnell/Sroka 2008, S. 168; SICI 2015). Trotz der immer weiter voranschreitenden Verbreitung der Schulinspektion ist es schwierig, gemeinsame Merkmale der internationalen Schulinspektionen festzulegen:

(...) it is difficult to find a common definition of inspection criteria or to reach an agreement on a particular school inspection model across different countries, given the differences in national traditions, culture and aspirations – among other factors that may influence a country's educational goals (Scheerens u. a. 2003, S. 324).

Ungeachtet der aufgezeigten Problematik formulieren Scheerens, Glas und Thomas (2003, S. 44) drei Merkmale von Schulinspektionen, die für die meisten europäischen Länder zutreffen: Es finden flächendeckende, nach einem festgelegten Turnus veranstaltete Schulbesuche statt, es werden Standards festgelegt und systematische, quasi-wissenschaftliche Methoden der Datengewinnung eingesetzt.

---

<sup>9</sup> Zur Verbreitung der Schulevaluation in Europa wird auf die Exekutivagentur Bildung, Audiovisuelles und Kultur (P9 EURYDICE 2009, S. 66-75) verwiesen.

### 3.1.1. Inspectie van het Onderwijs in den Niederlanden

In den Niederlanden wird die Aufsicht über das Bildungswesen im Auftrag des Ministeriums von der *Inspectie van het Onderwijs*, der Inspektion von Bildung und Unterricht wahrgenommen. Neben der Primar- und Sekundarstufe umfasst die Aufsicht auch die Berufs- und Erwachsenenbildung sowie den Hochschulbereich. Bezüglich des Kontextes des niederländischen Schulsystems sind zwei Besonderheiten hervorzuheben: Zum einen verfügen die Schulen über eine hohe Autonomie, die ihre religiösen, ideologischen und pädagogischen Prinzipien und den Umgang mit finanziellen Ressourcen umfasst. Zum anderen gehören die meisten Schulen privaten Schulträgern an und stehen durch die elterliche Schulwahlfreiheit in starker Konkurrenz zueinander (vgl. van Bruggen 2006, S. 110; van Ackeren 2003, S. 109f.; Schouten 2008, S. 235; Weerts 2007, S. 43).

Die zentrale Aufgabe der nationalen Inspektion als teilunabhängige Agentur ist es, die Qualität der Schulen auf System- und Einzelschulebene zu bestimmen, zu überwachen und zu befördern (vgl. van Ackeren 2003, S. 173). Damit unterliegt ihr ein breites Aufgabenspektrum mit den Feldern Kontrolle, Evaluation und Stimulierung der Einzelschulen in Richtung Qualitätsentwicklung sowie Rechenschaftslegung und öffentliche Berichterstattung über die System- und Einzelschulqualität. Die Kernfunktionen werden dabei in der Garantiefunktion (dass Schulen Unterricht in befriedigender Qualität leisten) und der Stimulierungsfunktion gesehen (vgl. Döbrich u.a. 2008, S. 194f.). Die Qualitätsbeurteilung der Schulen erfolgt auf der Grundlage eines Qualitätsverständnisses in Form von Kriterien und Indikatoren und unter Berücksichtigung des jeweiligen kulturellen, sozialen und demografischen Kontextes. Eine Einteilung schulischer Kontextmerkmale in sechs Klassen wird in den Niederlanden von dem staatlichen Test- und Prüfungsinstitut CITO vorgenommen. Insofern kann die Inspektion bei der Schulbewertung sowohl auf die Klassifizierung der Kontextmerkmale als auch auf die Testergebnisse von CITO zurückgreifen. Jede Schule wird somit vor dem Hintergrund der Referenzgruppe beurteilt, sowohl bezüglich der Lernergebnisse als auch aller anderen Qualitätsdimensionen des Inspektionsrahmens. Die Inspektionsberichte der Einzelschulen enthalten neben den Beurteilungen Empfehlungen für Maßnahmen zur Qualitätsverbesserung und werden elektronisch auf der Webseite des Ministeriums in Form einer Qua-

litätskarte veröffentlicht (vgl. Schouten 2008, S. 237; Weerts 2007, S. 46). Von jeder Schule gibt es in den Niederlanden eine Qualitätskarte, auf der allgemeine Informationen zu der Institution vermerkt sind, wodurch die Vergleichbarkeit mit anderen Schulen ermöglicht wird.

In den Niederlanden hat das Inspektionswesen eine lange, etwa 200-jährige Tradition, befindet sich jedoch insbesondere seit den 1990er Jahren und der Verabschiedung des Schulaufsichtsgesetzes (WOT) im Jahr 2002 in einem starken Wandel (vgl. Janssens 2007; Weerts 2007, S. 45). Zunächst lag der Schwerpunkt der Schulinspektion auf der intensiven Evaluation der Primarstufe und weitete sich 1997 auch auf den Sekundarbereich aus. Es wird unterschieden zwischen der *regulären Schulaufsicht (RST)* und der *integralen Schulaufsicht (IST)*, die sich durch eine unterschiedliche Besuchsfrequenz auszeichnen. Die RST wird in einem reduzierten Verfahren (eintägiger Schulbesuch; zwei- bis dreijähriger Zyklus) mit dem Fokus auf den Kernmerkmalen guten Unterrichts durchgeführt. Zeigen sich die schulischen Ergebnisse dabei als unauffällig, bleibt die Schule im RST-Zyklus, während bei negativen Ergebnissen eine intensivere Schulinspektion in Form der IST erfolgt. Zusätzlich existiert die *proportionale Schulaufsicht (PST)*, bei der ein Großteil der Verantwortung für die Qualitätssicherung den Schulen in Form von Selbstevaluation übertragen wird. Die Proportionalität stellt mittlerweile das wichtigste Prinzip der niederländischen Schulinspektion dar und ist durch folgenden Kerngedanken gekennzeichnet: „Nur so viel Schulinspektion wie nötig“ (van Bruggen 2006, S. 115). Gemeint ist damit das Verhältnis von bereits gemessener Schulqualität sowie der Intensität und Frequenz nachfolgender Inspektionen: Schulen mit hoher Qualität und positiver Entwicklung brauchen weniger und oberflächlichere Inspektionen als schwächere Schulen. Damit weitet sich der Inspektionszyklus starker Schulen auf drei bis vier Jahre aus. Bei schwachen Schulen erfolgen schnellere, tiefere und präzisere Inspektionen. Als zweites Kennzeichen der proportionalen Schulinspektion gilt das Verhältnis zwischen der internen und externen Evaluation der Schulen: Je zufriedenstellender die schulinterne Evaluation erfolgt, desto mehr wird die externe Kontrolle gelockert und nur in begrenztem Umfang durchgeführt. Indem sich die Schulinspektion auf eine stichprobenweise Beurteilung beschränkt, soll sich ihre Funktion in Richtung einer metaevaluativen Rolle entwickeln.

Insgesamt ist die niederländische Schulinspektion durch zwei Besonderheiten gekennzeichnet: Zum einen wird viel Wert auf eine gesicherte und akzeptierte Professionalität der Inspektoren gelegt. Vorausgesetzt wird, dass sie über langjährige Praxiserfahrungen als Lehrer oder Schulleiter verfügen, ein neunmonatiges Training absolvieren und sich an fünf Tagen im Jahr in der eigens dafür gegründeten Inspektionsakademie fortbilden. Ihre Aufgaben beschränken sich auf die Diagnose der Schulqualität und Formulierung von Empfehlungen, welche Handlungsfelder die Schulen in Angriff nehmen sollten, während eine professionelle Beratung von ihnen nicht geleistet wird. Zum anderen erfahren schulische Unterstützungsstrukturen im Anschluss an die Schulinspektionen eine große Bedeutung (vgl. Sparka/van Bruggen 2009): Diese zeigt sich an dem Verhältnis von 200 Inspektoren (und 230 Mitarbeitern) zu 2.500 Mitarbeitern in den Unterstützungseinrichtungen mit dem Grundsatz „Verbesserung und Innovation sind wichtiger als Diagnosen“ (van Bruggen 2006, S. 121).

In den letzten Jahren zeichnete sich eine Tendenz der niederländischen Schulinspektion dahingehend ab, dass die Flexibilisierung der Schulinspektion in Form von Teilinspektionen, die lediglich bestimmte Bereiche in den Blick nehmen, zunimmt und dass vollständige Schulbesuche generell abnehmen. Dies betrifft Schulen, die schon zwei oder drei Mal als *gut* in einer Inspektion bewertet wurden. Hingegen werden *Risikoschulen* (etwa 4 % aller Schulen sind Risikoschulen und 15 % sind gefährdete Schulen) von der Inspektion besonders begleitet und unter verschärfte Aufsicht gestellt (vgl. Luginbuhl/Webbink/de Wolf 2009, S. 221). Zeigen sich Unterstützungsleistungen und Vereinbarungen über den Verbesserungsprozess als wirkungslos, kann es zu einer Schulschließung kommen. Aktuell gibt es Überlegungen, den Schulen, Schulträgern und Eltern die Verantwortung über die Aufsicht in Form einer Supervision zu übergeben (vgl. Döbrich u.a. 2008, S. 200; Janssens 2007, S. 131).

In einer Gegenüberstellung der Stärken und Schwächen der niederländischen Inspektion lässt sich in Anlehnung an van Bruggen (2006, S. 122f.) konkludieren, dass ihre Stärken insbesondere in der Flexibilität der proportionalen Schulinspektion sowie der Professionalität der Inspektoren liegen. Als Schwäche kann hingegen das recht allgemeine Schulinspektionsverfahren gesehen werden, da es nur wesentliche Qualitätsaspekte beleuchtet und nach zwei bis drei Jahren eine Monotonie im System entsteht.

### 3.1.2. Inspection of schools in England

Die englische Schulinspektion ist mit ihrer Größe und Macht einzigartig unter den Inspektionssystemen Europas (vgl. van Ackeren 2003, S. 167). Im Vergleich mit anderen Ländern gilt sie als äußerst rigide und kompromisslos, da ihre Befunde öffentlich bekanntgegeben werden und weitreichende Konsequenzen für die inspizierten Schulen haben. Daher kann von einem high-stakes Schulinspektionssystem gesprochen werden (vgl. Cheesman 2007, S. 35; Kotthoff/Böttcher 2010, S. 305).

Zum besseren Verständnis des aktuellen englischen Schulinspektionssystems werden charakteristische Merkmale des Schulwesens sowie die ursprüngliche Schulinspektion knapp skizziert: Das Schulsystem in England ist durch ein an Bildungsstandards orientiertes nationales Curriculum, flächendeckende Tests der Schüler in den Kernfächern (Englisch, Mathematik und Naturwissenschaften) am Ende bestimmter Schulstufen sowie die öffentliche Bekanntgabe der Testergebnisse jeder Schule gekennzeichnet (vgl. Döbrich u.a. 2008, S. 172; Key 2007a, S. 22; Kotthoff 2003). Charakteristisch ist zudem, dass in England großes öffentliches Interesse an den Leistungen des Bildungssystems und der Schulen besteht, über die in der Lokalpresse detailliert berichtet wird (vgl. Key 2007b, S. 29f.). Die Medien sind es auch, welche Ranglisten (*league tables*) mit den schulischen Test- und Inspektionsergebnissen veröffentlichen. Diese hohe Transparenz ermöglicht es Eltern bei der Schulwahlentscheidung auf alle zentralen Informationen über die Schulen zurückzugreifen.

Seit über 150 Jahren existiert die *Her Majesty's Inspectorate* (HMI), welche sich lange Zeit durch sporadische Schulinspektionen auszeichnete (vgl. Key 2007a, S. 21). Durch die New-Public-Management-Reform wurde das HMI 1992 in seiner bisherigen Rolle durch *Ofsted* (Office for Standards in Education), dem Büro für Bildungsstandards, abgelöst. Das Ziel der Umstellung von HMI auf Ofsted war die Ausweitung der Frequenz und Intensität sowie ein kriteriengeleitetes Vorgehen des englischen Schulinspektionssystems. Als vom Ministerium unabhängige staatliche Einrichtung übernimmt Ofsted die Planung und Durchführung von Schulbesuchen der staatlich finanzierten Schulen und eines Teils der Privatschulen. Zusätzlich sind vom Ofsted beauftragte örtlich zuständige private Dienstleister (Regional Inspection Service Providers, RISP)

mit der organisatorischen Vorbereitung und Umsetzung der Inspektionen be-  
traut. Seit 2007 wurde der Zuständigkeitsbereich von Ofsted einhergehend mit  
der neuen Bezeichnung *Office for Standards in Education, Children's Services  
and Skills* ausgeweitet. Seitdem übernimmt es auch Inspektionen in vor-  
schulischen Einrichtungen, Lehrerbildungsinstituten, Adoptionsagenturen,  
Kinderheimen etc. Somit ist Ofsted seit 2007 für sämtliche Bereiche des staat-  
lich verantworteten Bildungssystems mit Ausnahme des Hochschulwesens ver-  
antwortlich (vgl. Döbrich u.a. 2008, S. 177).

Die englische Schulinspektion verfolgt insgesamt unter dem Leitgedanken  
„improvement through inspection“ (Key 2007a, S. 24) folgende Ziele: Verbes-  
serung der Qualität und Standards der Erziehungs- und Bildungseinrichtungen  
sowie Information der Öffentlichkeit, Regierung und Eltern über die Lage  
staatlich verantworteter Erziehung und Bildung (vgl. Döbrich u.a. 2008, S.  
174). Dazu werden in Inspektionen die Bildungsarbeit der Schule, die erzielten  
Bildungsstandards, die geistige, moralische und soziale Entwicklung der Schü-  
ler sowie der Beitrag der Schule zum allgemeinen Wohlbefinden ihrer Schüler  
geprüft. Ein wichtiges Prüfkriterium stellt zudem die Führungsqualität und das  
Management der Schule dar.

Ofsted beschäftigt etwa 5.300 Inspektoren, von denen aber die Mehrheit ne-  
benberuflich arbeitet (vgl. Key 2007a, S. 27). Diese müssen die Anforderung  
erfüllen, längere Zeit in einer verantwortlichen Position im Bildungswesen be-  
schäftigt gewesen zu sein und eine mehrmonatige Vorbereitung mit anschlie-  
ßender Testinspektion absolviert zu haben. Wie in den Niederlanden sind die  
englischen Schulinspektoren verpflichtet, sich an fünf Tagen im Jahr fortzubil-  
den. Die hauptberuflichen Ofsted-Inspektoren sind für das Management, die  
Qualitätssicherung der Inspektion und die Berichterstattung zuständig. Seit  
2005 nehmen sie auch an Schulbesuchen teil, welche zuvor vorrangig von den  
nebenberuflich tätigen Inspektoren durchgeführt wurden.

Das Schulinspektionsverfahren wurde seit der Gründung von Ofsted mehr-  
fach verändert, wobei besonders im Jahr 2005 durch massive Kürzungen des  
Budgets die Organisation der Schulinspektion modifiziert werden musste. Die  
Reformen von 2005 sind durch das Leitmotiv *downsizing* charakterisiert (vgl.  
Döbrich u.a. 2008, S. 177-180): Die Anzahl der Inspektionsteam-Mitglieder  
wurde reduziert, die Dauer der Schulbesuche wurde auf ein bis zwei Tage ver-

kürzt und die Bewertungsskala wurde aufgrund unzureichender Ausschöpfung von einer ehemals siebenstufigen auf eine vierstufige Notenskala dezimiert (vgl. Key 2007b, S. 20). Zudem wurde der Inspektionsbericht mit wesentlichen Befunden und klaren Empfehlungen auf sechs Seiten deutlich verschlankt und wird den Schulen nun zeitnah vorgelegt. Während vor 2005 eine exzessive Informationsgewinnung und zahlreiche Unterrichtsbesuche im Fokus der Schulinspektion standen, werden seit der Reform die wesentlichen Aspekte von Schulqualität beleuchtet und weniger Unterrichtsbesuche durchgeführt. Zudem verkürzte sich die Benachrichtigungsfrist der Schulen vor der Inspektion von ehemals sechs bis zehn Wochen auf zwei bis fünf Werktage, um den Druck der Schulen zu minimieren und eine möglichst realistische Schulsituation vorzufinden. Neben diesen Verschlankungen wurde dagegen der Anteil von Inspektionen unter der Leitung oder Mitwirkung eines Ofsted-Inspektors erhöht, ebenso wie die Schulbesuchsfolge von einem sechsjährigen auf einen dreijährigen Rhythmus wechselte.

Im Rahmen der Inspektionsvorbereitung wird ein *pre-inspection briefing* verfasst: Um Schwerpunktthemen und Hypothesen zur anstehenden Inspektion zu verfassen, werden der letzte Inspektionsbericht, die Leistungs- und Kontextdaten der Schüler und der Schule sowie – falls vorhanden – ein von der Schule ausgefülltes Formular zur Selbstevaluation herangezogen (vgl. Scheerens u. a. 2003, S. 328; 333-335; Döbrich u.a. 2008, S. 179). Im Zentrum der Schulvisitation stehen Gespräche mit allen schulischen Akteuren, Dokumentenanalysen sowie Unterrichtsbesuche. Direkt im Anschluss an die Schulvisitation wird den schulischen Akteuren das vorläufige Ergebnis vorgestellt. Der i. d. R. innerhalb der Folgeweche zugesandte Inspektionsbericht enthält eine Gesamteinschätzung der Effizienz der Schule und eine Bewertung der einzelnen Leistungsbereiche anhand der vierstufigen Notenskala von 1 (ausgezeichnet; outstanding) bis 4 (ungenügend; inadequate). Zusätzlich wird an die Schüler ein Brief adressiert. Die Schule kann zum Inspektionsbericht Stellung nehmen. Schließlich wird der Bericht zeitnah auf der Ofsted-Webseite veröffentlicht.

Der nachfolgende Umgang mit den Ergebnissen ist in einem hohen Ausmaß von der Gesamtbewertung der Schule determiniert. Als „ausgezeichnet“ bewertete Schulen werden öffentlich gewürdigt und erhalten Plaketten zur Anbringung am Schulgebäude. Allen anderen Schulen werden abhängig von der

Schwere der Probleme unterschiedliche Handlungsmöglichkeiten präsentiert und Sanktionen eingeleitet: Werden Schulen als „ungenügend“ bewertet, können sie entweder anhand der Kategorie *Aufforderung zur Verbesserung* (*notice to improve*) oder *besondere Maßnahmen* (*special measures*) klassifiziert werden. Im Schuljahr 2006/07 waren insgesamt 6 % der inspizierten Schulen von diesen beiden Kategorien des Verbesserungsbedarfs betroffen (vgl. Döbrich u.a. 2008, S. 180; Davies/Lee/Postlethwaite/Tarr/Thomas/Yee 1999, S. 130). Zur ersten Kategorie zählen Schulen, die den Schülern keine akzeptablen Erziehungs- und Bildungsstandards vermitteln, jedoch aus eigener Kraft diesbezügliche Verbesserungen einleiten können. Diese Kategorie umfasst auch Schulen, die akzeptable Erziehungs- und Bildungsstandards vermitteln, dies jedoch unter den von ihnen zu erwartenden Möglichkeiten vollbringen (vgl. Döbrich u.a. 2008, S. 181). Die Schulen dieser Kategorie erstellen einen Aktionsplan, erhalten nach sechs Monaten einen Zwischenbesuch und nach einem Jahr erfolgt eine Re-Inspektion. Werden Schulen unter *besondere Maßnahmen* gestellt, bedeutet dies, dass sie inakzeptable Erziehungs- und Bildungsstandards vermitteln und über keine eigenen Ressourcen für Verbesserungen verfügen. In diesen Schulen ist dringender Verbesserungsbedarf gefordert. Sie erhalten anschließend Unterstützungsleistungen in Form von Beratung sowie personeller und sächlicher Hilfe. Wenn diese Maßnahmen nicht greifen, kann im Sinne eines *Fresh Starts* ein Führungswechsel oder ein Austausch des Kollegiums erfolgen. In letzter Konsequenz kommt es zur Schulschließung, von der jährlich etwa 12 bis 20 Schulen betroffen sind (vgl. Key 2007a, S. 28).

Die neusten Trends der englischen Schulinspektion zeigen eine deutliche Annäherung an den internationalen Entwicklungstrend. Seit einigen Jahren werden im Vorfeld der Schulinspektion die Selbstevaluationen der Schulen als Ausgangspunkt der Inspektion berücksichtigt. Zudem hält die Flexibilität der Inspektionsanlässe, -dauer und -intensität Einzug in das englische Schulinspektionsmodell. Demnach können bei vorangegangenen zufriedenstellenden Leistungsergebnissen Kurzinspektionen (*reduced tariff inspections*, RTI) durchgeführt werden oder bei aufgedeckten Problemen in der schulischen Selbstevaluation umfassendere Inspektionen erfolgen.

Abschließend werden die Stärken und Schwächen der englischen Schulinspektion gegenübergestellt: Als Stärke ist die absolute Transparenz über Test-

und Inspektionsergebnisse zu sehen, die als Teil des Verantwortungsprinzips Eltern und Schüler mit so vielen Informationen wie möglich versorgt. Die Identifizierung und Verbesserung der Problemschulen (*failing schools*) wird als Erfolg der englischen Schulinspektion verbucht (vgl. Key 2007b, S. 17) und ist nicht zuletzt auf die Unterstützungsleistungen für schwache Schulen der lokalen Behörden zurückzuführen. Die Inspektion überprüft dabei den Fortschritt des Verbesserungsprozesses und bewertet die schulischen Möglichkeiten, weitere Fortschritte zu erzielen. Die Mehrheit der Schulen schafft es dann, innerhalb von zwei Jahren aus dieser Kategorie aufzusteigen und als *gute* Schulen bezeichnet zu werden (vgl. Key 2007a, S. 29f.). Durch dieses Vorgehen hat sich die Qualität des englischen Bildungssystems in den letzten 20 Jahren deutlich verbessert (vgl. Key 2007b, S. 35): Die Test- und Prüfungsergebnisse haben sich positiv entwickelt und die Mehrzahl der Schulen wird als *gut* oder besser eingestuft. Eine positive Einschätzung der englischen Schulinspektion gibt folgendes Zitat wieder:

Inspection has been identified as a particular powerful catalyst for change and improvement in England. Inspection is a clear policy lever that can be used to help promote school improvement and raise standards for disadvantaged students (Sammons 2006, S. 51f.).

Ob die im Rahmen der *special measures* eingeleiteten Veränderungen allerdings nachhaltig sind oder nur kurzfristige und kosmetische Auswirkungen haben, bleibt fraglich (vgl. Perryman 2006). Zu den negativen Effekten der rigiden englischen Schulinspektion zählen Stress und Besorgnis auf Seiten der Schulleitung und Lehrer (vgl. Perryman 2007; Key 2007a, S. 30). Neben dem Effekt der Demoralisierung kann eine negative Schulbewertung und ihre Öffentlichmachung (polemisch auch als *name, blame and shame policy* bezeichnet) existenzielle Konsequenzen für die Schulen haben, wenn aufgrund der Wettbewerbssituation Eltern ihre Kinder von diesen Schulen nehmen und in erfolgreicheren Schulen anmelden oder es zu einer Fluktuation des Personals kommt. Letztlich wird die englische Schulinspektion mit der Frage nach Alternativen konfrontiert: „(...) could [there] be other, more effective and less costly, ways of helping schools to improve their practice and outcomes?“ (Scheerens u.a. 2003, S. 331).

### 3.1.3. Vergleichende Gegenüberstellung der niederländischen und englischen Inspektionsverfahren

Bei der vergleichenden Gegenüberstellung der niederländischen und englischen Inspektionsverfahren werden zunächst die Gemeinsamkeiten hervorgehoben: Sowohl in den Niederlanden als auch in England existieren mächtige Schulinspektionsbehörden (Inspectie van het onderwijs und Ofsted), welche einen vom jeweiligen Bildungsministerium relativ unabhängigen Status besitzen und der jeweiligen Regierung rechenschaftspflichtig über die Lage und den Fortschritt des Bildungssystems sind (vgl. Böttcher/Kotthoff 2007a, S. 12; Döbrich u.a. 2008, S. 228). Durch ihre lange Tradition nehmen die Schulinspektionen beider Länder eine Vorbildfunktion für andere, insbesondere europäische Länder ein. Ein weiteres gemeinsames Kennzeichen ist, dass sich beide Inspektionssysteme an Standards orientieren, die sämtliche Verfahrensschritte des gesamten Inspektionsvorgangs regeln. Zudem liegen den Schulbewertungen transparente Qualitätsrahmen mit zahlreichen Kriterien zugrunde. Schließlich zeigt sich eine weitere Gemeinsamkeit beider Länder darin, dass die schulischen Inspektionsergebnisse online veröffentlicht werden und damit der interessierten Öffentlichkeit und den Medien zugänglich sind.

Neben den genannten Gemeinsamkeiten überwiegen jedoch die Unterschiede der Schulinspektionen beider Länder. Die deutlichsten sind in den Steuerungsphilosophien begründet (vgl. Dederling 2012, S. 77). In England findet eine starke Orientierung an der Wettbewerbssteuerung statt, die zwar auch in den Niederlanden vorzufinden ist, jedoch in einer abgeschwächten Form. Mit dieser Steuerungsphilosophie sollen primär über Druck Effizienzsteigerungen im Bildungswesen erzielt werden. Während in England nur als *ungenügend* bewertete Schulen staatliche Unterstützung erhalten, existiert in den Niederlanden ein komplexes Unterstützungssystem, welches den Schulen unabhängig von ihrem gemessenen Qualitätsstand gewährt wird (vgl. Sparka/van Bruggen 2009).

Werden Kontrolle und Qualitätsentwicklung als gegensätzliche Pole des Merkmals *Zielsetzung der Schulinspektion* gesetzt, so lässt sich England eindeutig dem Pol *Kontrolle* zuordnen, während sich die Niederlanden zwischen beiden Polen bewegen (vgl. Dederling 2012, S. 78). Im Spektrum zwischen ei-

nem differenzierten System an Sanktionen und keinen Konsequenzen steht England eher auf der Sanktionsseite, indem es die Inspektionsergebnisse veröffentlicht und damit schwache Schulen brandmarkt (vgl. Böttcher/Kotthoff 2007a, S. 16). Auch wenn die Inspektionsberichte in den Niederlanden ebenfalls im Internet veröffentlicht werden, besteht ihr Anliegen in der Transparenz der gemessenen Ergebnisse und nicht in einer öffentlichen Diskreditierung der Schulen.

Die Unterschiede beider Schulinspektionssysteme zeigen sich auch in ihren verschiedenen Zuständigkeitsbereichen: In England ist seit 2005 eine deutliche Erweiterung des Zuständigkeitsbereichs der Inspektion zu konstatieren, da sie seitdem auch Vorschulen und Lehrerbildungsinstitutionen umfasst. Während diese Institutionen in den Niederlanden nicht in den Verantwortungsbereich der Schulinspektion fallen, wird dort dafür der Hochschulsektor bewertet. Ein weiterer Unterschied besteht in dem Verhältnis der Schulinspektion zu anderen Systemen der schulischen Rechenschaftslegung (vgl. Böttcher/Kotthoff 2007a, S. 13): England stützt sich bei der Schulinspektion in starkem Ausmaß auf die nationalen Testergebnisse der Schüler. In den Niederlanden werden zwar auch die CITO-Testergebnisse herangezogen, jedoch spielt die schulische Selbstevaluation im Inspektionsprozess eine zentrale Rolle. Der Grundsatz, dass schulische Selbstevaluation und externe Schulinspektion komplementäre Prozesse darstellen, ist für beide Länder kennzeichnend.

Zusammenfassend lässt sich folgender internationale Trend der Schulinspektion konstatieren: Die Inspektion entwickelt sich zunehmend zu einer Metaevaluation. Je mehr innerhalb der Schulen valide Selbstevaluationsinstrumente existieren und genutzt werden, desto eher kommt die externe Evaluation in stark abgemilderter Form zum Einsatz. Damit folgt der Trend einer zunehmenden Flexibilität des Schulinspektionsverfahrens und verlässt starre und nivellierende Verläufe. In den folgenden zwei Kapiteln wird auf der Grundlage der Schulinspektionskonzepte und -erfahrungen im europäischen Ausland das deutsche Schulinspektionssystem beschrieben und spezifisch am Beispiel Nordrhein-Westfalens dargestellt.

## 3.2. Schulinspektion in Deutschland

In einem ersten Schritt wird eine begriffliche Konkretisierung der Schulinspektion in Deutschland vorgenommen. Daran schließt die Skizzierung der Funktionen und Zielsetzungen an, die mit der Implementierung der Schulinspektion verbunden sind. Anschließend werden bundesländerübergreifend gültige Merkmale der Schulinspektion dargestellt, die in der vergleichenden Erörterung der Schulinspektionskonzepte und -verfahren der einzelnen Bundesländer münden, bevor spezifisch auf die Qualitätsanalyse NRW eingegangen wird.

### 3.2.1. Definition

Auch wenn in den einzelnen Ländern unterschiedliche Begrifflichkeiten und Verfahren der Schulinspektion existieren, kann als Versuch der Bündelung folgende grundlegende Definition angeführt werden:

Schulinspektion bedeutet Blick in oder auf Einzelschulen auf der Grundlage einer Zusammenschau vorhandener, intern und/oder extern gewonnener Daten. Die Inspektionen erfolgen meist auf der Grundlage von Verfahren mit wiederkehrenden Standardelementen, die in öffentlich zugänglichen Handbüchern beschrieben sind (Maritzen 2008, S. 87).

Ergänzend wird folgende Definition herangezogen:

[Bei der Schulinspektion] handelt es sich um ein neues, von der bisherigen Tätigkeit der staatlichen Schulaufsicht abgegrenztes Verfahren der systematischen evaluativen Bestandsaufnahme der Arbeitsbedingungen, Arbeitsweisen und Arbeitsergebnisse der einzelnen Schulen (Rürup 2008, S. 467).

In beiden Fällen handelt es sich um eine recht allgemeine und weite Definition der Schulinspektion mit jeweils unterschiedlichen Nuancierungen. Der Begriff Schulinspektion wird eher umschrieben, in einzelne Elemente zergliedert und von herkömmlichen Evaluationsverfahren abgegrenzt. Bei Maritzens Versuch einer Definition wird diese Unschärfe im Vokabular deutlich („Die Inspektionen erfolgen *meist...*“) und erschwert es, eine allgemeingültige Definition der Schulinspektion zu skizzieren. Mit einer anderen Nuancierung verdeutlicht Rür-

rups Definition zum einen den neuartigen Charakter und die Abgrenzung der Schulinspektion von herkömmlichen schulaufsichtlichen Tätigkeitsfeldern, zum anderen wird Maritzens „Blick in oder auf Einzelschulen“ spezifiziert, indem eine Differenzierung in schulische Arbeitsbedingungen, Arbeitsweisen und Arbeitsergebnisse erfolgt.

Trotz dieser Unterschiede beinhalten beide Definitionen die wesentlichen Kernpunkte der Schulinspektion (vgl. Maritzen 2008, S. 86; Rürup 2008, S. 470): Beide Definitionen sehen die Einzelschule im Fokus des Verfahrens und nicht die Makroebene oder einzelne schulische Akteure. Es handelt sich generell um eine systematische, standardisierte und datengestützte Evaluation der einzelschulischen Gesamtleistung, genauer der schulischen Prozessqualitäten. Die für die Qualitätsbewertung der Schulen erforderlichen Daten existieren entweder bereits oder werden eigens im Rahmen der Inspektion durch den Einsatz verschiedener Instrumente erhoben. Die Standardisierung bezieht sich zum einen auf die Instrumente und Abläufe der Schulinspektion, zum anderen auf transparente Kriterienkataloge der Schulqualität, welche eine normative Wirkung ausüben. Zu den zentralen Instrumenten der Schulinspektion gehören die Analyse schulischer Dokumente (z. B. Schulprogramme), ein Schulrundgang sowie leitfadengestützte Interviews mit der Schulleitung, den Lehrkräften, Schülern, Eltern, dem nicht-lehrenden Schulpersonal, ggf. Schulträgern und dualen Partnern. Im Zentrum des mehrtägigen Schulbesuchs stehen Unterrichtsbeobachtungen anhand standardisierter Beobachtungsbögen. Formal erfolgt der Schulinspektionsabschluss mit der Übermittlung eines schriftlichen Inspektionsberichts an die Schule. Somit fungiert die Einzelschule als Hauptadressat der Inspektionsergebnisse und ist für den Abschluss von Zielvereinbarungen verantwortlich.

### **3.2.2. Ziele und Funktionen**

Beide zu Beginn des Kapitels zitierten Definitionen der Schulinspektion beinhalten nur eine äußere Beschreibung, ohne die mit der Schulinspektion verbundenen Ziele und den intendierten Nutzen zu konkretisieren. Gemäß Landwehr (2011, S. 37) können vier Hauptfunktionen der Schulinspektion unterschieden werden: Wissensgewinnung, Entwicklungsanstoß für die Schulentwicklung, Rechenschaftslegung/Kontrolle sowie Normendurchsetzung. Die Gewinnung

datengestützten Wissens gilt als grundlegende Funktion der Schulinspektion, wobei weniger die Vermittlung neuartiger Erkenntnisse als vielmehr die Offizialisierung oftmals bereits vorhandenen Wissens intendiert wird. Im Rahmen der Normendurchsetzungsfunktion rücken durch die Ankündigung der Inspektion spezifische, von der Schulinspektion bestimmte Themenfelder in den Fokus der Schule. Die wesentliche Intention und Stimulierungsfunktion der Schulinspektion liegt eindeutig in der *Qualitätsentwicklung* der Schulen: Dabei soll sie als Instrument der Impulsgebung, Optimierung und Unterstützung der Schul- und Unterrichtsentwicklung der Einzelschule fungieren. Es ist vorgesehen, dass die von der Schulinspektion ausgehenden Impulse in verbindlichen Zielvereinbarungen zwischen Schulen und Schulaufsicht münden. Zugleich dienen die Ergebnisrückmeldungen der Rechenschaftslegung über den Qualitätsstand der Schulen und des Schulsystems gegenüber der Schulöffentlichkeit und Schulaufsicht. Im Gegensatz zu den Niederlanden und England herrscht in Deutschland ein geringes Maß der Wettbewerbssteuerung und es folgen keine gravierenden Konsequenzen bei negativen Inspektionsresultaten. Zudem werden die Ergebnisse nicht im Internet veröffentlicht, sondern sind nur an die Schulöffentlichkeit adressiert.

Neben der wissenschaftlichen Beschreibung der Zielsetzung(en) der Schulinspektion findet sich auch in den offiziellen Verlautbarungen der Ministerien oder Schulverwaltungen eine vielfältige Funktionszuweisung. Demnach erfüllen Schulinspektionen in Anlehnung an Maritzen (2008, S. 87) folgende Funktionen:

- Spiegel- oder Feedbackfunktion aus externer Sicht
- Qualitätssicherungsfunktion im Rahmen staatlicher Gewährleistungsverantwortung
- Unterstützungsfunktion insbesondere für Schulleitungen und Lehrkräfte
- Impuls- oder Katalysatorfunktion für die Schul- und Unterrichtsentwicklung
- Erkenntnisfunktion hinsichtlich der Wirkungen schulischer Arbeit

Die Funktionen lassen sich auf verschiedenen Ebenen verorten: einerseits auf der Mesoebene (Einzelschule) und andererseits auf der Makroebene (Gesamtsystem). Maritzen deutet diesen Funktionenmix und die fehlenden Zielprioritä-

ten als deutlichen Hinweis auf erhebliche Unsicherheiten hinsichtlich der Zielsetzung von Schulinspektionen sowie auf strategische oder inhaltlich programmatische Unklarheiten (vgl. Maritzen 2008, S. 87). Auch Bos u.a. (2006, S. 97) stellen in ihrer Bestandsaufnahme zur Schulinspektion in den deutschen Bundesländern fest, dass die Implementierungsschritte sowie die Funktionen und Ziele von Inspektionen „ein gewisses Maß an Diffusität und Unbestimmtheit“ aufweisen. Des Weiteren konstatieren sie, dass die Zielformulierungen der Schulinspektion auch bei weiteren Maßnahmen der Qualitätssicherung zur Anwendung kommen (vgl. ebd., S. 108). Als Instrument der Qualitätssicherung müsse sich die Schulinspektion jedoch von anderen Instrumenten hinsichtlich ihrer Ziele und Funktionen abgrenzen und andererseits sich mit ihnen kompatibel zeigen – so eine Forderung Maritzens (vgl. Maritzen 2011, S. 88).

Diese o. g. Kritikpunkte abmildernd wird darauf hingewiesen, dass sich Evaluationsverfahren generell anhand der vier Ziele Erkenntnisgewinn, Kontrolle, Entwicklung und Legitimation unterscheiden lassen, die sich zwar analytisch trennen lassen, jedoch in der Praxis deutlich zusammenhängen (vgl. Stockmann 2007, S. 36-39). Entscheidend ist das Setzen einer Ziel- und Funktionspriorität, welche die anderen Ziele unterordnet und eine Systematik festlegt.

In enger Verknüpfung zu den Funktionen stehen Annahmen über die Wirksamkeit und möglichen Wirkungen der Schulinspektion. In erster Linie kann die Schulinspektion dann als wirksam angesehen werden, „wenn tatsächlich Schulentwicklung stattfindet oder (...) die zusammengetragenen Informationen dazu dienen, die Aufsicht und Kontrolle über die Schule zu ermöglichen bzw. zu verbessern“ (Husfeldt 2011, S. 265). Es sind dabei drei Optionen möglich, wie Schulentwicklung angestoßen werden kann: Entweder über Wettbewerb, Konsequenzen oder Einsicht (vgl. Böttger-Beer/Koch 2008). Einsicht bedeutet in diesem Zusammenhang Konfrontation mit unerwarteten Informationen und liegt als alleiniger Wirkmechanismus der Schulinspektion in Deutschland zugrunde. Durch den Reformimpuls und die „Unruhe“ sollen Handlungsrouninen der schulalltäglichen Praxis durchbrochen und Veränderungspotentiale freigesetzt werden, welche schließlich zu einer systematischen Weiterentwicklung der Qualität schulischer und unterrichtlicher Prozesse führen.

### 3.2.3. Qualitätsverständnis der Schulinspektion

Ohne auf die spezifischen Qualitätsrahmen einzelner Bundesländer einzugehen, soll im Folgenden aufbauend auf die Ausführungen zur Qualität von Schule und Unterricht in Kapitel 2.1 geklärt werden, welches Qualitätsverständnis der Schulinspektion zugrunde liegt und wie es sich in das Qualitätsmodell im Bildungsbereich (vgl. Ditton 2000; s. Abb. 1) einbettet.

Schulqualität im Rahmen der Schulinspektion erfordert in erster Linie ein gemeinsames Verständnis darüber, was gute – im Sinne effektiver – Schulen und guten Unterricht auszeichnet. Während über die Notwendigkeit eines normativen Rahmens in den einzelnen Bundesländern Einigkeit besteht, herrscht hinsichtlich der Bestimmung spezifischer Indikatoren und Kriterien kein Konsens. Im Prinzip ist eine beliebige Anzahl und inhaltliche Variation der Qualitätsbereiche, -aspekte und -indikatoren denkbar. Die Vielzahl an unterschiedlichen Operationalisierungen von Schulqualität liegt darin begründet, dass keine allgemeingültige normative Bestimmung von Schulqualität vorliegt. Insgesamt ist die Schulinspektion auf empirische Erkenntnisse zur Effektivität von Schulen angewiesen, die eine Deduktion spezifischer Kriterien zulassen.

Die empirischen Erkenntnisse in unterschiedlichem Ausmaß aufgreifend hat jedes Bundesland einen verbindlichen Qualitätsrahmen entwickelt, welcher die Schulqualität auf unterschiedlichen Ebenen dimensioniert. Bei einer – zunächst analytischen Sichtung – zeigt sich, dass die existierenden Qualitätsrahmen der Länder eine erhebliche Tiefe und einen hohen Strukturierungsgrad aufweisen. Hierbei wird eine übersichtliche Anzahl an Qualitätsdimensionen (i. d. R. drei bis sieben) in eine höhere Anzahl an Qualitätsmerkmalen untergliedert, welche wiederum in eine nicht mehr überschaubare Anzahl an Kriterien ausdifferenziert werden (z. B. umfasst das Qualitätstableau in NRW 153 Kriterien; s. MSW NRW 2006). Dieses Verständnis schließt mit ein, dass es sich nicht um statische oder finale Modelle handelt, da diese kontinuierlich weiterentwickelt und optimiert werden. Die Tendenz geht in einigen Bundesländern zur Minimierung der Anzahl der Kriterien, um damit der Übersichtlichkeit und Handhabbarkeit – sowohl für die Inspektoren als auch für die schulischen Akteure – Rechnung zu tragen (vgl. Wendt/Bos 2011, S. 231). Insgesamt erfüllt das in den Qualitätsrahmen manifestierte Qualitätsverständnis zwei Funktionen: Zum

einen bietet es im Rahmen von Schulinspektionen eine Grundlage für die Erfassung der Leistung der Einzelschule. Zum anderen erfüllt es eine Orientierungsfunktion für Schulen und Lehrkräfte und die Möglichkeit der schulinternen Nutzung, indem administrativ gesetzte Erwartungen an Schulqualität und konkrete Evaluationskriterien definiert werden (vgl. Rürup 2008, S. 469).

Im Rahmen der konzeptionellen Grundlage der Schulinspektion stehen Merkmale der Prozess- und Ergebnisqualität im Vordergrund, die durchaus dem wissenschaftlichen Diskussionsstand der Schuleffektivitäts- und Unterrichtsforschung und politisch vorgegebenen Schulqualitätsansprüchen entsprechen. Zusätzlich werden Inputmerkmale mitberücksichtigt. Inhaltlich weisen nahezu alle Qualitätsrahmen der Länder identische Dimensionen auf: Ergebnisse der Schule, Unterricht, Schulkultur, Professionalisierung der Lehrkräfte, Pädagogische Führung und Qualitätsentwicklung (vgl. Botz/Heese/Schirow 2011, S. 114-116). Dabei stellt insbesondere das durch die Schuleffektivitätsforschung gestützte Modell (s. Abb. 1) für die Systematisierung von Schul- und Unterrichtsqualität ein Orientierungsschema dar und folgt eindeutig dem Prozess-Produkt-Paradigma. Dies ermöglicht, dass die Schulen zu den schulischen Kontext-, Prozess- und Ergebnismerkmalen eine kriterienorientierte Bewertung durch geschulte Inspektoren erhalten.

Für die Entwicklung der Qualitätsrahmen beanspruchen die Länder nicht zuletzt wissenschaftliche Expertisen. In diesem Zusammenhang gehört es mittlerweile zum anerkannten Usus, die von Helmke und Meyer entwickelten „Zehnerkataloge“ (Meyer 2010, S. 168) als Grundlage für die Entwicklung der Unterrichtsbeobachtungen heranzuziehen (s. Tab. 1; vgl. Dobbstein 2008). In einer Überprüfung des niedersächsischen Orientierungsrahmens auf sein Verständnis von Unterrichtsqualität stellt Kiper (2008) heraus:

Wenn wir die vorgestellten Überlegungen [didaktische Modelle; Ideen des deutschen Bildungsrates; Pädagogische Psychologie; empirische Unterrichtsforschung; Reformpädagogik; Arbeitswelt; naive Annahmen über Lernen; Rahmensteuerung durch die KMK] mit Blick auf den Niedersächsischen Orientierungsrahmen Schulqualität bündeln, so wird sichtbar, dass in diesen aus den verschiedenen Richtungen Kategorien aufgenommen wurden (S. 110; Aufzählungen in Klammern aus Kiper 2008).

Zusammenfassend wird konstatiert, dass es sich bei dem Qualitätsverständnis der Schulinspektion und seiner Implikation in die jeweiligen Qualitätsrahmen um ein zumindest teilweise wissenschaftsbasiertes Konstrukt handelt. Die Länder bemühen sich ausdrücklich darum, auf wissenschaftlichen Erkenntnissen gestützte Qualitätsrahmen zu erstellen (vgl. Bonsen/Bos/Rolff 2008, S. 15). Trotz der Zweckdienlichkeit schulischer Qualitätsrahmen darf nicht außer Acht gelassen werden, dass Schulqualität als facettenreiches Konstrukt zu verstehen ist, welches durch empirische Verfahren nie vollständig erfasst werden kann, da sich die Messinstrumente durch einen fragmentarischen Charakter auszeichnen. Überdies hat jedes Bundesland sein eigenes normatives Qualitätsverständnis guter Schulen und guten Unterrichts, so dass sich die Operationalisierung einzelner Dimensionen teilweise deutlich unterscheidet. Die jeweiligen Qualitätsrahmen weisen zwar auch eine relativ große Schnittmenge auf, erschweren jedoch die Vergleichbarkeit gemessener Schulqualität in den einzelnen Ländern und eine systematische Konstruktvalidierung. Somit stellen Studien zur Erfassung und zum systematischen Vergleich der Qualitätsrahmen der Schulinspektionen im Hinblick auf ihre Selektion, Operationalisierung und Gewichtung von Qualitätsdimensionen und -subdimensionen noch Forschungsdesiderate dar.

Von einigen Wissenschaftlern wird ein kritisches Bild der Qualitätsrahmen gezeichnet: So bezeichnet Böttcher (2009, S. 209) die Qualitätsrahmen im Kontext der Schulinspektion als „schriftliche Taxonomien, die (...) lediglich intuitiv Qualitätsbereiche (...) benennen“. Seiner Meinung nach gehen sie nicht über die Ebene der Deskription hinaus. Des Weiteren bemängeln Gärtner und Pant (2011, S. 13):

Die Frage der Inhaltsvalidität der Qualitätsrahmen der Länder in Deutschland ist ungeklärt. Insgesamt wird in den Qualitätsrahmen Schul- und Unterrichtsqualität kaum mit dem Verweis auf aktuelle Ergebnisse der Bildungsforschung definiert, welche den empirischen Beleg dafür liefern, dass die hier beschriebenen Qualitätsaspekte tatsächlich in einem Zusammenhang zum Bildungs- und Erziehungsauftrag der Schulen stehen.

Die Schulinspektion als Instrument des Bildungsmonitorings kann nur dann als wissenschaftlich seriös gelten, wenn die Messergebnisse als verlässliche Indizien für die Schul- und Unterrichtsqualität fungieren und aus ihnen Konsequenzen zur Qualitätsverbesserung abgeleitet werden können (vgl. Klieme/Tippelt 2008, S. 13). Diesem Kerngedanken folgend entwickeln sich das „Grundlagenwissen über Qualität (...) einerseits, konkrete Qualitätssicherung in der Praxis andererseits (...) – im optimalen Fall – durch- und miteinander“ (ebd.). Eine erste dahingehende Tendenz zeigt sich in einem abgestuften Modell der Unterrichtsqualität, welches auf der Grundlage von 32 Hamburger Schulinspektionen und 2.240 Unterrichtsbeobachtungen mit Methoden der Item-Response-Theorie entwickelt wurde (vgl. Pietsch 2010; Pietsch/Schulze/Schnack/Krause 2011). Es ermöglicht, einzelne Merkmale von Unterricht aufeinander aufzubauen. Die Abstufung von Unterrichtsqualität wird dabei inhaltlich folgendermaßen beschrieben (vgl. Pietsch 2010, S. 140):

1. Stufe: Lernklima und pädagogische Strukturen sichern
2. Stufe: Klassen effizient führen und Methoden variieren
3. Stufe: Schüler motivieren, aktives Lernen und Wissenstransfer ermöglichen
4. Stufe: Differenzieren, Schüler wirkungs- und kompetenzorientiert fördern

Auf der Grundlage dieser Stufen kann zum einen die Gesamtheit des Unterrichts an einer Schule qualitativ eingeordnet werden und zum anderen lassen sich konkrete Entwicklungsschritte zur Unterrichtsverbesserung, d. h. zur Erreichung der nächsthöheren Stufe, projektieren. Der Gewinn des Modells liegt zudem in seiner Anwendungsorientierung, von der Schulpraktiker und die Schulöffentlichkeit profitieren können (vgl. ebd., S. 140). Damit werde – so Pietsch – eine standardisierte, regional geeichte Messskala gebildet, die auch für die Bildungsadministration leicht nachvollziehbare Rückmeldungen zur Unterrichtsqualität beinhaltet (vgl. ebd., S. 142). Dies zeigt eindrücklich, dass sich aus der Praxis der Schulinspektion neue wissenschaftliche Erkenntnisse bezüglich guter Schulen und guten Unterrichts generieren lassen und das gewonnene Datenmaterial großes Potential für weiterführende Forschung vermuten und erwarten lässt.

### 3.2.4. Schulinspektion in den deutschen Bundesländern

Die Schulinspektion in Deutschland ist durch eine erhebliche Variantenvielfalt, die sich in der Pluralität der Länderkonzepte deutlich zeigt, gekennzeichnet. Auch wenn sie im europäischen Vergleich ein noch recht junges Instrument ist, kann sie auf eine längere Tradition zurückblicken: Ihre Anfänge sind bereits im 18. Jahrhundert zu erkennen und um die Wende zum 19. Jahrhundert erfolgten flächendeckende Visitationen der Schulen in Deutschland, wenn auch nicht in annähernd so systematischer Form, wie sie aktuell durchgeführt werden (vgl. Geißler 2008). In diesem Kapitel werden Spezifika und Differenzen der Schulinspektionsverfahren und -konzepte in den Bundesländern präsentiert, um das Spektrum und die Divergenz des Instruments zu beleuchten und für den Gebrauch der Bezeichnung *Schulinspektion in Deutschland* zu sensibilisieren.

Die nahezu synchrone und rasante Implementierung der Schulinspektion in den einzelnen Bundesländern erweckt den irreführenden Eindruck, es handele sich – analog zu den bundesweit geltenden Bildungsstandards – um eine Vereinbarung der KMK<sup>10</sup>. Auch der Eindruck, die Schulinspektion als unmittelbare Reaktion auf die PISA-Befunde zu verstehen, erweist sich als nicht zutreffend (vgl. Döbert/Rürup/Dedering 2008, S. 137). Gegen diese Auffassung spricht, dass Bremen als erstes Bundesland im Jahr 1995 – fünf Jahre vor der ersten PISA-Studie – eine Schulinspektion in Ansätzen einführte. Vielmehr gilt als übereinstimmender Auslöser der Implementierung der Schulinspektion die Notwendigkeit schulischer Qualitätsverbesserung. Die nahezu zeitgleiche Einführung der Schulinspektion und die grundsätzliche Ähnlichkeit der Schulinspektionskonzepte in den Bundesländern geht zu einem bedeutenden Anteil auf „policy borrowing“ (Brauckmann u.a. 2010, S. 132) zwischen den Bundesländern zurück. Da die Zuständigkeit für das Bildungswesen im föderalistisch strukturierten Deutschland den einzelnen Bundesländern obliegt, existieren jeweils unterschiedliche Konzepte und Verfahren der Schulinspektion. Das für Deutschland eher zügige Implementierungstempo der Schulinspektion interpretiert Maritzen als „Symptom für das Erschrecken der *Steuerleute* über die chronische Wissensarmut im Schulwesen“ (Maritzen 2006, S. 13).

---

<sup>10</sup> Eine Erörterung der Angleichungstendenzen länderübergreifender sowie differenzierender länderspezifischer Steuerungskonzepte ist bei Böttcher und Rürup (2010) zu finden.

Mittlerweile sind mehrere Bestandsaufnahmen der deutschen Schulinspektion vorgenommen worden (vgl. z. B. Bos/Holtappels/Rösner 2006; Bos/Dedering/Holtappels/Müller/Rösner 2007). So bedeutend die ersten Bestandsaufnahmen als Grundlage und Wegbereitung, bspw. für die Forschung, auch waren, so schnell waren sie in ihren Einzelheiten z. T. wieder hinfällig: Angesichts der hohen Dynamik und Variabilität des neuen Instruments waren sie nur kurz nach ihrer Publikation nicht mehr aktuell. Die zuletzt veröffentlichten systematischen Synopsen der Konzepte und Implementierungen der Schulinspektion in den einzelnen Bundesländern wurden im Jahr 2008 (vgl. Döbert u.a. 2008) und 2009 (vgl. Hovestadt 2009) erstellt. Beide Bestandsaufnahmen beinhalten eine vergleichende Perspektive der Schulinspektion nach bestimmten Aspekten<sup>11</sup>.

Seitdem die Implementierungsphase etwa zwischen den Jahren 2006 und 2008 abgeschlossen wurde, ist die Schulinspektion weitestgehend in die Phase der Konsolidierung übergegangen. Schleswig-Holstein hat sich diesem Entwicklungsprozess als bislang einziges Bundesland entzogen, da es zum Herbst 2009 die Schulinspektion aus Kostengründen abgeschafft hat. Nach der Phase der Implementierung und Konsolidierung befindet sich die Schulinspektion in einigen Bundesländern bereits in der nächsten Entwicklungsstufe: Der erste Inspektionsturnus ist in den Ländern Bayern, Berlin, Brandenburg und Hessen abgeschlossen und der zweite Zyklus wurde eingeleitet. Manche Bundesländer, bspw. Brandenburg, nehmen den zweiten Zyklus zum Anlass, ein modifiziertes Inspektionsverfahren mit inhaltlichen und technischen Optimierungen einzuführen. Teilweise ändert sich mit dem zweiten Turnus auch die Bezeichnung der Schulinspektion, z. B. wird sie in Hessen als *Zweite Inspektion* bezeichnet.

Aus der Konvergenz der Implementierung der Schulinspektion in den deutschen Bundesländern lässt sich nicht auf übereinstimmende Konzepte und Verfahren schließen (vgl. Maritzen 2008, S. 86). Es überwiegen eher die Unterschiede: „Jedes Land hat sich ‚sein‘ spezifisches Instrument der externen Eva-

---

<sup>11</sup> Folgende Aspekte wurden von Döbert u.a. (2008) vergleichend untersucht: Zielsetzung, normativer Rahmen, Institutionalisierung, Ausstattung und Arbeitsstand, Schulauswahl und (zeitlicher) Verlauf, eingesetzte Verfahren und Instrumente sowie Rückmeldung, Rolle der Schulaufsicht und Datenverwendung. Bei Hovestadt (2009) ist eine vergleichende Erörterung folgender Merkmale zu finden: institutionelle Zuständigkeit, Evaluatoren, Referenzrahmen, Ergebnisse der Evaluation, Phase nach der Evaluation und Verfahren bei schlechtem Evaluationsergebnis.

luation von Schulen (...) konstruiert“ (Döbert u.a. 2008, S. 150). Die Gemeinsamkeiten liegen eher auf globaler Ebene. Werden spezifische Merkmale der einzelnen Inspektionskonzepte fokussiert, sind z. T. beachtliche Differenzen festzustellen. Eine Nutzung der länderspezifisch aggregierten Daten der Schulinspektion im Rahmen eines Bildungsmonitorings auf nationaler Ebene wird damit erschwert oder gar unmöglich (vgl. ebd.).

Im Folgenden werden Gemeinsamkeiten und Unterschiede der Schulinspektionen in den einzelnen Bundesländern anhand bestimmter Merkmale skizziert. Auffällig sind die unterschiedlich verwendeten Begrifflichkeiten für die Bezeichnung der Schulinspektion. In lediglich vier Bundesländern (Berlin, Hamburg, Hessen, Niedersachsen) wird die Bezeichnung *Schulinspektion* verwendet, während ein Großteil der Bundesländer die Bezeichnung *Externe Evaluation* oder damit vergleichbare Bezeichnungen (z. B. *Fremdevaluation* in Baden-Württemberg) gewählt hat. Allein das Bundesland Nordrhein-Westfalen nimmt mit der Bezeichnung *Qualitätsanalyse* eine Sonderrolle ein und konnotiert mit dieser Begriffswahl eine eher schonende und milde Form des Verfahrens.

Als weiterer länderspezifischer Unterschied gilt die institutionelle Anbindung der Schulinspektion entweder in Referaten der Bildungsministerien, Landesinstituten oder Qualitätsagenturen (vgl. ebd., S. 147f.). Dadurch, dass keine unabhängigen Institutionen geschaffen wurden, wird in Deutschland eine eher konservative Lösung der Institutionalisierung verwirklicht (vgl. Maritzen 2008, S. 91). Analog zur institutionellen Anbindung wird die Personalrekrutierung gehandhabt, indem sie durch die Anlehnung an bereits bestehende Strukturen – in diesem Fall die Schulaufsicht – vollzogen wird. Die Inspektoren handeln zwar personell getrennt von der herkömmlichen Schulaufsicht, rekrutieren sich aber z. T. aus ihr. Die Auswahl der Schulinspektoren vollzieht sich nach unterschiedlichen Gesichtspunkten, z. B. im Rahmen eines Assessment-Centers (Berlin). In der Mehrzahl der Bundesländer sind hauptamtliche Inspektoren beschäftigt (z. B. Hamburg, Hessen, Niedersachsen), jedoch existieren auch Modelle mit nebenberuflichen Inspektoren oder semiprofessionellen Mitgliedern, bspw. Wirtschaftsvertretern oder Eltern (z. B. Bayern). Die Inspektion wird in Teams durchgeführt, die sich aus zwei bis vier Mitgliedern zusammensetzen. Die einzelnen Länder haben einen Inspektionszyklus zwischen drei und sechs, überwiegend von vier Jahren, vorgesehen (vgl. Döbert u.a. 2008, S. 140).

In ihrer Synopse der Schulinspektion in den Bundesländern konstatieren Döbert, Rürup und Dederling (2008, S. 145), dass vor allem in der organisatorischen und prozessualen Inspektionspraxis große Differenzen bestehen. Diese zeigen sich in der Schulauswahl, im Umfang, Inhalt, in den Erhebungsverfahren und Ergebnisberichten. Im Hinblick auf die Schulauswahl werden das Los- bzw. Zufallsprinzip, das Prinzip der Freiwilligkeit oder ein kriteriengeleitetes Verfahren herangezogen. In Bremen bspw. bestimmt die Schulaufsicht die zu inspizierenden Schulen und in NRW gilt das Prinzip der Regionalität, dem zufolge zunächst alle Schulen einer bestimmten Region evaluiert werden (vgl. Döbert u.a. 2008, S. 141). In einigen Ländern herrscht auch die Praxis, mehrere Formen der Schulauswahl zu kombinieren (z. B. Bayern).

In allen Bundesländern kommen Dokumentenanalysen, Schulrundgänge, Unterrichtsbeobachtungen und Interviews mit den schulischen Akteuren als Instrumente der Schulinspektion zum Einsatz. Zudem werden in 13 Bundesländern standardisierte, überwiegend onlinegestützte schriftliche Befragungen der schulischen Akteure durchgeführt. Die Bundesländer Hessen, Sachsen und Saarland verwenden ausschließlich Papierfragebögen. Der Umfang der Schulbesuchstage ist von der Schulgröße abhängig und variiert zwischen einem Tag (z. B. Baden-Württemberg) bis zu fünf Tagen (Niedersachsen). Die Unterrichtseinsichtnahmen erfolgen anhand standardisierter Beobachtungsbögen. Sie werden überwiegend in zwanzigminütigen Sequenzen, aber auch in ganzen Unterrichtsstunden (z. B. Bayern und Sachsen, z. T. Thüringen) vorgenommen. In Tabelle 3 ist der Umfang der Unterrichtsbeobachtungsbögen in sechs Bundesländern anhand der Anzahl der Seiten und den zu bewertenden Indikatoren dargestellt (aus: Müller/Pietsch 2011, S. 43). Es zeigt sich, dass es z. T. beträchtliche Unterschiede der Anzahl der Beobachtungsbögen gibt und dass der Unterricht in NRW am differenziertesten bewertet wird.

**Tab. 3: Vergleich des Umfangs der Unterrichtsbeobachtungsbögen in sechs Bundesländern**

Land	Berlin	Brandenburg	Hamburg	Hessen	Niedersachsen	NRW
Anzahl der Seiten	1	1	2	3	1	4
Anzahl der Items	16	18	30	13	20	42

Als Grundlage für die Bewertungen der Schulen liegen allen Bundesländern Qualitätskataloge zugrunde, allerdings mit jeweils unterschiedlichen Bezeichnungen (z. B. *Orientierungsrahmen Schulqualität* in Brandenburg oder *Qualitätsstableau der externen Evaluation* in Bayern). Die einzelnen Qualitätskataloge enthalten unterschiedlich viele Qualitätsbereiche/-felder/-dimensionen in einer Varianz von drei bis sieben. Darüber hinaus werden diese übergeordneten Bereiche z. T. in zwei nachgeordnete Ebenen von Aspekten, Kriterien oder Indikatoren unterteilt (vgl. Überblick über formale und inhaltliche Merkmale der Qualitätsrahmen der Bundesländer in Botz/Heese/Schirow 2011, S. 114-116). Trotz der Differenzen in der Quantität und Benennung weisen die Qualitätsrahmen ähnliche Inhalte auf (vgl. ebd.):

- Ergebnisse der Schule
- Unterricht
- Schulkultur
- Professionalisierung der Lehrkräfte
- Qualitätsentwicklung

Insgesamt werden in den Qualitätsrahmen die Prozessqualitäten der Schulen hervorgehoben. Das Kerngeschäft der Schule – der Unterricht – nimmt entgegen seiner zentralen Bedeutung in den Qualitätsrahmen keinen entsprechend größeren Raum ein, sondern stellt einen von mehreren Qualitätsbereichen dar. Die Bewertung der einzelnen Qualitätsdimensionen erfolgt anhand unterschiedlicher Bewertungskategorien. Diese reichen von drei- bis sechsstufigen Skalen bis hin zu offenen Bewertungskategorien (vgl. Döbert u.a. 2008, S. 146).

Das Ergebnis der Schulinspektion ist in allen Bundesländern ein schriftlicher Bericht der Befunde, allerdings mit unterschiedlichem Umfang und unterschiedlicher Darstellungsart. Der Bericht enthält entweder eine reine Ergebnisdarstellung oder darüber hinausgehende konkrete Empfehlungen an die Schule. Die letztgenannte Variante ist in den vier Bundesländern Baden-Württemberg, Bayern, Schleswig-Holstein und Thüringen vorzufinden. In Bremen werden Empfehlungen nur auf Wunsch der Schule mitgeteilt und stellen somit kein verbindliches Element der Abschlussberichte dar. Im Saarland endet die Schulinspektion erst mit der Beratung der Schulen und nicht bereits mit der Ergebnismeldung in Form eines Inspektionsberichts (vgl. Hovestadt 2009).

Nach der Ergebnismeldung und somit mit dem Abschluss der Schulinspektion wechselt die Zuständigkeit: Es beginnt ein überwiegend schulinterner Prozess der Schulentwicklung. I. d. R. ist ein Abschluss von Zielvereinbarungen zwischen den Schulen und der Schulaufsicht vorgesehen (vgl. Kapitel 4.1 und 4.2). Einige Bundesländer verpflichten die Schulaufsicht zur Unterstützung der Schulen bei der Umsetzung der Zielvereinbarungen (z. B. Thüringen).

Die z. T. beträchtlichen Differenzen der Schulinspektionsmerkmale in den einzelnen Bundesländern erschweren es, ein zusammenhängendes Bild der deutschen Schulinspektion zu zeichnen. Mittlerweile existieren einige Typisierungen von Schulinspektionen, die eine Bündelung bestimmter Merkmale vorschlagen. Dazu zählt z. B. die von Maritzen (2008) vorgenommene Differenzierung der vier idealtypischen Modellvarianten der Inspektionsverfahren (vgl. Maritzen 2008, S. 91): Netzwerk-, Experten-, Inspektions- und Zertifizierungsmodell. Ein weiterer Typisierungsvorschlag wird von Rürup (2008, S. 472-474) unterbreitet: Demnach können Ländermodelle danach differenziert werden, ob sie eher vorgeben, wissenschaftlich beeinflusstes Wissen zu generieren (wahrheitsorientiertes Modell) oder ob sie in ihre Evaluationen pragmatische Sichtweisen einbeziehen (angemessenheitsorientiertes Modell). Seinen Typisierungsvorschlag einschränkend verweist Rürup (2008, S. 475) darauf, dass sich kein Bundesland mit allen Merkmalen eindeutig einer der beiden Typisierungen zuordnen lässt. Diese Typen stellen eher gegensätzliche Extreme dar, deren Grenzen fließend verlaufen. Als einzige Ausnahme einer annähernd eindeutigen Zuordnung benennt Rürup das Bundesland Nordrhein-Westfalen, welches am offensichtlichsten für eine Wahrheitsorientierung steht. In den Bundesländern Baden-Württemberg, Bayern, Bremen und Thüringen hingegen überwiegen in der Schulinspektionspraxis eher Merkmale, welche sich den angemessenheitsorientierten Inspektionsmodellen zuordnen lassen.

Zusammenfassend zeigt sich für die Schulinspektion in Deutschland, dass sie in den einzelnen Ländern äußerlich durchaus große Ähnlichkeiten und Übereinstimmungen aufweist, in Detailfragen jedoch bundeslandabhängig von Heterogenität gekennzeichnet ist und unterschiedlich konzeptioniert und implementiert wird. Folglich kann kaum von *der* Schulinspektion in Deutschland gesprochen werden, sondern vielmehr von mehreren unterschiedlichen Schulinspektionsmodellen in den einzelnen Bundesländern.

### 3.2.5. Kritik an der Schulinspektion

Die mit der Implementierung der Schulinspektion intendierten Ziele müssten per se den Schluss nahelegen, es handele sich bei der Schulinspektion um ein allseits akzeptiertes und wirksames Instrument der Schulentwicklung, da es positiv konnotiert ist und auf ein allgemeinverbindliches Ideal (gute Schulqualität) abzielt. Jedoch steht dieser Annahme einige Kritik gegenüber, die überwiegend am Verfahren und an den Ergebnissen der Schulinspektion geübt wird: So haftet z. B. den zwanzigminütigen Unterrichtsbesuchen der Vorwurf an, sie seien zu kurz und nicht repräsentativ. Auch sehen Kritiker in der Schulinspektion ein kostenintensives Verfahren, welches eine Disproportionalität der erzielten Ergebnisse zu den hohen Kosten aufweist. Im Folgenden sollen einige kritische Punkte<sup>12</sup> der Schulinspektion beleuchtet werden, die zum einen den strukturellen Mangel, zum anderen die zurückgemeldeten Informationen und schließlich den Widerstand der Lehrkräfte zum Gegenstand haben.

Die größte Kritik der Schulinspektion wird in ihrem strukturellen Mangel und ihrer Anschlusslosigkeit gesehen. In den meisten Bundesländern endet die Schulinspektion formal mit der Bereitstellung des Inspektionsberichts. Genau an dieser Stelle sehen Kritiker die größte Schwachstelle des Instruments: Die fehlende Nachsorge und Unterstützung im Anschluss an die Rückmeldung der Ergebnisse lässt sich mit der ärztlichen Anamnese, auf die keine Therapie und Medikation folgt, vergleichen. Die Schulen fühlen sich häufig mit den Befunden allein gelassen und äußern Beratungs- und Unterstützungsbedarf. Mittlerweile wird in den Bundesländern in Form von Zielvereinbarungen zwischen Schulen und Schulaufsicht ein Anstoß zur Umsetzung der Ergebnisse in schulentwicklungsfördernde Handlungen gegeben. Die Forderung nach Unterstützungsleistungen finanzieller und personeller Art bleibt hingegen bestehen.

Zudem wird Kritik an den erhobenen schulischen Dimensionen und der Informationsrückmeldung geäußert. Zum einen wird gefragt, ob die Schulen überhaupt die Informationen erhalten, die sie benötigen. Die Inspektion wird

---

<sup>12</sup> Weitere Kritikpunkte können in den sog. Nebenwirkungen oder den nicht-intendierten Effekten neuer Steuerung im Schulsystem gesehen werden. Die diesbezüglichen Forschungserkenntnisse stammen vorrangig aus angloamerikanischen Ländern. Eine Systematik nicht-intendierter Effekte neuer Steuerung im Schulsystem sowie theoretische Erklärungsmodelle werden in Bellmann und Weiß (2009) skizziert.

nicht, so wie es in den Niederlanden praktiziert wird, flexibel und in Abhängigkeit von schulinternen Evaluationsergebnissen durchgeführt. Das Verfahren verläuft in Deutschland noch stark standardisiert und nivellierend. Kritiker sehen darin eine Unterwanderung der Individualität und Profilbildung der Schulen (vgl. Middendorf 2010). Anstatt alle Qualitätsaspekte zu bewerten fordern sie, nur bestimmte Aspekte einer intensiveren Analyse zu unterziehen. Zum anderen lässt sich fragen, inwiefern die Evaluationsergebnisse und Inspektionsberichte von den Schulen als hilfreich empfunden werden. Die Frage zielt auf die Art der Darstellung und den Umfang der Berichte: Sind die Informationen verständlich? Liegt eine Informationsüberflutung vor? Bleiben Fragen offen?

Schließlich zeigt sich ein weiterer Kritikpunkt der Schulinspektion aus der Perspektive der Lehrkräfte: Sie erleben Stress und Angst, sehen ausgelöst durch die Schulinspektion eine hohe Arbeitsbelastung auf sich zukommen und bereiten sich z. T. exzessiv auf den Schulbesuch vor (vgl. Perryman 2007). Zur Verdeutlichung, welche Reaktionen die Qualitätsanalyse NRW bei Lehrkräften auslöst, werden einige Zitate aus einem Lehrerforum exemplarisch angeführt.<sup>13</sup>

Bei wem war die QA [Qualitätsanalyse] im Unterricht und wie ist das?

Aber man weiß doch nicht, an welchem Tag die kommen, oder? Für mich wäre das der Albtraum, ich hab diese Vorführstunden im Ref [Referendariat] gehasst (klar, wer nicht) und war nervös ohne Ende.

Keine Sorge. Diese Besuche sind harmloser als man denkt. Die wollen nix haben, die reden eigentlich kaum mit dir und verhalten sich sehr zurückhaltend.

Unsere SL [Schulleitung] nervt uns schon seit 2 Jahren mit der drohenden QA.

Bei uns ist dies eine Art Schreckgespenst. Über allem, was wir in Hinblick auf Schulentwicklung etc. machen, hängt dieses Thema.

QA (oder wie es bei uns heißt: Kuharsch...).

---

<sup>13</sup> Zitate entnommen aus folgenden Onlineforen: <http://www.lehrerforen.de/index.php?page=Thread&threadID=28213> und <http://85.214.97.239/thread.php?postid=208174> [Zugriff am 27.10.2011].

Die exemplarischen Lehrer-Zitate zeigen verschiedene Reaktionen in Bezug auf die Schulinspektion: Zum einen besteht z. T. noch Informationsbedarf zum Vorgang der Inspektion, selbst noch im Jahr 2011, zum anderen werden beschwichtigende Kommentare, aber auch Angst, Skepsis, offensichtliche Ablehnung und Verhöhnung der Schulinspektion deutlich. Das in den Zitaten erkennbare „routinierte Abweisen und Umgehen extern (oder auch intern) erzeugter Hinweise auf Entwicklungsbedarf“ (Terhart 2002, S. 104) überrascht vor dem Hintergrund, dass gerade die Leistungsüberprüfung und -bewertung zu den alltäglichen professionellen Aufgaben der Lehrkräfte gehört. Entgegen der Annahme, dass in der Institution Schule die Evaluation der Schulqualität zu ihrer Verbesserung akzeptiert ist, bezieht sich die Akzeptanz der Lehrkräfte vorrangig auf die Evaluation von Schülerleistungen und auf „traditionelle Formate von Rückmeldungen in der Schule und im Unterricht“ (Merkens 2007, S. 217), während der Leistungsstand der eigenen Arbeit und der Institution Schule z.T. davon ausgenommen werden. Evaluationen im schulischen Kontext, darunter auch die Schulinspektion, sind immer mit Veränderungen verbunden und können Widerstand von Seiten der Lehrkräfte hervorrufen. Widerstand wird

oft als die irrationale Abwehr logisch sinnvoller Veränderungen angesehen, somit als Störung und einseitige Opposition, die sich in Nörgeleien, Aufsässigkeit, betretenem Schweigen, Kündigungen, Krankheit und indirekter Sabotage äußern kann (Steins 2009, S. 186).

Anstatt als Hilfsmittel, können Evaluationen von Lehrkräften als Gängelung wahrgenommen werden (vgl. ebd., S. 185). Dies führt dazu, dass der gesamte Evaluationsprozess nicht akzeptiert wird, die Ergebnisse angezweifelt werden und das Instrument letztlich wirkungslos bleibt (vgl. Lambrecht/Kotthoff/Maag Merki 2008, S. 290). Widerstand tritt bei Interessengegensätzen zwischen verschiedenen Personen oder -gruppen auf und kann vielfältige Gründe haben. Bspw. entsteht er durch die Wahrnehmung eines Kontrollverlustes über die eigene Arbeit oder durch die Abwehr von Angst, die durch die Befürchtung einer Überforderung eintritt (vgl. Berkemeyer/Bos 2009, S. 530). Steins (2009, S. 187-192) führt drei Theorien an, die den Widerstand von Lehrkräften gegen Evaluationen erklären: Reaktanztheorie, Theorie der kognizierten Kontrolle sowie Selbstaufmerksamkeitstheorie.

Dem Widerstand im Zusammenhang mit der Schulinspektion kann durch folgende Faktoren entgegengewirkt werden (vgl. Steins 2009; Böttcher/Kotthoff 2007, S. 225): Das Verfahren (Ablauf, Zeitraum, Konsequenzen) und sein Nutzen sind zu begründen, die Intentionen einer Evaluation sind glaubwürdig zu vermitteln, die Konsequenzen von Evaluationsergebnissen sollten als fair und unterstützend wahrgenommen werden und die Evaluation sollte möglichst langsam implementiert werden. Zudem erweist es sich als vorteilhaft, durch Partizipationsmöglichkeiten der Beteiligten ihre Wahrnehmung und Bewahrung der Kontrolle zu sichern. Schließlich sollte die Evaluation selbst zum Gegenstand einer Evaluation werden.

Resümierend zeigt sich, dass die Schulinspektion einige wichtige Kriterien bereits berücksichtigt (Begründung des Verfahrens, Evaluation der Schulinspektion), andere allerdings noch nicht verwirklicht werden, z. B. die Glaubwürdigkeit der Intention für die schulischen Akteure herauszustellen. Zu erwarten ist, dass die Schulinspektion analog zu allen neuen Verfahren und Denkweisen verschiedene Phasen von der Verhöhnung über die Ablehnung bis hin zur Akzeptanz durchläuft.

### **3.3. Die Qualitätsanalyse an Schulen in Nordrhein-Westfalen**

Die Anfänge der Schulinspektion in NRW sind bereits im Jahr 2002 in einem Kooperationsprojekt der Schulaufsicht mit Experten der niederländischen Schulinspektion erkennbar, welches die Übertragbarkeit des niederländischen Inspektionsverfahrens auf NRW prüfen sollte (vgl. Thürmann 2007, S. 189). Im Schuljahr 2005/06 erfolgte die Pilotierung der Schulinspektion in NRW mit 50 freiwilligen Schulen (vgl. ebd., S. 179). Seit dem Schuljahr 2006/07 wird die *Qualitätsanalyse* auf der Grundlage des Schulgesetzes vom 27.06.2006 (SchulG, § 3 Abs. 4 und § 86 Abs. 4 und 5) flächendeckend durchgeführt.

#### **3.3.1. Zielsetzungen und normativer Rahmen**

Die rechtlichen Rahmenbedingungen der Qualitätsanalyse NRW sind in einer Rechtsverordnung (QA-VO vom 27.04.2007) geregelt. Darin heißt es in § 1 Abs. 1 – Ziele und Aufgaben:

Qualitätsanalyse dient dem Ziel, die Qualität von Schulen zu sichern und nachhaltige Impulse für deren Weiterentwicklung zu geben. Dazu liefert sie detaillierte Kenntnisse über die Qualität der einzelnen Schulen und darüber hinaus über die Qualität des nordrhein-westfälischen Schulsystems insgesamt. [...] Die Ergebnisse sollen für gezielte Maßnahmen der Qualitätsverbesserung in den einzelnen Schulen sowie für entsprechende Unterstützungsleistungen der Schulaufsichtsbehörden und Steuerungsmaßnahmen des Ministeriums genutzt werden.

Die im ersten Satz der Verordnung angeführten Impulse sollen insbesondere von dem Qualitätsbericht ausgehen, der die zentralen Ergebnisse der Qualitätsanalyse einer Schule darstellt. Es ist vorgesehen, dass aus der Analyse des Berichts und der in einer Zielvereinbarung mit der Schulaufsicht fixierten Maßnahmenplanung Handlungen im Sinne der Schul- und Unterrichtsentwicklung umgesetzt werden. Die Ergebnisse sollen zudem für Unterstützungsleistungen der Schulaufsicht und für Steuerungsmaßnahmen des Ministeriums genutzt werden. Die in der Verordnung angesprochenen detaillierten Kenntnisse über

die Qualität der einzelnen Schulen werden auf der Grundlage eines normativen Rahmens erhoben. Das Qualitätstableau NRW orientiert sich am niederländischen Vorbild und wurde in Kooperation mit der niederländischen Schulinspektion erarbeitet. Es weist eine hohe inhaltliche Übereinstimmung zu den Qualitätsrahmen anderer Bundesländer auf, z. B. Niedersachsen, Berlin und Brandenburg. Bei der Entwicklung wurden Schulpraktiker, Schulaufsichtsmitarbeiter, Mitarbeiter des Schulministeriums aus den Bereichen Grundsatzfragen sowie Bildungsforschung, externe Experten und Wissenschaftler einbezogen (vgl. Wey 2008, S. 72f.). Das Ministerium für Schule und Weiterbildung (MSW) NRW bezeichnet es als „wissenschaftlich abgesichert“ (MSW NRW 2009a, S. 7). Das Qualitätstableau NRW dient als Bezugsrahmen für die Qualitätsanalyse, jedoch auch als Verständigungsgrundlage für die schulinterne Evaluation. Es umfasst folgende sechs Qualitätsbereiche ([QB]; vgl. MSW NRW 2008a, S. 184-194):

- QB 1 – Ergebnisse der Schule
- QB 2 – Lernen und Lehren – Unterricht
- QB 3 – Schulkultur
- QB 4 – Führung und Schulmanagement
- QB 5 – Professionalität der Lehrkräfte
- QB 6 – Ziele und Strategien der Qualitätsentwicklung

Die sechs Qualitätsbereiche werden in 28 Qualitätsaspekte und 153 Qualitätskriterien untergliedert. Dabei werden allein sieben Qualitätsaspekte dem Qualitätsbereich 2 zugeordnet, der sich von den anderen Qualitätsbereichen dadurch abhebt, dass drei Qualitätsaspekte auf der Grundlage der 41 Indikatoren des Unterrichtsbeobachtungsbogens bewertet werden.

Zeitgleich mit der Entstehung dieser Arbeit wurde das Qualitätstableau überarbeitet und modifiziert. Im Zuge der „Neuausrichtung der Qualitätsanalyse“ im Jahr 2013 wurde es um zwei Qualitätsaspekte, zwei Qualitätskriterien und fünf Unterrichtsindikatoren reduziert. Zudem wurde eine Unterscheidung in obligatorische und fakultative Prüfkriterien vorgenommen. Dabei wurde die Anzahl der obligatorischen Prüfkriterien von 153 auf ca. 50 Prüfkriterien reduziert (vgl. MSW NRW 2015a).

### 3.3.2. Institutionalisation und Arbeitsstand

Im MSW NRW wurde eigens für die Qualitätsanalyse ein Referat – Referat 414 – Qualitätsanalyse an Schulen, Qualitätstableau und IT-Verfahren der Qualitätsanalyse, Schulung der Qualitätsteams, Zielvereinbarungen nach Qualitätsanalyse – instituiert. Auf der Ebene der fünf nordrhein-westfälischen Regierungsbezirke Arnsberg, Detmold, Düsseldorf, Köln und Münster sind in den Bezirksregierungen in den Abteilungen 4 (Schule) die Dezernate 4Q – Qualitätsanalyse an Schulen – angesiedelt. Jedes dieser Dezernate 4Q verfügt über einen Hauptdezernenten und ein Team von Qualitätsprüfern, pädagogischen Mitarbeitern sowie Verwaltungsmitarbeitern. Die Dezernate sind sowohl personell als auch funktionell von der herkömmlichen Schulaufsicht abgekoppelt.

Die personelle Ausstattung mit Qualitätsprüfern ist seit 2006 kontinuierlich gestiegen. Zunächst wurden landesweit lediglich 40 Stellen für Qualitätsprüfer ausgewiesen; im Jahr 2008 waren es insgesamt 65 Stellen, ein Jahr später bereits 90 Stellen. Von diesen 90 Stellen waren 50 Stellen auf dem Wege der Abordnung entnommen. Im Endausbau sollen nach Beschluss der Landesregierung 105 Qualitätsprüfer für die Qualitätsanalyse tätig sein. Ende des Jahres 2009 waren 75 Stellen besetzt und für die übrigen Stellen liefen Ausschreibungs- und Besetzungsverfahren (vgl. Landesregierung<sup>14</sup> NRW 2010, S. 118). Insgesamt rekrutierte sich die erste Besetzung der Qualitätsanalyse aus der herkömmlichen Schulaufsicht, während in den folgenden Einstellungsverfahren auch Schul- und Seminarleitungen berücksichtigt wurden (vgl. Döbert u.a. 2008, S. 104). Die Auswahl der Qualitätsprüfer erfolgt grundsätzlich nach dem Prinzip der Bestenauslese im Sinne des Beamtenstatusgesetzes. Dabei werden sowohl fachliche Befähigungen als auch persönliche Anforderungen berücksichtigt. Bei einer positiven Bewerbungsentscheidung durchlaufen die Qualitätsprüfer eine sechsmonatige Vorbereitungsphase mit theoretischen und praktischen Ausbildungsabschnitten (vgl. Krufft 2008, S. 117-119). Sie dient dem sicheren Umgang mit den Instrumenten der Qualitätsanalyse, einheitlichen Bewertungen nach standardisierten Kriterien, der Erzeugung gesicherter Ergebnisse sowie der Professionalisierung des Auftretens (vgl. MSW NRW 2009a, S. 7; Koch-Riotte 2008). Die Qualifizierung wird mit einer selbständig durchgeführ-

---

<sup>14</sup> Bis Juli 2010 wurde die Landesregierung von der CDU/FDP-Koalition geführt.

ten Qualitätsanalyse abgeschlossen. Erst nach dem erfolgreichen Abschluss der Qualifizierungsphase wird über den Einsatz als Qualitätsprüfer entschieden. Grundsätzlich fungieren die Qualitätsprüfer nicht als Berater der Schulen, vielmehr gehört die sachliche und neutrale Bewertung der Schulqualität zu ihren prioritären Aufgabenfeldern (vgl. Kruft 2008, S. 120). Ein Qualitätsteam besteht aus mindestens zwei Qualitätsprüfern und kann bei Qualitätsanalysen großer Schulen auch aus drei oder vier Qualitätsprüfern zusammengesetzt sein. Dabei verfügt der federführende Qualitätsprüfer über die Lehramtsbefähigung der besuchten Schulform. Die Teams arbeiten in einer flexiblen Konstellation, d. h. es existiert ein rotierendes Verfahren beim Einsatz von Qualitätsprüfern (vgl. Amonat 2008, S. 54).

Im Jahr 2006 wurde vom MSW das Ziel vorgegeben, im Jahr 2011 die Basismessung aller Schulen erreicht zu haben. Aus heutiger Sicht steht fest, dass das Vorhaben, alle 6.800 nordrhein-westfälischen Schulen innerhalb von fünf Jahren zu inspizieren, angesichts der personellen Ausstattung sehr ambitioniert war und nicht erfüllt wurde. Zum Stand der von der Qualitätsanalyse besuchten Schulen liegen Angaben aus dem Jahresbericht der Qualitätsanalyse NRW (vgl. MSW NRW 2009a) und aus der Antwort der Landesregierung NRW (2010, S. 174) auf die Große Anfrage 40<sup>15</sup> vor (s. Tab. 4)

**Tab. 4: Übersicht der Qualitätsanalysen nach Schulform und Bezirksregierung bis zum Ende des Schuljahres 2008/09**

Bezirks- regierung	Schulform							$\Sigma$
	GS	HS	RS	GY	GE	FS	BK	
<b>Arnsberg</b>	78	32	31	20	28	35	28	<b>252</b>
<b>Detmold</b>	73	27	30	23	6	26	10	<b>195</b>
<b>Düsseldorf</b>	66	31	29	25	15	29	24	<b>219</b>
<b>Köln</b>	71	20	4	28	11	30	23	<b>187</b>
<b>Münster</b>	45	31	26	8	9	21	20	<b>160</b>
<b><math>\Sigma</math></b>	<b>333</b>	<b>141</b>	<b>120</b>	<b>104</b>	<b>69</b>	<b>141</b>	<b>105</b>	<b>1.013</b>

GS = Grundschule; HS = Hauptschule; RS = Realschule; GY = Gymnasium;  
GE = Gesamtschule; FS = Förderschule; BK = Berufskolleg;  $\Sigma$  = insgesamt

<sup>15</sup> Die Große Anfrage 40 wurde 2009 von der SPD gestellt.

Im Schuljahr 2006/07 sind 227 Qualitätsanalysen und im Schuljahr 2007/08 325 weitere Qualitätsanalysen durchgeführt worden (vgl. MSW 2009a, S. 31). Bis zum Schuljahr 2008/09 sind etwa 1.000 Schulen und nahezu 30.000 Unterrichtssequenzen besucht und mehr als 8,5 Millionen Einzeldaten erhoben worden – dies mit z. T. beträchtlichen Differenzen in Abhängigkeit der Schulformen und Bezirksregierungen (s. Tab. 4; vgl. Landesregierung NRW 2010, S. 174). Von den 1.013 Schulen wurden elf Nachanalysen<sup>16</sup> durchgeführt, davon jeweils vier in der Schulform Grundschule und Förderschule (vgl. ebd., S. 171). Insgesamt werden der massive Rückstand und die große Kluft zwischen geplanten und tatsächlich durchgeführten Qualitätsanalysen mit der angespannten Haushaltslage, personellen Ausstattung und optimierten Qualifizierung der Qualitätsprüfer begründet (vgl. ebd., S. 175).

### 3.3.3. Schulauswahl und (zeitlicher) Verlauf

Die Auswahl der öffentlichen Schulen erfolgt nach dem Aspekt der Regionalität (vgl. Döbert u.a. 2008, S. 105). Für die Analysen werden bestimmte Regionen – Kreise und Städte – zufällig ausgewählt und alle Schulen eines regionalen Kontextes werden in einer zeitnahen Sequenz besucht. Durch dieses Verfahren sollen für die Schulen Vor- und Nachteile in einer lokalen Wettbewerbssituation vermieden werden (vgl. Müller 2010, S. 298). Im Regierungsbezirk Münster wurde bspw. im Jahr 2006 mit den Qualitätsanalysen im Kreis Warendorf und der Stadt Recklinghausen begonnen (vgl. Ruhwinkel/Sticka 2008).

Während die Qualitätsanalyse für die öffentlichen Schulen obligatorisch ist, hat sie für Schulen in privater Trägerschaft Angebotscharakter. Dazu wird eine Kooperationsvereinbarung zwischen dem MSW und den Schulen in privater Trägerschaft geschlossen, welche eine Ergänzung des für öffentliche Schulen geltenden Qualitätstableaus um einen siebten Qualitätsbereich gemäß der Eigenprägung solcher Schulen enthält (vgl. Middendorf 2010; von Moritz 2011). Bei katholischen Schulen lautet dieser siebte Qualitätsbereich „Eigenprägung katholischer Schulen in freier Trägerschaft“. Evangelische Schulen haben als weiteren Qualitätsbereich ein „Evangelisch diakonisches Profil“ definiert.

---

<sup>16</sup> Folgende Referenzwerte schwach bewerteter Schulen können exemplarisch aus drei Bundesländern berichtet werden (vgl. Sommer/Stöhr/Thomas 2010, S. 210): Berlin 4 %; Brandenburg zwischen 5 % und 10 %; Niedersachsen 6,2 %.

Der zeitliche Ablauf von der Ankündigung der Qualitätsanalyse bis zum Erhalt des schriftlichen Qualitätsberichts umfasst ca. fünf Monate. Dabei verläuft die Qualitätsanalyse in drei Phasen von der Vorbereitung über die Durchführung bis zur Phase nach dem Schulbesuch. Im Einzelnen nimmt eine Qualitätsanalyse folgenden Verlauf: Etwa drei Monate vor dem Schulbesuch erfolgt die Ankündigung der Qualitätsanalyse. Sechs Wochen vor dem Schulbesuch sind die Schulen angehalten, der Bezirksregierung das Schulportfolio sowie sämtliche Daten, Konzepte und Dokumente zuzuschicken. Etwa vier Wochen vor dem Schulbesuch findet im Rahmen einer Vorinformation ein erstes Zusammentreffen statt, jedoch noch nicht der eigentliche Schulbesuch. Dabei erfolgt ein Planungsgespräch mit der Schulleitung, ein Schulrundgang, ein Interview mit dem Schulträger sowie eine Informationsveranstaltung für die Schulöffentlichkeit. Der drei- bis viertägige Schulbesuch stellt das Kernstück der Qualitätsanalyse dar. Es werden Unterrichtsbeobachtungen und Interviews mit den schulischen Akteuren durchgeführt. Am letzten Schulbesuchstag erfolgt eine erste mündliche Rückmeldung in Form einer etwa einstündigen PowerPoint-Präsentation an das Lehrerkollegium und die Schulleitung. Etwa einen Monat nach dem Schulbesuch wird der Schulleitung und dem Schulträger ein Berichtsentwurf gestellt, zu dem die Adressaten Stellung beziehen können. Etwa zwei Monate nach dem Schulbesuch erhalten die Schule, der Schulträger und die schulformbezogene Schulaufsicht den abschließenden Qualitätsbericht.

### **3.3.4. Verfahren und Instrumente**

Im Rahmen der Qualitätsanalyse kommen folgende Instrumente zum Einsatz: Dokumentenanalyse, Schulrundgang, Unterrichtsbeobachtungen und Leitfadenterviews. Im Vorfeld des Schulbesuchs werden von den Qualitätsteams und pädagogischen Mitarbeitern die schulischen Dokumente gesichtet. Die Dokumentenanalyse umfasst neben dem Schulprogramm das schulformspezifische Schulportfolio, welches die schulischen Leistungs- und Entwicklungsdaten, den pädagogischen Arbeitsstand sowie schulische Rahmendaten enthält (vgl. MSW NRW 2008b; Peuker 2008). Beim Schulrundgang erhält das Qualitätsteam einen über die Dokumente hinausgehenden ‚realen‘ Eindruck der Schule. Dabei werden neben ausgewählten Unterrichts- und Verwaltungsräumen die Sporthalle, das Schulgelände etc. anhand eines Leitfadens begutachtet. Insgesamt soll

dadurch die materielle Ausstattung der Schulen nach aktuellen didaktisch-methodischen Standards geprüft werden.

Die Unterrichtsbeobachtungen erfolgen in NRW anhand eines standardisierten, schulformübergreifenden Unterrichtsbeobachtungsbogens und stehen im Zentrum der Qualitätsanalyse (s. MSW NRW 2008c). In den Unterrichtsbesuchen sollen mindestens 50 % der Lehrkräfte berücksichtigt werden, bzw. bei kleineren Schulen mindestens 20 Unterrichtssequenzen (vgl. Lennartz 2010). Jede Unterrichtssequenz hat eine Dauer von 20 Minuten und erfolgt in der ersten oder letzten Unterrichtshälfte. Der Besuchsplan wird vorab von den Verwaltungsmitarbeitern unter Berücksichtigung unterschiedlicher Fächer, Jahrgänge und schulformspezifischer Strukturen erstellt und den Schulen vorab nicht bekanntgegeben. Mittlerweile erfolgen die Unterrichtsbewertungen durch eine direkte Eingabe der entsprechenden Bewertungsstufen in einen Tablet-PC, welcher eine unmittelbare elektronische Verarbeitung der Eingaben sichert. Die Bewertungen der einzelnen Qualitätsindikatoren des Unterrichtsbeobachtungsbogens sowie der Qualitätskriterien werden anhand einer vierstufigen Bewertungsskala (++ „beispielhaft erfüllt“; + „eher erfüllt“; - „eher nicht erfüllt“; -- „überhaupt nicht erfüllt“) vorgenommen.<sup>17</sup> Daneben kann auch eine Nullwertung vorgenommen werden, wenn zu einer Kategorie keine Beobachtung gemacht werden kann und ihr Fehlen nicht zu einer negativen Bewertung führen soll (vgl. Dorn 2008, S. 98).

Schließlich finden während des Schulbesuchs leitfadengestützte Interviews mit der Schulleitung, den Lehrkräften, ausgewählten Schülern sowie Eltern statt. Zudem werden das nichtlehrende Personal und ggf. duale Partner befragt. Die Interviews finden jeweils getrennt statt, dauern je nach Befragungsgruppe maximal 90 Minuten und orientieren sich inhaltlich am Qualitätstableau. Die Schule übernimmt dabei die Aufgabe, die Interviewgruppen nach repräsentativen Kriterien zusammenzustellen.

Für jede der Datenquellen wird eine entsprechende Bewertung der Qualitätsaspekte auf einer vierstufigen Skala vorgenommen (vgl. MSW NRW 2009a, S. 10; MSW NRW 2008d, S. 217):

---

<sup>17</sup> Seit Beginn des Schuljahres 2013/14 werden die Indikatoren des neuen, modifizierten Unterrichtsbeobachtungsbogens mit den beiden Bewertungsstufen „trifft zu“ und „trifft nicht zu“ bewertet.

- *Bewertungsstufe 4 – vorbildlich*: Die Schule erfüllt nahezu alle Kriterien dieses Qualitätsaspekts optimal oder gut.
- *Bewertungsstufe 3 – eher stark als schwach*: Die Schule weist bei diesem Qualitätsaspekt mehr Stärken als Schwächen auf. Die Schule kann die Qualität einiger Kriterien noch weiter verbessern; die wichtigsten Kriterien werden erfüllt.
- *Bewertungsstufe 2 – eher schwach als stark*: Die Schule weist bei diesem Qualitätsaspekt mehr Schwächen als Stärken auf; die wesentlichen Kriterien sind noch verbesserungsfähig.
- *Bewertungsstufe 1 – erheblich entwicklungsbedürftig*: Bei allen Kriterien des Qualitätsaspekts sind Verbesserungen erforderlich.

Jeder Bewertungsstufe wird eine spezifische Farbe zugeordnet (s. o.), welche den Schulen eine rasche Orientierung bezüglich ihrer bewerteten Qualität ermöglicht. Das Farbschema hat einen hohen Wiedererkennungswert, verleitet aber andererseits dazu, die Komplexität der Schulbewertung auf eine Farbe zu reduzieren. Die zusammengetragenen Einzelbewertungen werden in ELBA (Elektronischer Berichtsanhang), einem Excel-gestützten Programm, verrechnet. Es wird ein Bewertungsvorschlag entwickelt, der im Ausnahmefall von den Qualitätsprüfern geändert werden kann (vgl. Dorn 2008). Dem Programm ELBA liegen bestimmte Bewertungsvorschriften zugrunde. Bspw. wird ein Qualitätsaspekt mit der Stufe 3 bewertet, wenn mehr als 50 % der bewerteten Kriterien mit + oder ++ bewertet sind (vgl. MSW NRW 2009a, S. 58-61; Dorn 2008, S. 97). Es liegen jedoch auch sog. Sperrkriterien zugrunde, die im Vergleich zu anderen Kriterien höher gewichtet werden (vgl. Wey 2008, S. 71).

### 3.3.5. Rückmeldung und Ergebnisnutzung

Am letzten Schulbesuchstag erhält die Schule eine datengestützte mündliche Rückmeldung zu den erhobenen Ergebnissen. Der einige Wochen später folgende schriftliche Qualitätsbericht stellt eine reine Beschreibung der Ergebnisse dar und enthält keine konkreten Handlungsempfehlungen. In den ersten Jahren der Qualitätsanalyse umfasste er ca. 50 bis 60 Seiten, mittlerweile weist er jedoch einen reduzierten Umfang auf. Er wird der Schulleitung, Schulkonferenz, Lehrerkonferenz, Schulpflegschaft, dem Schülerrat, Schulträger sowie der

Schulaufsicht zur Verfügung gestellt. Neben umfangreichen Datenanteilen und -anhängen enthält er Textelemente, welche entlang der Qualitätsaspekte des Qualitätstableaus gegliedert sind. Hervorzuheben ist, dass die Schule mit dem Bericht – und grundsätzlich mit der Qualitätsanalyse – keine Globalbewertung ihrer Schulqualität anhand einer einzelnen Bewertungsstufe erhält. Der Bericht enthält die Bewertungen der 28 Qualitätsaspekte, zeichnet ein differenziertes Bild der Schule und ermöglicht durch seinen Impulscharakter den Schulen, möglichst konkrete Handlungsfelder zu identifizieren, damit Schul- und Unterrichtsentwicklungsmaßnahmen abgeleitet und aktiviert werden können.

Im Anschluss an die Qualitätsanalyse erfolgt eine ministeriell administrierte Evaluation des Verfahrens, indem die Schulen einen standardisierten Bogen zur Vorbereitung und Durchführung des Schulbesuchs sowie zum Qualitätsbericht beantworten (vgl. MSW NRW 2008e). Mit der Berichtsübergabe wird die Verantwortung für die Nutzung der Ergebnisse zum großen Teil an die Schulen übergeben. Um den Prozess nach der Qualitätsanalyse zu steuern, wurde vom MSW ein institutioneller Zielvereinbarungsprozess zwischen den Schulen und der Schulaufsicht verordnet. Bei der Umsetzung der in den Zielvereinbarungen festgehaltenen Maßnahmen können die Schulen neben der Beratungskompetenz der Schulaufsicht auf Fortbildungsmaßnahmen der 53 neu gegründeten Kompetenzteams<sup>18</sup> zurückgreifen. In einem Bilanzgespräch wird schließlich die Zielerreichung überprüft. Der Zielvereinbarungsprozess wird an dieser Stelle nicht weiter beleuchtet, sondern in Kapitel 4.2 wieder aufgegriffen.

### **3.3.6. Zentrale Befunde der bisherigen Qualitätsanalysen**

Das MSW NRW hat Ende des Jahres 2009 erstmals einen Jahresbericht zu den Ergebnissen, zum Verfahren und den Instrumenten der Qualitätsanalyse veröffentlicht (vgl. MSW NRW 2009a). Die Befunde des Jahresberichts stammen aus dem Schuljahr 2007/08. Im Hinblick auf die zentralen Befunde der Quali-

---

<sup>18</sup> In NRW wurden in jedem der 53 Schulämter Kompetenzteams für alle Schulformen eingerichtet. Sie sind für Fortbildungen an Schulen zuständig. Dabei legt das MSW abwechselnde Fortbildungsschwerpunkte, z. B. zu den Themen Unterrichtsentwicklung und individuelle Förderung, fest. Die Kompetenzteams werden unter der Mitarbeit von Moderatoren von Schulaufsichtsbeamten geleitet, organisiert und betreut.

tätsanalyse zeichnet das MSW NRW (2009a) auf der Grundlage der ersten 325 Schulbesuche folgendes (nicht-repräsentative) Bild:

Zu den besonderen Stärken der besuchten Schulen zählen dabei das Schulklima, die Lernatmosphäre und die Zufriedenheit aller Beteiligten mit der eigenen Schule insgesamt, die Schülerberatung und -unterstützung sowie die außerschulische Kooperation. Ebenfalls sehr positiv bewertet wird der Qualitätsaspekt ‚Führungsverantwortung der Schulleitung‘. Handlungsfelder bestehen hingegen in einzelnen Aspekten der Qualitätsbereiche 2 und 6, dort z. B. bei der internen Evaluation. Diese Ergebnisse decken sich weitgehend mit denen anderer Länder, die vergleichbare Instrumente für die externe Evaluation verwenden (S. 16).

Dabei kann eine schulformübergreifende Tendenz beobachtet werden: Relativ homogene Ergebnisse liegen für die Qualitätsaspekte 1.5 *Zufriedenheit der Beteiligten*, 2.3 *Unterricht – fachliche und didaktische Gestaltung* und 2.4 *Unterricht – Unterstützung eines aktiven Lernprozesses* vor. Hingegen zeigt sich eine Streuung zwischen einer und einer halben Bewertungsstufe bei den Aspekten 2.6 *Individuelle Förderung und Unterstützung*, 2.7 *Schülerbetreuung*, 3.5 *Außerschulische Kooperation* und 6.2 *Schulinterne Evaluation*. Insgesamt liegen die Stärken der Schulen eher im Bereich sog. weicher Kriterien der Schulqualität, wie z. B. dem Schulklima und der Zufriedenheit der schulischen Akteure, und *nicht* im Unterricht – dem Kernbereich der Schulen. Neben den schwachen Bewertungen des Unterrichts liegen weitere im Qualitätsbereich 6 *Ziele und Strategien der Qualitätsentwicklung* und werfen die Frage auf, wie gut die Schulen angesichts ihrer als eher schwach bewerteten Evaluationskultur in der Lage sind, die Schulinspektionsergebnisse in Maßnahmen der Schul- und Unterrichtsentwicklung zu überführen.

### 3.3.7. Zusammenfassung

Die Vorstellung der Qualitätsanalyse NRW macht deutlich, dass es sich dabei um kein finales Instrument handelt, sondern dass es sich in einem ständigen Prozess der Revision und Optimierung befindet, welche sich bspw. in der Reduzierung des Berichtsumfangs oder am überarbeiteten Unterrichtsbeobachtungsbogen zeigen. Die rasante Weiterentwicklung, Veränderung und Umstrukturierung der Schulinspektion zeigt sich nicht nur im Falle der Qualitätsanalyse NRW, sondern ist auch in anderen Bundesländern und europäischen Ländern (z. B. Niederlande und England, s. Kapitel 3.1) zu beobachten.

In NRW ist nach fünf Jahren Praxis die Mehrzahl der Schulen noch nicht inspiziert worden. Durch die lange Dauer der ersten Inspektionswelle ist anzuzweifeln, dass sich ein zweiter Inspektionsturnus nach gleichem Schema anschließt. Da sich NRW bereits bei der Implementierung der Qualitätsanalyse an den Niederlanden orientiert hat, ist auch bezüglich des zweiten Turnus eine Flexibilisierung des Instruments nach niederländischem Vorbild denkbar. Entsprechende Tendenzen sind in der Neuausrichtung der Qualitätsanalyse NRW erkennbar, welche beginnend ab dem Schuljahr 2013/14 in den darauffolgenden zwei Jahren vollständig umgesetzt werden soll (vgl. Jäger 2013). Kennzeichen dieser Neuausrichtung sind unter anderem ein verschlankter Ablauf, eine bürokratische Entlastung der Schulen, die Differenzierung von obligatorischen und fakultativen Prüfkriterien im Qualitätstableau, das zusätzliche Angebot eines Gesprächs zur Interpretation des Qualitätsberichts und der Ergebnisse sowie die stärkere Gewichtung der internen Evaluation (vgl. ebd.).

Die bislang erzielten Ergebnisse zeigen der Systemebene die charakteristischen Stärken und Schwächen der Schulen auf. Insbesondere die Ergebnisse zum Unterricht und zu den Zielen und Strategien der Qualitätsentwicklung erweisen sich z. T. als alarmierend und zeigen deutlichen Handlungsbedarf auf. Damit hat die Qualitätsanalyse eine ihrer Funktionen und Ziele erfüllt, indem sie sowohl den Schulen als auch der Systemebene detaillierte Kenntnisse über die Qualität von Schulen zur Verfügung gestellt hat. Inwiefern auch das weitere Ziel der Impulsgebung für die Schulentwicklung erfüllt wird, wird im Rahmen von wissenschaftlicher Begleitforschung zu eruieren sein. Der bisherige Forschungsstand zur Schulinspektion wird im nächsten Kapitel aufgearbeitet.

### 3.4. Forschungsstand zur Schulinspektion

Der internationale Forschungsstand zu den Wirkungen und zur Wirksamkeit von Schulinspektionen kann als „mixed picture“ (Ehren/Visscher 2006, S. 52) bezeichnet werden. Er ist durch heterogene und z. T. gegensätzliche Befunde gekennzeichnet, welche zu einem gewissen Anteil auch auf die unterschiedlichen Inspektionssysteme, Untersuchungsdesigns und -zeiträume zurückzuführen sind. Folglich können keine generalisierbaren und eindeutigen Schlüsse aus dem bisherigen Forschungsstand zu Schulinspektionen gezogen werden.

Zur Schulinspektion in Deutschland liegen unzählige und öffentlich zugängliche Auswertungen und Erfahrungsberichte der Ministerien, Qualitätsagenturen und Landesinstitute vor. Diese überwiegend in Form von Jahresberichten veröffentlichten Ergebnisse stellen zwar evaluationsbasiertes Wissen bereit, liefern jedoch keine wissenschaftlich konzipierten und an theoretische Modelle angelehnten Erkenntnisse. Hingegen ist die wissenschaftliche Erforschung der Schulinspektion in Deutschland bisher durch überschaubare Aktivitäten sowie einen schmalen Forschungsstand gekennzeichnet und liegt weit hinter dem internationalen Forschungsstand zurück. Dies ist insofern nachvollziehbar, als dass es sich dabei hierzulande um ein neues Steuerungsinstrument handelt und als dass sich eine „empirische Forschungslandschaft i. e. S.“ (Brauckmann u.a. 2010, S. 133) aktuell noch entfaltet und insgesamt die Erforschung der Schulinspektion und ihrer Wirkungen als dringendes Desiderat gefordert wird.

Aktuell sind in Deutschland verstärkte Forschungsbemühungen zum Thema Schulinspektion zu erkennen, so dass in den nächsten Jahren diesbezüglich neue und bedeutende Ergebnisse zu erwarten sind. Im Rahmen des Forschungsschwerpunkts *Steuerung im Bildungssystem* (SteBis) werden zwei vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Forschungsprojekte zum neuen Steuerungsinstrument *Schulinspektion* durchgeführt. Es handelt sich dabei unter anderem um das Verbundprojekt *Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene* (Laufzeit 2010 bis 2017). Hierbei werden in vier Bundesländern neben der Konzeptualisierung und Implementierung der jeweiligen Schulinspektion die intendierten und nicht-intendierten Wirkungen untersucht (vgl. BMBF 2010a). Zudem werden einzelne Schulen bei der Verarbeitung der

Inspektionsergebnisse wissenschaftlich begleitet und der Rezeptionsprozess auch aus Sicht der Administration und des schulischen Umfelds beleuchtet. Das zweite vom BMBF im Schwerpunkt SteBis geförderte Projekt *Externe Evaluation/Schulinspektion und Qualitätsentwicklung von Schulen* (Laufzeit 2010 bis 2013) untersucht den Einfluss der Schulinspektionsverfahren in zwei Bundesländern auf die Entwicklungsarbeit in Schulen (vgl. BMBF 2010b).

Auf internationaler Ebene wird aktuell das von der EU geförderte Forschungsprojekt *Impact of School Inspection on Teaching and Learning* (ISITL) mit einer Laufzeit von 2011 bis 2013 unter Beteiligung von neun Mitgliedsstaaten, allerdings ohne deutsche Partizipation, durchgeführt (vgl. Huber/Skedsmo/Durrer 2011). Zur Analyse des Einflusses von Schulinspektionen auf die Lehr- und Lernqualität sollen anhand eines Quasi-Experiments unterschiedliche europäische Schulinspektionen verglichen und daraus Empfehlungen für die Verbesserung der Lehre und Schülerleistungen abgeleitet werden.

Die o. g. aktuellen Forschungsprojekte verlaufen in Deutschland und Europa erstmals (bundes)länderübergreifend und mit einer längeren Laufzeit und unterstreichen damit die Bedeutung des Themas Schulinspektion in der Bildungsforschung. Besonders auffällig ist, dass alle drei gen. Großprojekte die von der Schulinspektion ausgehenden Wirkungen auf unterschiedliche Zielgrößen, bspw. die Schulentwicklung oder Verbesserung von Schülerleistungen, fokussieren. Die Wirkungsfrage der Schulinspektion bildet insgesamt ein für die Legitimierung des Instruments dringend gefordertes Kriterium.

### **3.4.1. Systematisierungen der Schulinspektionsforschung**

Die Schulinspektion bietet als Forschungsgegenstand äußerst facettenreiche Möglichkeiten und führt zu entsprechenden Forschungsaktivitäten. Mittlerweile liegen sogar Systematisierungsversuche der Schulinspektionsforschung vor. Auf die Implementierung des neuen Steuerungsinstruments hat die Forschung in Deutschland bislang in erster Linie mit vergleichenden Analysen der Schulinspektionskonzepte und -verfahren auf nationaler und internationaler Ebene reagiert (s. Kapitel 3.1 und 3.2; vgl. Döbert u.a. 2008; Döbrich u.a. 2008; Böttcher/Kotthoff 2007; Bos u.a. 2006; Kotthoff/Böttcher 2010; van Ackeren 2003). Diese Analysen, welche im engeren Sinne Forschungsdokumentationen darstellen, sind als erste Generation der wissenschaftlichen Auseinandersetzung

mit der Schulinspektion zu interpretieren. Sie nähern sich dem neuen Steuerungsinstrument auf deskriptivem Wege und versuchen, es in seiner Komplexität zu durchdringen und zu strukturieren. Die zweite und stetig wachsende Generation der deutschen Schulinspektionsforschung geht vorrangig empirisch vor und beleuchtet die Schulinspektion multiperspektivisch und -methodisch. Sie fokussiert individuelle Einschätzungen und Erfahrungen mit der Schulinspektion aus der Perspektive der Inspektion, Schulaufsicht und Schulen. Dabei stehen jeweils bundeslandbezogen einerseits die Instrumente der Schulinspektion auf dem Prüfstand (vgl. Pietsch/Tosana 2008; Gärtner/Pant 2011; Müller/Pietsch 2011; Tosana/Lambrech/Perels/Bardowiecks 2011) und andererseits werden die Effekte und Wirkungen von Schulinspektionen beleuchtet (vgl. z. B. Gärtner 2011; Sommer 2011).

Im Folgenden werden für die Darstellung des Forschungsstands die Systematiken von Böttcher und Keune (2010) sowie Husfeldt (2011a) zugrunde gelegt. Sie sind durch eine Dreiteilung charakterisiert und ähneln sich inhaltlich in starkem Maße bzw. sind im Falle des Forschungstyps III redundant (s. Tab. 5).

**Tab. 5: Systematiken der Schulinspektionsforschung**

Typ	Husfeldt (2011a): Studien zu Wirkungen und Wirksamkeit der Schulinspektion	Böttcher/Keune (2010): Studien zum Einfluss von Schulinspektion auf Schulentwicklung
I	Forschungen zu Einstellungen und Wirksamkeitserwartungen	Akzeptanz von Schulinspektion
II	Forschungen zu Reaktionen auf Inspektionen und Folgemaßnahmen	Einfluss von Schulinspektion auf Schul- und Unterrichtsentwicklung
III	Forschungen zu Schülerleistungen	Einfluss von Schulinspektion auf Schülerleistungen

Beide Typisierungen der internationalen Schulinspektionsforschung sind nicht als Metaevaluationen zu verstehen. Sie resultieren vielmehr aus einer überblickhaften Literaturrecherche für die Konzeption eigener Forschungsaktivitäten. Böttcher und Keune haben bei ihrer Recherche den Fokus auf die Entwicklungsfunktion der Schulinspektion gelegt und 31 nationale und internationale Zeitschriften der Jahrgänge 2005 bis 2009 zu Originalstudien zum Thema *Einfluss von Schulinspektion auf Schulentwicklung* gesichtet (vgl. Böttcher/Keune

2010, S. 152f.). Husfeldt (2011a, S. 271) führt „eine Ordnung der vorhandenen Forschungsansätze“ als Ziel der Systematisierung an.

Der *Forschungstyp I* zeichnet sich bei beiden Systematiken durch sog. weiche Effektivitätsaspekte aus, zu denen die Akzeptanz, Einstellungen und Wirksamkeitserwartungen der schulischen Akteure bezüglich Schulinspektion zu zählen sind. Der *Typ II* konzentriert sich auf die Frage der Nutzung der Ergebnisse und der Wirkungen von Schulinspektion auf die Schul- und Unterrichtsentwicklung. Die Befunde dieses Forschungstyps können eindeutige Erkenntnisse zur Wirksamkeit und die Wirkungen von Schulinspektion liefern. Dabei stehen oftmals die Anzahl und die Inhalte der Maßnahmen, welche als Reaktion auf die Ergebnismeldungen eingeleitet werden, im Mittelpunkt. Dem *Typ III* sind Studien zuzuordnen, welche sich mit den Wirkungen von Schulinspektion auf Schülerleistungen und die Leistungsentwicklung befassen.

Neben den o. g. Systematisierungsversuchen lassen sich zwei weitere finden: De Wolf und Janssens (2007) unterscheiden vier Typen von Studien zu den Wirkungen und zur Wirksamkeit von Schulinspektion:

1. Satisfaction of teachers and principals
2. Behaviour of teachers and principals
3. Improvement of school policy
4. Improvement of student results

Der erste und vierte Forschungstyp entsprechen den Typen I und III in den Systematiken von Husfeldt (2011a) sowie Böttcher und Keune (2010), die anderen zwei Typen unterteilen jeweils den Typ II (s. Tabelle 5) in das Verhalten des Schulpersonals sowie in die Entwicklung der Schulpolitik. Schließlich kann eine Systematisierung von Brauckmann u.a. (2010, S. 133) angeführt werden, welche die bislang vorliegenden Studien in zwei Kategorien unterteilt:

1. Systematische empirische Wirkungsforschung
2. Untersuchungen zur Messgüte der verwendeten Inspektionsinstrumente

Während die erste Kategorie eine Parallele zu den zuvor präsentierten Systematisierungsversuchen darstellt, stellt die Kategorie zur Messgüte der Inspektionsinstrumente, welche die Reliabilität, Objektivität und Validität des Inspektionsinstrumentariums umfasst, eine Erweiterung dar. Untersuchungen zur Messgüte der verwendeten Inspektionsinstrumente sind insbesondere in den letzten vier

Jahren verstärkt durchgeführt worden und geben Hinweise darauf, inwieweit die gemessenen Ergebnisse überhaupt als zuverlässig gelten können.

Die aufgezeigten Systematisierungsversuche der Schulinspektionsforschung sind inhaltlich größtenteils redundant und zeigen auch formal Übereinstimmungen, auch wenn sich die Systematiken im Einzelnen in der Gliederung und Zuordnung der Forschungsinhalte z. T. unterscheiden. Lediglich die von Brauckmann u.a. (2010) vorgenommene Typisierung stellt eine Ergänzung und Erweiterung der übrigen Forschungsschwerpunkte dar.

Die folgende Darstellung des Forschungsstands zu Schulinspektionen präsentiert keine Gesamtübersicht der Untersuchungen und Untersuchungsergebnisse zum Thema Schulinspektion. Aufgrund der jahrelangen internationalen Erfahrungen mit Schulinspektionen stellt sich der Forschungsstand äußerst komplex dar. Die hier gegebene Übersicht konzentriert sich daher insbesondere auf Untersuchungen, welche die deutsche Schulinspektion betreffen. Ein weiterer Schwerpunkt wird auf die Schulentwicklungsfunktion der Schulinspektion gelegt. Im Folgenden werden ausgewählte Studien entlang einer Systematisierung von Forschungstypen der Schulinspektion wiedergegeben (vgl. Böttcher und Keune 2010, Husfeldt 2011a, b; Brauckmann u.a. 2010).

### **3.4.2. Forschungstyp I: Einstellungen und Akzeptanz der Schulinspektion**

Ein Großteil der vorliegenden Forschungsbefunde zur deutschen Schulinspektion beruht auf subjektiven Einschätzungen zur Durchführungs- und Ergebnisakzeptanz aus der Perspektive von Schulleitungen und Lehrkräften (vgl. Gärtner/Hüsemann/Pant 2009; Gärtner/Wurster 2009; Sommer 2011). Kennzeichnend für solche Untersuchungen ist nach Brauckmann u.a. (2010, S. 134), dass sie bislang i. d. R. von Wissenschaftlern durchgeführt wurden, welche sich aus der Schulinspektion rekrutieren oder die Schulinspektion wissenschaftlich begleiten. Auch wenn es sich dabei um subjektive Einschätzungen handelt, die nicht generalisierbar sind, liefern sie dennoch für die erwartete Datennutzung und Maßnahmenumsetzung wichtige Anhaltspunkte (vgl. Visscher/Coe 2003, S. 322). Es liegen für die Schulinspektion Forschungsbefunde zu einem starken positiven Zusammenhang zwischen der Akzeptanz der erhobenen Ergebnisse und der Datennutzung vor (vgl. Pietsch 2011, S. 11f.). Umgekehrt lassen sich

keine positiven Wirkungen erwarten, wenn bei den schulischen Akteuren die Verfahrens- und Ergebnisakzeptanz der Schulinspektion ausbleibt.

Die z. T. kritischen Befunde zur von den schulischen Akteuren wahrgenommenen Akzeptanz der Schulinspektion aus England (vgl. Brimblecombe/Shaw/Ormston 1996) lassen sich nicht auf die deutsche Situation übertragen, da die Schulinspektion hierzulande keine solchen Sanktionen in Abhängigkeit der Inspektionsergebnisse nach sich zieht. Aus Deutschland kann insbesondere von drei Untersuchungen berichtet werden, welche die Einstellungen der schulischen Akteure und ihre Akzeptanz bezüglich der Schulinspektion untersuchen. Es handelt sich dabei um eine Auswertung niedersächsischer Evaluationsbögen im Anschluss an die Schulinspektion (vgl. Sommer 2011), um eine Schulleiterbefragung in Brandenburg (vgl. Gärtner u.a. 2009) sowie eine Befragung von Lehrkräften und Schulleitungen in Berlin (vgl. Gärtner/Wurster 2009). Darüber hinaus enthalten viele der von den Ministerien und Qualitätsagenturen verfassten Jahresberichte zur Schulinspektion Ergebnisse zur Zufriedenheit sowie Akzeptanz der Schulinspektion aus der Perspektive der schulischen Akteure. Hierzu kann exemplarisch angeführt werden, dass sich im Rahmen der Evaluationsbefragung der Qualitätsanalyse NRW (N = 396 Schulen) die überwiegende Mehrheit der Schulen (zwischen 82 % und 90 %) insgesamt zufrieden mit der Durchführung der Qualitätsanalyse zeigt (vgl. MSW NRW 2009a, S. 18). Allerdings wurde das Verhältnis von Gesamtaufwand und Ertrag der Qualitätsanalyse im Schuljahr 2007/08 zu 46 % als kritisch bewertet (vgl. ebd., S. 20).

Auf der Grundlage einer Befragung im Anschluss an die Schulinspektion in Niedersachsen wurden die Antworten von Schulleitungen und Lehrkräften (N = jeweils ca. 500) hinsichtlich der Akzeptanz von Schulinspektionen ausgewertet (vgl. Sommer 2011). Als Indikatoren der Inspektionsbeurteilung wurde einerseits die professionelle Durchführung der Inspektion, andererseits der Anregungsgehalt dieser für die Schulentwicklung herangezogen. Erstere wurde von den Schulleitungen überwiegend positiv beurteilt (nahezu 90 %). Die Lehrkräfte schätzten diesen Aspekt hingegen zurückhaltender ein (völlige Zustimmung unter 50 %). Die positivere Akzeptanzbekundung von Schulleitungen im Vergleich zu Lehrkräften zeigt sich zudem beim Anregungsgehalt der Inspektion für die Schulentwicklung, welche 70 % der Schulleitungen und ca. 55 % der Lehrkräfte als erfüllt einschätzen. In einem zweiten Schritt hat Som-

mer die einzelnen Items der Evaluationsfragebögen der Schulleitungen und Lehrkräfte faktorenanalytisch zu Skalen zusammengefasst. Die Dimension *Akzeptanz* der Schulinspektion wird dabei durch die vier Skalen *Verfahrensakzeptanz*, *Ergebnisakzeptanz*, *Validität* sowie *Bedeutung der Inspektion* operationalisiert (vgl. Sommer 2011, S. 159). Die Ergebnisse deuten darauf hin, dass die Akzeptanz sowohl bei den Schulleitungen als auch bei den Lehrkräften insgesamt hoch ausgeprägt ist. Alle Skalenmittelwerte liegen auf der vierstufigen Antwortskala<sup>19</sup> unterhalb des theoretischen Skalenmittelwerts von  $MW = 2,5$ . Dabei werden wiederholt die einzelnen Akzeptanzskalen von den Schulleitungen positiver bewertet als von den Lehrkräften. Die positivste Skalenausprägung ist bei den Schulleitungen bezüglich der Akzeptanz des Inspektionsverfahrens zu beobachten ( $MW = 1,3$ ), bei den Lehrkräften hingegen bei der Bedeutung der Inspektion ( $MW = 1,6$ ). Die Daten der niedersächsischen Evaluationsbefragung zeigen, dass die Verfahrensakzeptanz im Vergleich zur Ergebnisakzeptanz der Schulinspektion – übereinstimmend für die Schulleitungs- und Lehrkräftebefragung – mit höheren Skalenmittelwerten bewertet wird.

Ergänzend können folgende Erkenntnisse zur wahrgenommenen Diagnosegüte der Schulinspektion angeführt werden: Gärtner, Hüsemann und Pant (2009) haben 182 Schulleitungen der in den Schuljahren 2005/06 und 2006/07 in Brandenburg inspizierten Schulen zu den Wirkungen von Schulinspektionen schriftlich befragt. Als ein Teilaspekt wurde die wahrgenommene diagnostische Güte des Inspektionsverfahrens erhoben. Größtenteils sprechen die Schulleitungen den Inspektoren die professionelle Fähigkeit zu, die Schul- und Unterrichtsqualität gut – im Sinne von zuverlässig – einschätzen zu können. Diese generelle Bestätigung fällt im ersten Inspektionsjahr geringer aus als im zweiten Jahr und steht zudem im Zusammenhang mit der Bewertung der Schulen. Mit zunehmend positiven Bewertungen der Schulen wurde von den Schulleitungen auch die Diagnosefähigkeit der Inspektoren als zuverlässiger eingeschätzt. Dennoch ist der berechnete Einfluss der Inspektionsergebnisse auf den Akzeptanzgrad mit einer erklärten Varianz von 8 % als gering zu bewerten.

Eine etwas differenziertere Erhebung der Akzeptanz der Schulinspektion aus der Perspektive von Schulleitungen ( $N = 133$ ) und Lehrkräften ( $N = 396$ ) ist bei

---

<sup>19</sup> Vierstufige Antwortskala (vgl. Sommer 2011): 1 = stimme zu; 2 = stimme eher zu; 3 = stimme eher nicht zu; 4 = stimme gar nicht zu.

Gärtner und Wurster (2009) zu finden. Die Schulen, aus denen sich die entsprechenden Zielgruppen der schriftlichen Befragung rekrutieren, wurden in den Schuljahren 2006/07 und 2007/08 in Berlin inspiziert. Insgesamt zeigen die Befragungsbefunde, dass die Ergebnisakzeptanz der Schulinspektion sowohl aus der Perspektive der Schulleitungen als auch der Lehrkräfte überwiegend positiv bewertet wird. Das Item „Inwieweit stimmen die Inspektionsergebnisse mit Ihren eigenen Einschätzungen überein?“ beurteilen die Schulleitungen zu 19 % mit der Skalenausprägung<sup>20</sup> „vollständig“ und zu 71 % mit der Ausprägung „überwiegend“, Lehrkräfte dagegen zu 10 % mit „vollständig“ und zu 74 % mit „überwiegend“. Die direkte Abfrage der Akzeptanz wurde zudem sowohl aus der Selbstsicht als auch der Perspektive der übrigen schulischen Akteure auf einer sechsstufigen Antwortskala (1 = überhaupt nicht akzeptiert bis 6 = vollkommen akzeptiert) vorgenommen. Die Ergebnisse zeigen, dass die Schulinspektion unterschiedlich wahrgenommen wird. Sowohl aus der Selbstsicht als auch aus der Perspektive der Lehrkräfte wird der Schulleitung die größte Akzeptanz der Schulinspektion bescheinigt, gefolgt von der Elternschaft. Die Gruppe der Lehrkräfte schätzt die Akzeptanz der Schulinspektion im Vergleich zu den anderen schulischen Akteuren geringer, jedoch immer noch moderat ein. Weitere Indikatoren für die Akzeptanz können auch in der Wahrnehmung von negativen Folgen, der Einschätzung des Nutzens der Schulinspektion, der zeitlichen Belastung sowie der Diagnosegüte der Inspektoren gesehen werden. Dabei werden potentielle unerwünschte Folgen der Schulinspektion (z. B. Demotivation, Konflikte, Belastungen) aus der Perspektive der Schulleitungen und Lehrkräfte als relativ gering ausgeprägt wahrgenommen. Bedenklich ist, dass der Nutzen der Schulinspektion für die eigene pädagogische Arbeit sowie für die Schule sowohl von den Schulleitungen als auch Lehrkräften als eher gering eingeschätzt wird. Insbesondere die Schulleitungen empfinden die zeitliche Belastung durch die Schulinspektion als sehr hoch (43 %), während der entsprechende Wert bei den Lehrkräften bei 14 % liegt. Analog zu den Ergebnissen von Gärtner, Hüsemann und Pant (2009) wird die Diagnosegüte der Inspektoren von den Schulleitungen generell positiv beurteilt, während Lehrkräfte diese tendenziell kritischer einschätzen.

<sup>20</sup> Vierstufige Antwortskala (vgl. Gärtner/Wurster 2009): 1 = gar nicht, 2 = wenig, 3 = überwiegend, 4 = vollständig; zusätzliche Kategorie: weiß nicht.

Zusammenfassend zeigen die Ergebnisse des Forschungstyps I, dass Schulleitungen und Lehrkräfte überwiegend positive Akzeptanzbekundungen gegenüber der Schulinspektion äußern. Dabei ist bei Schulleitungen tendenziell eine höhere Akzeptanzwahrnehmung als bei Lehrkräften zu beobachten. Die Ergebnisse stützen sich vor allem auf Angaben von Schulleitungen. Für die übrigen schulischen Akteure wie Eltern und Schüler liegen nur vereinzelte Befunde zur Akzeptanz und wahrgenommenen Qualität der Schulinspektion in Deutschland vor (z. B. in Sommer 2011). Trotz positiver Akzeptanzbekundungen stellt sich die Frage, wie belastbar die Ergebnisse vor dem Hintergrund sind, dass sie von den in die Schulinspektion involvierten Mitarbeitern oder im Rahmen von wissenschaftlichen Begleituntersuchungen im ministeriellen Auftrag erhoben und ausgewertet wurden. Die sozial erwünschte Beantwortung spielt bei Akzeptanzfragen oftmals eine nicht zu vernachlässigende Rolle (vgl. Gruschka 2010, S. 76). Ob Lehrkräfte ein bestimmtes Ausmaß an Akzeptanz aufweisen müssen, damit die Ergebnisse überhaupt genutzt werden, ist bislang noch ungeklärt und stellt ein Forschungsdesiderat dar. Dass auch die Akzeptanz der Schulinspektion und ihrer Ergebnisse eine zentrale Rolle bei der Ergebnisnutzung und der Einleitung von Folgemaßnahmen spielt, wird im nächsten Abschnitt zum Forschungstyp II demonstriert.

### **3.4.3. Forschungstyp II: Auswirkungen der Schulinspektion auf die einzelschulische Schulentwicklung**

Die Untersuchungen des Forschungstyps II untersuchen die im Anschluss an die Schulinspektionen eingeleiteten Maßnahmen zur Schul- und Unterrichtsentwicklung. Auf internationaler Ebene zeigen sich als diesbezügliche Einflussfaktoren zum einen die professionelle Herangehensweise der Schulleitung im Vorfeld der Schulinspektion und zum anderen die Interaktion zwischen den Lehrkräften und Schulinspektoren während des Schulbesuchs (vgl. Brimblecombe/Ormston/Shaw 1995). Positive Effekte von Schulinspektionen auf Schulentwicklungsprozesse zeigen sich auf der Grundlage englischer Befunde insbesondere in Schulen, welche von der Schulinspektion als nur schwach bewertet wurden (vgl. Matthews/Sammons 2004). Von diesem Befund ausgehend wird eine Unterscheidung zwischen starken und schwachen

Schulen und die Unterscheidung von kurz- und langfristigen Effekten von Schulinspektion gefordert (vgl. Wolf/Janssens 2007, S. 388).

Wenngleich auf internationaler Ebene ein genereller Zusammenhang von Schulinspektion und Schulentwicklung konstatiert wird, deuten mehrere Studien darauf hin, dass der Einfluss eher gering ausfällt (vgl. z. B. Gray/Wilcox 1995). Im Folgenden werden eine niederländische und mehrere deutsche Untersuchungen und die entsprechenden Befunde aus den Bundesländern Bayern, Baden-Württemberg, Berlin, Brandenburg, Hamburg und Hessen präsentiert.

Auf der Grundlage niederländischer Schulinspektionen haben Ehren und Visscher (2008) eine Fallstudie an zehn Schulen unter Berücksichtigung der Ausprägung des schulischen Innovationspotentials (hoch vs. niedrig) sowie unterschiedlicher Inspektorenstile (direktiv vs. reserviert) durchgeführt. Dabei wurde die Anzahl der nach der Inspektion eingeleiteten Maßnahmen als Indikator für die Schulentwicklung herangezogen. Die Ergebnisse verdeutlichen, dass alle zehn Schulen die zurückgemeldeten Inspektionsergebnisse zur Schulentwicklung genutzt haben und noch sechs Monate nach der Schulinspektion mit der Erarbeitung von Entwicklungsmaßnahmen beschäftigt waren. Bei den abgeschlossenen Entwicklungsmaßnahmen handelte es sich insbesondere um einfacher umzusetzende und eher nebensächliche Maßnahmen, zu denen die Schulinspektoren Empfehlungen ausgesprochen hatten. Zwischen der Anzahl der ergriffenen Maßnahmen und dem Akzeptanzgrad der Schulen bezüglich der Rückmeldungen bestand kein Zusammenhang. Zudem zeigte der Vergleich von Schulen mit hohem und niedrigem Innovationspotential keine Unterschiede bezüglich der Anzahl umgesetzter Maßnahmen. Folgende Aspekte zeigten ebenfalls keinen Zusammenhang auf: Anzahl der rückgemeldeten Stärken und Schwächen, Umfang der Rückmeldung, Anzahl der Empfehlungen, Anzahl der Treffen zwischen Schulen und Schulinspektoren. Hingegen trägt die Art der Rückmeldung dann signifikant zur Schulentwicklung bei, wenn die schwachen Inspektionsbewertungen berücksichtigt werden und Vereinbarungen über Verbesserungen der schwachen Aspekte innerhalb einer bestimmten Frist getroffen werden. Insgesamt gilt die Studie von Ehren und Visscher (2008) und ihre Ergebnisse in der Forschungscommunity als bedeutsam und regt weitere Forschung an. Allerdings kann der herangezogene Indikator in Form der eingelei-

teten Maßnahmenanzahl kritisch bewertet werden, da er keine Aussagen über die Qualität und den Erfolg der Maßnahmen macht.

Eine frühe deutsche Studie zu den Konsequenzen aus Schulinspektionen wurde im Jahr 2006 von der bayerischen Qualitätsagentur durchgeführt (vgl. Huber 2008; 2006). Befragt wurden bayerische Schulleitungen (N = 311) zu den Umsetzungsprozessen der Rückmeldungen im Anschluss an die Schulinspektion. Zum Zeitpunkt der Befragung gaben 43 % der Schulleitungen an, bislang keine Maßnahmen umgesetzt zu haben und ein Drittel der Schulen (32 %) befand sich in der Phase der Maßnahmenumsetzung. Dabei wurden zwischen einer und sechs Maßnahmen umgesetzt und die Schulen benötigten unterschiedlich viel Zeit, um eine Zielfindung herbeizuführen. Im Durchschnitt nahm dieser Prozess fünf Monate in Anspruch. Die Überprüfung der Maßnahmen wurde nach durchschnittlich 19,5 Monaten eingeleitet. Auch hier deutet die Streuung auf große Differenzen hin. Insgesamt äußert sich die Mehrheit der Schulleitungen sehr positiv und zufrieden zum Erfolg der eingeleiteten Maßnahmen. Auch bei der Globaleinschätzung der Schulinspektion zeigt sich ein recht positives Bild, jedoch mit erheblichen Streuungen.

Gärtner, Hüsemann und Pant (2009; s. o.) lieferten auf der Grundlage einer Schulleiterbefragung in Brandenburg (N = 182) erste Erkenntnisse zu den Auswirkungen von Schulinspektionen auf Schulentwicklungsmaßnahmen im Vorfeld, während und im Anschluss an die Schulinspektion. Berücksichtigt wurden dabei Schulen, welche in den Schuljahren 2005/06 und 2006/07 inspiziert wurden. Gemäß der Selbstauskunft der befragten Schulleitungen wirkte sich die Ankündigung einer Schulinspektion vor allem auf Aktivitäten im Bereich der Dokumentation der Schulunterlagen aus. Nach der Schulinspektion gestaltete sich die Informationsweitergabe der Berichtsergebnisse sehr unterschiedlich und reichte vom passiven Auslegen des Berichts bis zur aktiven Einforderung von Rückmeldungen aus den Gremien. Laut Angaben der Schulleitungen wurden im Durchschnitt knapp vier und maximal elf Maßnahmen im Bereich der Schul- und Unterrichtsentwicklung im Anschluss an die Rückmeldungen der Schulinspektionsergebnisse eingeleitet, wobei im ersten Inspektionsjahr signifikant mehr Maßnahmen (durchschnittlich 4,4 Maßnahmen) eingeleitet wurden als im zweiten Jahr (durchschnittlich 3,2 Maßnahmen). Inhaltlich können hierbei die Maßnahmenschwerpunkte den Themenfeldern Schulpro-

grammarbeit und schulinterne Evaluation zugeordnet werden. Etwas mehr als ein Achtel der Schulen gab an, keine Maßnahmen eingeleitet zu haben. Nahezu die Hälfte der befragten Schulleitungen nimmt einen Einfluss der Inspektion auf die Unterrichtspraxis wahr. Der von Ehren und Visscher (2008) eruierte Befund, dass die Anzahl der eingeleiteten Maßnahmen nicht mit der Schulbewertung zusammenhängt, wird von der Brandenburger Schulleiterbefragung untermauert.

Weitere Befunde zur Wirkung von Schulinspektion auf Schulentwicklung liegen auf der Grundlage einer quasi-experimentellen Feldstudie von Gärtner (2011) vor. Darin wurden Schulleitungen und Lehrkräfte zu zwei Messzeitpunkten, jeweils zu Beginn des Schuljahres 2008/09 und 2009/10, zur Schulqualität in Anlehnung an die Qualitätsrahmen in Berlin und Brandenburg befragt. Das quasi-experimentelle Design der Studie ermöglichte, die schulischen Entwicklungen ein und zwei Jahre nach der Inspektion sowie im Jahr der Inspektion selbst zu untersuchen und den Entwicklungen in nicht inspizierten Schulen als Kontrollgruppe gegenüberzustellen. Die Ergebnisse zeigen, dass die Schulinspektion relativ wenig Einfluss auf die Veränderung der Schulqualität ausübt, da diese im zeitlichen Verlauf relativ konstant bleibt. Als interessant erweist sich die Entwicklung in der Kontrollgruppe, bei welcher sich im Vergleich zu den inspizierten Schulen die stärkste Veränderung zwischen den beiden Messzeitpunkten beobachten lässt. Diesen Befund interpretiert Gärtner als Hinweis darauf, dass sich die Wirkungen der Schulinspektion bereits in der Ankündigung und Vorbereitung auf die Schulinspektion zeigen und die Impulsgebung durch die Schulinspektion nicht erst in der Phase nach der Schulinspektion erfolgt (vgl. Gärtner 2011, S. 158).

Ein weiteres interessantes Forschungsergebnis zeigt sich in der Typisierung von Schulen im Umgang mit der Schulinspektion und ihren Ergebnissen (vgl. Wurster/Gärtner 2010). Auf der Grundlage der bis zum Ende des Schuljahres 2007/08 in Berlin und Brandenburg befragten Schulleitungen und Lehrkräfte in 580 Schulen konnten vier Schultypen identifiziert werden:

1. Schultyp: (Selbst-)Zufriedene Schulen (ca. 18 %)
2. Schultyp: Aktive Schulen (ca. 29 %)
3. Schultyp: Unzufriedene Schulen (ca. 29 %)
4. Schultyp: Resignierte Schulen (ca. 24 %)

Die Schultypen sind hinsichtlich ihrer Inspektionsergebnisse, schulinternen Kommunikation, Ergebnisreflexion, Nutzung externer Unterstützung zur Realisierung der Schulentwicklungsmaßnahmen, Akzeptanz und ihres empfundenen Nutzens der Schulinspektion zu differenzieren. Bei der Beschreibung der Schultypen fällt auf, dass die Schulen mit eher negativen Merkmalsausprägungen (Schultyp 3 und 4) zu über 50 % vertreten sind. Die unterschiedlichen Schultypen zeigen die Notwendigkeit auf, den Schulen im Anschluss an die Inspektion differenzierte Unterstützungsleistungen anzubieten.

Eine weitere Untersuchung liegt in Form einer Schulleitungsbefragung der in den Schuljahren 2007 bis 2011 inspizierten Schulen in Hamburg (N = 166) vor und gibt Einblicke in die Nutzung und wahrgenommene Nützlichkeit der Schulinspektionsergebnisse (vgl. Pietsch 2011). Analog zu den o. g. Ergebnissen, z. B. von Gärtner, Hüsemann und Pant (2009), wurde auch in dieser Untersuchung die Diagnosegüte der Schulinspektoren von den Schulleitungen zu einem hohen Anteil als zuverlässig eingeschätzt. Es zeigt sich zudem, dass in der Mehrzahl der Schulen (76 %) aus den Inspektionsergebnissen konkrete Maßnahmen abgeleitet wurden. Hingegen haben 21 % der Schulen die Ergebnisse nicht für die Initiierung konkreter Maßnahmen genutzt. Im Hinblick auf die wahrgenommene Nützlichkeit der Schulinspektion wurden von den Schulleitungen insbesondere die Identifikation von Schwachstellen und eine Zusammenschau der Schulqualität genannt. Zu einem etwas geringeren Anteil, aber dennoch positiv konnotiert, wurde die Nützlichkeit der Inspektionsbefunde in Hinblick auf die Schul- und Unterrichtsentwicklung bewertet. Hingegen haben ca. 20 % der Schulleitungen die Inspektionsergebnisse generell für nicht aussagekräftig und anschlussfähig befunden. Die empfundene Relevanz der Schulinspektionsergebnisse ist jedoch insofern zentral, als von ihr zu einem wesentlichen Anteil die Nutzung der Ergebnisse determiniert ist:

Es gibt somit einen stark positiven Zusammenhang zwischen der wahrgenommenen Relevanz der Inspektionsbefunde und der Einleitung konkreter Entwicklungsmaßnahmen auf Ebene der Einzelschule. Mit anderen Worten ist die Chance, dass konkrete Maßnahmen aus diesen Informationen abgeleitet werden, an Schulen, an denen die Ergebnisse der Inspektion als wichtig erachtet werden, mehr als achtmal höher als an Schulen, an denen die Ergebnisse als irrelevant

betrachtet werden. Diesen Befund kann man jedoch auch anders lesen: Gelingt es der Schulinspektion Hamburg nicht, den Schulen relevante Informationen zurückzumelden, sinkt die Chance, dass konkrete Maßnahmen aus den Inspektionsbefunden abgeleitet werden, um fast 90 Prozent auf ein Achtel (Pietsch 2011, S. 11f.).

Dieser Befund betont die Notwendigkeit, Untersuchungen des Forschungstyps I und die Sicherstellung einer hohen Akzeptanz der beteiligten Akteure nicht zu vernachlässigen, da die Akzeptanz, die Einstellungen und die wahrgenommene Nützlichkeit der Inspektion und ihrer Ergebnisse wichtige Prädiktoren und Gelingensbedingungen für die Ergebnisnutzung darstellen. Gleichwohl dürfen diese eher weichen Effektivitätskriterien sowie grundsätzlich die Schulinspektion per se – wie ein weiteres Ergebnis Pietschs verdeutlicht – nicht überbewertet werden: „Etwas mehr als 12 Prozent der Nutzungsunterschiede lassen sich über Merkmale der Inspektion selbst erklären“ (ebd., S. 16). Dies bedeutet, dass vor allem andere, noch unbekannte Faktoren Einfluss auf die Nutzung der Ergebnisse haben. Exemplarisch können das innerschulische Veränderungssetting und bereitgestellte Unterstützungsmaßnahmen angeführt werden (vgl. ebd.).

Eine den Forschungshorizont zu den Wirkungen der Schulinspektion erweiternde Studie bildet der vom hessischen Kultusministerium gegebene Auftrag zur *Schulinspektion in Hessen – Wirkungen auf die Qualitätsentwicklung von Schulen und die Arbeit der Schulaufsicht* (vgl. Böttcher/Keune/Neiwert 2010; Böttcher/Keune 2011, S. 132-139). Dabei wurden zwischen 2008 und 2009 standardisierte Lehrerbefragungen (N = 400) an 15 Sekundarschulen, schriftliche Befragungen und Experteninterviews mit 15 Schulleitungen sowie 15 Schulaufsichtsmitarbeitern durchgeführt. Zusätzlich wurden Dokumente in Form von Schulinspektionsberichten sowie Zielvereinbarungen der 15 untersuchten Schulen berücksichtigt. Die zahlreichen Erkenntnisquellen wurden in einem Strukturierungsmodell zur Schulentwicklung nach einer Schulinspektion gebündelt. Dieses Modell weist fünf idealtypische Phasen des Wandels durch Schulinspektion auf: Initialisierung, Konzipierung, Mobilisierung, Umsetzung und Verstetigung (vgl. Böttcher/Keune 2011, S. 134f.). Im Hinblick auf die Schulentwicklungsprozesse sind sich alle befragten Schulleitungen einig, dass durch die Schulinspektion innerschulische Prozesse ausgelöst wurden. Hingegen ist die Sicht der befragten Lehrkräfte auf die Maßnahmen und Prozesse, die

durch die Schulinspektion angestoßen wurden, deutlich kritischer: Der Mittelwert der Skala *Schulentwicklung durch Schulinspektion* von 1,9<sup>21</sup> liegt unterhalb des theoretischen Skalenmittelwerts von 2,5 und signalisiert eine eher skeptische Haltung. Ca. 80 % der befragten Lehrkräfte sind der Meinung, dass die Schulinspektion keinen innerschulischen Wandlungsprozess hervorgerufen habe und 60 % verneinen die Aussage, die Schulen seien gestärkt aus der Schulinspektion hervorgegangen. Auch die Skala *Umsetzung von Maßnahmen und Prozessen* liegt mit einem Mittelwert von 2,3 unterhalb der Skalenmitte. Als bedenklich erweist sich, dass etwas mehr als die Hälfte der befragten Lehrkräfte (58,6 %) angibt, nach der Schulinspektion keine neuen Maßnahmen und Prozesse eingeführt zu haben.

Zielvereinbarungen zwischen Schulen und Schulaufsicht bilden eine Schnittstelle zwischen den Inspektionsbefunden und den darauf bezogenen Maßnahmen der Schul- und Unterrichtsentwicklung. Die Befunde mehrerer Studien zeigen, dass ein starker Zusammenhang zwischen den gemessenen Inspektionsergebnissen und diesbezüglichen Maßnahmeneinleitungen besteht, die in den Zielvereinbarungen dokumentiert werden. In einer bayerischen Schulleitungsbefragung (vgl. Huber 2008; s. o.) gaben 98 % der Schulleitungen an, die zentralen Empfehlungen des Inspektionsberichts in den Zielvereinbarungen aufgegriffen zu haben. Dabei wurde der Zielvereinbarungsprozess zwischen Schulen und Schulaufsicht als kooperativ wahrgenommen. Etwas mehr als die Hälfte der befragten Schulleitungen war der Ansicht, dass der Zielvereinbarungsprozess die Akzeptanz der Lehrkräfte gegenüber der Schulinspektion erhöht habe.

Im Rahmen der Berliner Schulleitungs- und Lehrkräftebefragung (vgl. Gärtner/Wurster 2009) wurden in 86 % der Schulen nach der Ergebnisrezeption Folgemaßnahmen eingeleitet. In 26 % der Schulen stehen alle Ziele, in 17 % die meisten und in 41 % einige Ziele aus den Zielvereinbarungen im Zusammenhang mit dem Inspektionsbericht. Hierzu halten Gärtner, Hüsemann und Pant (2009, S. 15) fest: „Schulleitungen scheinen die Inspektion zum Anlass zu nehmen, bestimmte Aktivitäten der Schul- und Unterrichtsentwicklung einzuleiten, zu deren Implementation sie ohne diesen äußeren Anlass nicht in der Lage waren“.

---

<sup>21</sup> Vierstufige Antwortskala (vgl. Böttcher/Keune 2010): 1 = trifft gar nicht zu; 2 = trifft eher nicht zu; 3 = trifft eher zu; 4 = trifft voll zu.

Inhaltlich lassen sich die vereinbarten Ziele i. d. R. allen Bereichen der jeweiligen Qualitätsrahmen zuordnen. In Bayern (vgl. Huber 2008) bilden *Schulleitung und Schulmanagement, Organisation der Arbeitsabläufe* sowie die *Zusammenarbeit im Kollegium* inhaltliche Schwerpunkte der Schulentwicklungsmaßnahmen. Im Bereich der Unterrichtsentwicklung sind dies Themenfelder wie *Individualisierung, Methodenvielfalt, Leistungserhebung* und *selbständiges Lernen*. Im Rahmen der Schulleitungs- und Lehrkräftebefragung in Berlin (vgl. Gärtner/Wurster 2009) bezieht sich die große Mehrheit der Maßnahmen aus dem Bereich Unterrichtsentwicklung auf die Themenfelder *Unterrichtsgestaltung/Lehrerhandeln im Unterricht, Sozial- und Methodenkompetenzen* sowie *Schülerunterstützung und -förderung*. Im Bereich der Schulentwicklung sind dies *Ziele und Strategien der Qualitätsentwicklung, Lehrerprofessionalisierung, Personalentwicklung* und *Schulkultur*.

Bisher wurden von der Schulinspektionsforschung in Deutschland primär die intendierten Wirkungen beleuchtet. Erste Erkenntnisse zu den nicht-intendierten Konsequenzen von Schulinspektionen liefert eine qualitative Fallanalyse mit Leitfadeninterviews mit den schulischen Akteuren im Kontext der baden-württembergischen Fremdevaluation (vgl. Lambrecht/Kotthoff/Maag Merki 2011). Die Ergebnisse eines ausgewählten schulischen Fallportraits zeigen, dass nicht nur die offiziellen Funktionen der Schulinspektion, sondern auch weitere, schulspezifische und schwer vorhersagbare Funktionen in Erscheinung treten können. In der untersuchten Schule wurden die Ergebnisse der Schulinspektion von der Schulleitung gegenüber dem Kollegium als strategisches Werkzeug und Argumentationsstütze eingesetzt. Die damit einhergehende Folgeerscheinung war eine „Externalisierung eines innerschulischen Konflikts“, bei der die Schulinspektion als „Schiedsrichter“ fungieren sollte (ebd., S. 188). Die sowohl von der Schulleitung als auch den Lehrkräften erwartete Entlastung durch die Inspektion ist allerdings nicht eingetreten.

Zusammenfassend zeigen die Ergebnisse des Forschungstyps II, dass der überwiegende Teil der frühen Befunde (z. B. Huber 2006, 2008; Gärtner/Hüsemann/Pant 2009) auf der Grundlage quantitativer Maßnahmenenerfassungen im Anschluss an die Schulinspektion zustande gekommen ist. Die Bestimmung der Qualität und des Erfolgs der eingeleiteten Maßnahmen gilt bislang noch als Forschungsdesiderat. Aktuell sind innovative Forschungsdesigns zu erkennen,

welche über die Bestimmung der Maßnahmenanzahl hinausgehen, nicht-intendierte Effekte der Schulinspektion beleuchten sowie verstärkt Zusammenhänge und Gelingensbedingungen der Schulinspektion hinsichtlich der Initiierung von Schulentwicklung fokussieren.

#### **3.4.4. Forschungstyp III: Einfluss der Schulinspektion auf die Schülerleistungen**

In Untersuchungen des Forschungstyps III erfolgt die Analyse der Effekte der Schulinspektion auf die Lernzuwächse von Schülern. Die Schwierigkeit solcher Vorhaben besteht darin, den Einfluss der Schulinspektion auf die Schülerleistungen zu isolieren, insbesondere vor dem Hintergrund, dass diese unabhängig von der Schulinspektion kontinuierlich steigen (vgl. Matthews/Sammons 2004, S. 24). Womöglich sind darauf auch die im Hinblick auf diese Fragestellung insgesamt uneinheitlichen Ergebnisse zurückzuführen.

Internationale Befunde, die den Zusammenhang zwischen Schulinspektionen und Schülerleistungen untersuchen, sind insgesamt rar und widersprüchlich. Sie wurden bisher nahezu ausschließlich in England durchgeführt, da dort eine längere Tradition der Schulinspektion und der standardisierten Leistungsfeststellung in Form von GCSE<sup>22</sup> herrscht. Die Ergebnisse von vier englischen Studien, die anhand unterschiedlich komplexer Designs den direkten Zusammenhang zwischen Schulinspektion und den Prüfungsergebnissen von Schülern messen, deuten darauf hin, dass die Ofsted-Besuche keinen positiven Einfluss auf die Schülerleistungen ausüben (vgl. Cullingford/Daniels 1999; Wilcox/Gray 1996; Shaw u.a. 2003; Rosenthal 2004). Zum Teil zeigen die Studien sogar, dass sich die Testleistungen im Jahr nach der Inspektion tendenziell sogar verschlechtern (vgl. Cullingford/Daniels 1999; Rosenthal 2004). Für die ausbleibenden positiven bzw. die diagnostizierten negativen Effekte der Schulinspektionen machen die Forscher den Stress der schulischen Akteure vor und während der Inspektion sowie die aufwendige Vorbereitung im Vorfeld einer Schulinspektion verantwortlich. Darin zeigen sich die nicht intendierten Wirkungen resp. Nebeneffekte von Schulinspektionen. Shaw u.a. (2003) konnten durch eine logistische Mehrebenenanalyse auf der Basis eines Vergleichs

---

<sup>22</sup> General Certificate of Secondary Education

zwischen inspizierten und nichtinspizierten Schulen (N = ca. 3.000 Schulen; Inspektionszeitraum zwischen 1992 und 1997) herausfinden, dass für bestimmte privilegierte Schulen, z. B. Mädchenschulen, schwache positive Effekte der Schulinspektion zu beobachten sind. Für die Mehrzahl der herkömmlichen Schulen (comprehensive schools) sind allerdings keine Unterschiede zwischen den Vergleichsgruppen bzw. z. T. sogar negativere Testleistungen in den inspizierten Schulen festzustellen. Weitere Ergebnisse des Forschungstyps III basieren auf Längsschnitt-/Paneldaten aus England (N = 2.000 Schulen) zwischen 1994 bis 1998 und lassen sich in die Reihe früherer Befunde einordnen, dass Schulinspektionen entweder keinen positiven Effekt oder im Inspektionsjahr sogar tendenziell einen hindernden Effekt auf die Schülerleistungen im GCSE haben (vgl. Rosenthal 2004). Für zukünftige Auswertungen der Schülerleistungsdaten in England lässt sich vermuten, dass die Effekte der Schulinspektion auf die Schülerleistungen größer und positiver ausfallen könnten, da sich der Stress und die empfundene Belastung der schulischen Akteure durch die Ofsted-Reformen im Jahr 2005 verringert haben dürften.

Auch in den Niederlanden sind in den letzten Jahren Auswertungen von Schülerleistungen im Zusammenhang mit der Schulinspektion durchgeführt worden. In einer umfangreichen Studie haben Luginbuhl, Webbink und de Wolf (2009) CITO-Testdaten<sup>23</sup> von zwölfjährigen Schülern an ca. 6.000 Schulen im Primarbereich zwischen den Jahren 1999 und 2003 ausgewertet. Sie konnten belegen, dass Schulinspektionen schwache positive Effekte auf Schülerleistungen haben. Diese wurden eher nach intensiveren Schulinspektionsformen, im Fach Mathematik und in Schulen mit einem höheren Anteil an benachteiligten Schülern festgestellt. Die Forscher konkludieren: „inspections do no harm, but seem to have little or no effect on student performance“ (Luginbuhl/Webbing/de Wolf 2009, S. 235).

Studien zu den Auswirkungen von Schulinspektionen auf Schülerleistungen sind generell auch in Deutschland durchführbar, da die Datengrundlage in Form von Ergebnissen aus zentralen Lernstandserhebungen und Abschlussprüfungen gegeben ist. Wie in allen inspektionserprobten Ländern sind die Ergebnisse solcher Studien nur schwer abzusichern, da sich der Effekt der Schulinspektionen

<sup>23</sup> Das staatliche Institut für Testentwicklung in den Niederlanden (CITO) entwickelt landesweite Tests am Ende eines Schuljahres.

spektion nicht vollständig isolieren lässt. Erschwerend kommt hinzu, dass in Deutschland in den letzten Jahren ein Konglomerat an neuen Steuerungsinstrumenten eingeführt wurde, von denen jeweils positive Wirkungen erwartet werden. Trotz dieser Schwierigkeiten ist hierzulande jüngst eine erste Studie zum Zusammenhang zwischen Schulinspektionen und Schülerleistungen in Angriff genommen worden (vgl. Gärtner/Pant 2011). Dabei wurden VERA-Ergebnisse inspizierter Schulen in Brandenburg aus den Schuljahren 2005/06 (N = 54 Grundschulen) sowie 2006/07 (N = 70 Grundschulen) in Form eines Leistungsindex auf Schulebene aggregiert, um sie mit den Inspektionsergebnissen in Beziehung zu setzen. Auf der Grundlage der VERA-Ergebnisse wurde der Prozentrang einer Schule (von 1 = beste Schule bis 100 = schlechteste Schule) gebildet und die Schulen getrennt nach den Gesamtbewertungen der Schulinspektion durch einen Mediansplit in eine obere und untere Hälfte aufgeteilt. Die Befunde zeigen unterschiedliche Entwicklungen je nach Jahr der Inspektion auf. Während sich die im Schuljahr 2005/06 inspizierten Schulen hinsichtlich ihrer VERA-Ergebnisse im Jahr der Inspektion selbst und den darauf folgenden Schuljahren nicht unterscheiden und somit eine Entwicklung ausbleibt, zeigen sich in den im Schuljahr 2006/07 inspizierten Schulen im Längsschnitt Unterschiede bezüglich der VERA-Ergebnisse. Insbesondere die der oberen Hälfte (Prozentrang 1 - 50) zugehörigen Schulen erzielten im Inspektionsjahr und den beiden darauf folgenden Schuljahren bessere VERA-Ergebnisse. Bei den Schulen in der unteren Hälfte zeigt sich ein gegenläufiger Trend.

Insgesamt liefern die Ergebnisse zum Zusammenhang zwischen den Ergebnissen von Schulinspektionen und den Schülerleistungen keine eindeutigen Befunde: Es sind sowohl Verbesserungen als auch keine Entwicklungen oder leichte Verschlechterungen erkennbar. Noch basieren in Deutschland die Befunde auf geringen Stichprobenzahlen. Für kommende Analysen sind größere Stichproben sowie ein Längsschnittdesign von mehreren Schuljahren gefordert. Zudem ist es sinnvoll, gemessene Schülerleistungen nicht mit zu zeitnahen Schulinspektionsbesuchen in Verbindung zu bringen, da Schulentwicklung nicht von heute auf morgen geschieht, sich nicht unmittelbar in den Schülerleistungen niederschlägt und für ihre zielgerichtete und von allen Akteuren getragene Entfaltung mehrere Jahre benötigt.

### **3.4.5. Forschungstyp IV: Messgüte der Schulinspektionsinstrumente**

Aktuell mehren sich in der Forschungslandschaft zum Thema Schulinspektion Untersuchungen zur Validität der Schulinspektion. Insbesondere Untersuchungen zur Zuverlässigkeit von Unterrichtsbeobachtungen im Rahmen von Schulinspektionen stehen im Fokus. So konnten Pietsch und Tosana (2008) auf der Basis von 14 Inspektoren und 13 Schulen in Hamburg zeigen, dass die Bewertung der Unterrichtsqualität zu rund 9 % der Gesamtvariation inspektorenbedingte Verzerrungen aufweist und dieser Anteil trotz umfangreicher Schulungen Bestand hat. Jüngst wurde eine länderübergreifende Studie zur Beurteilerübereinstimmung (absoluter Konsens) und Inter-Beobachter-Reliabilität (relativer Konsens) von Schulinspektoren bezogen auf Unterrichtsbeobachtungen durchgeführt (vgl. Müller/Pietsch 2011). In die Studie wurden 167 Inspektoren aus den sechs Bundesländern Berlin, Brandenburg, Hamburg, Hessen, Niedersachsen und Nordrhein-Westfalen miteinbezogen. Die Inspektoren hatten die Aufgabe, drei videografierte Unterrichtssequenzen mit den jeweils eigenen Unterrichtsbeobachtungsbögen und denen der anderen Bundesländer zu bewerten. Rund 50 % der Unterschiede der Unterrichtsqualität einzelner Sequenzen gingen auf die Beurteilereffekte der Inspektoren zurück. Dieser Effekt erweist sich auf der Basis einzelner Unterrichtssequenzen als stark, wird jedoch auf Schulebene deutlich relativiert. Darüber hinaus erweist sich die Übereinstimmung der Unterrichtsbeobachter als hoch und ihre Austauschbarkeit realisierbar, sowohl bei den bekannten als auch den fremden Unterrichtsbeobachtungsbögen, was auf ein nahezu allgemeingültiges Verständnis von Unterrichtsqualität hindeutet. Dabei lassen sich Schulungseffekte der Inspektoren mit den ihnen bekannten Unterrichtsbeobachtungsbögen belegen, während berufsbezogene Merkmale nur wenig Erklärungskraft bezüglich der beobachteten Bewertungsunterschiede haben. Hinreichend reliable Ergebnisse zur Unterrichtsqualität erfordern eine durchschnittliche Unterrichtsbewertung von vier Inspektoren.

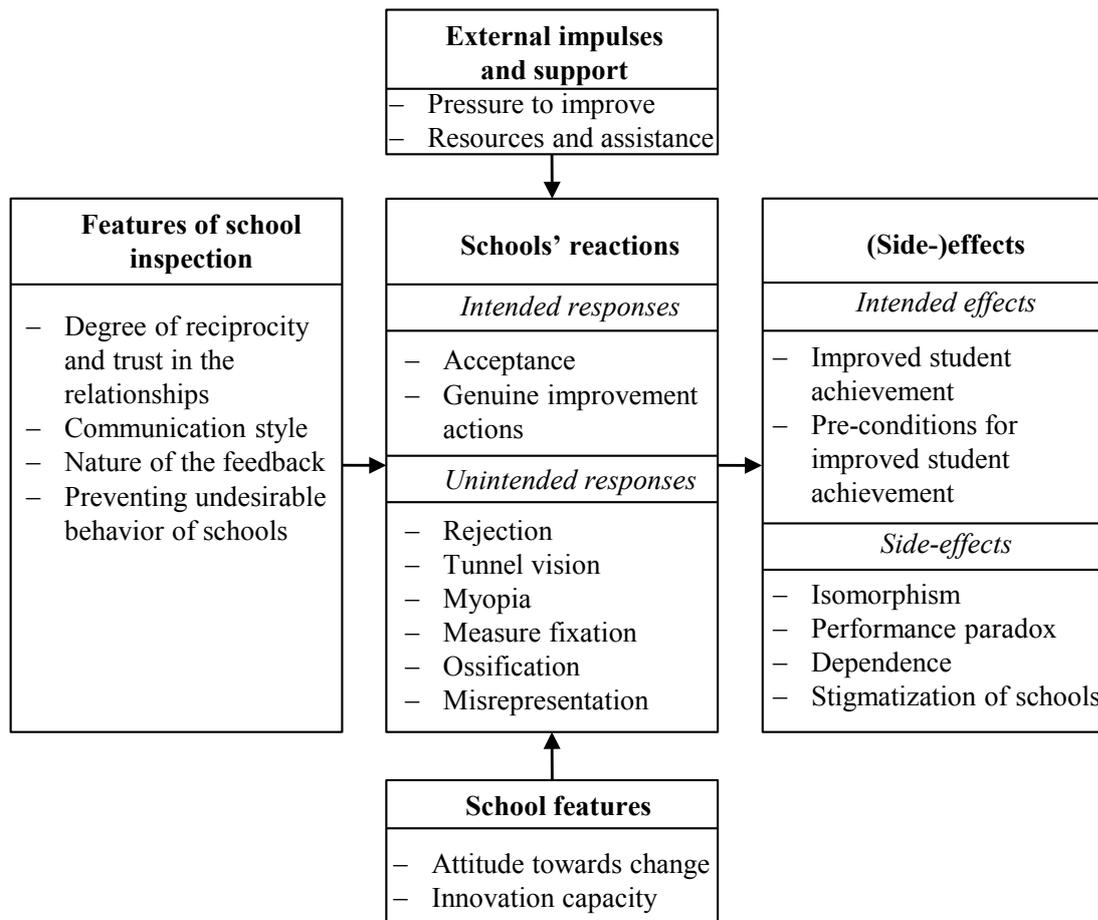
Während das Instrument Unterrichtsbeobachtung in der Tendenz über eine zuverlässige Messgüte verfügt, zeigt sich bei den Interviews im Rahmen der Schulinspektion ein anderes Bild: In einem hessischen und einem Hamburger Fallbeispiel wurden einerseits die Übereinstimmung zwischen Gesprächsinhal-

ten, Protokollen und Referenzrahmen und andererseits die Selektions- und Interpretationseffekte in Interviewprotokollen der Schulinspektion untersucht. Für die Analyse wurden 47 Protokolle eines zehnminütigen Ausschnitts eines Schülerinterviews in Hessen ausgewertet sowie fünf Protokolle vollständiger Schulleitungsinterviews aus Hamburg (vgl. Tosana/Lambrech/Perels/Bardowiecks 2011). Diese Protokolle sind jeweils auf der Basis von leitfadengestützten, d. h. teilstandardisierten Interviews entstanden. Die Ergebnisse zeigen deutlich, dass leitfadengestützte Interviews keine objektive Datenquelle darstellen, da sie durch Auslassungen, Selektions- und Interpretationseffekte Verzerrungen ergeben. Ein Grund für diese Effekte liegt im Spannungsfeld begründet, in welchem sich die Interviewer und Protokollanten bewegen müssen: Einerseits müssen sie sich an den standardisierten Leitfaden halten, andererseits sollten sie flexibel auf die Antworten ihrer Interviewpartner reagieren.

Insgesamt zeichnen die Ergebnisse zur Messgüte der Schulinspektionsinstrumente ein heterogenes Bild und zeigen die Notwendigkeit auf, die Zuverlässigkeit einzelner Instrumente eingehender zu untersuchen. Da das Instrumentarium im Rahmen der Schulinspektion vielfältig ist und durch die Aggregation der unterschiedlichen Datenquellen auf der Ebene der einzelnen Qualitätsaspekte eine einzige Bewertungsstufe zustande kommt, stellt die Untersuchung der Zuverlässigkeit des Gesamtergebnisses ein Forschungsdesiderat dar.

### **3.4.6. Wirkungsmodell zur Schulinspektion**

Zur Einordnung der Forschungstypen zur Schulinspektion und der Determinanten, welche bei der Wirkung von Schulinspektion eine Rolle spielen, wird in diesem Abschnitt der Fokus auf das von den niederländischen Wissenschaftlern Ehren und Visscher (2006) entwickelte theoretische Wirkungsmodell gelegt. Es wurde auf der Grundlage bisher vorliegender internationaler Forschungsbefunde und theoretischer Überlegungen entwickelt und dient der Forschung zur Schulinspektion als ein vielfach herangezogenes Modell. Das Modell trägt dazu bei, spezifische exogene und endogene Elemente und deren Beziehungen, die zur Wirkung der Schulinspektion beitragen, strukturiert darzustellen. Abbildung 4 veranschaulicht das Modell. Im Zentrum steht ein abgestufter Prozess mit den drei Elementen Schulinspektionsmerkmale, Reaktionen der Schule und (Neben)Wirkungen (vgl. Ehren/Visscher 2006, S. 66).



**Abb. 4: Wirkungsmodell zur Schulinspektion**

Der mittlere Prozessschritt (Reaktionen der Schule) wird einerseits von externen Impulsen und Unterstützungsleistungen, andererseits von innerschulischen Merkmalen beeinflusst (s. Abb. 4; vgl. Ehren/Visscher 2006, S. 66). In Abhängigkeit der Ausprägungen der Schulinspektionsmerkmale bzw. der Eingangsbedingungen und der schulischen Reaktionen löst der Schulinspektionsprozess unterschiedliche Wirkungen aus. Als Ausprägungen der Schulinspektion erfassen Ehren und Visscher (2006) das Ausmaß des wechselseitigen Vertrauens in die Beziehungen, den Kommunikationsstil zwischen den verschiedenen Akteuren, die Art der Rückmeldung sowie die Vorbeugung unerwünschter Reaktionen der Schulen. Analog zu den Wirkungen der Schulinspektion, welche sich in erwünschte und unerwünschte Ausprägungen differenzieren lassen, können auch die schulischen Reaktionen ausfallen. Zu den erwünschten Reaktionen zählen die Akzeptanz sowie die tatsächliche Umsetzung von Entwicklungs-

maßnahmen zur Qualitätssicherung und -steigerung. Hingegen sind auch folgende unerwünschte Reaktionen denkbar: Ablehnung der Ergebnisse, Tunnelblick, Kurzsichtigkeit, Maßnahmenfixierung, Erstarrung sowie Verzerrung der Ergebnisse. Beeinflusst werden die schulischen Reaktionen – sowohl die intendierten als auch die nicht-intendierten – von externer Seite durch den auf die Schulen ausgeübten Druck zur Entwicklung sowie durch Ressourcen und Unterstützungsleistungen. Schulinterne Einflussfaktoren sind durch die Einstellung gegenüber Veränderungen und durch die Innovationsfähigkeit gekennzeichnet. Bei den durch die Merkmale der Schulinspektion und den schulischen Reaktionen ausgelösten Wirkungen sind als erwünschte Effekte die Steigerung der Schülerleistungen sowie die diesbezüglichen Bedingungen denkbar. Als unerwünschte Wirkungen werden ein schwacher Zusammenhang zwischen den gemessenen und den tatsächlichen schulischen Leistungen sowie die Abhängigkeit und Stigmatisierung der Schulen angeführt.

Bislang sind einzelne Teilbereiche des Modells durch die empirische Forschung untersucht worden. Die zuvor präsentierten Forschungsbefunde lassen sich größtenteils dem Prozessschritt *school's reactions – Intended responses* zuordnen. Die übrigen Prozessschritte und Rahmenfaktoren sind im Forschungsfeld unterrepräsentiert. Insbesondere in Deutschland existieren kaum Studien zu nicht-intendierten Effekten der Schulinspektion. Die Vorteile des Modells können in seiner Strukturierung und Benennung zentraler wirkungsbeeinflussender Faktoren gesehen werden. Es fehlen allerdings konkrete Wirkmechanismen, welche z. B. die Beantwortung der Frage „Welche spezifischen Schulinspektionsmerkmale und schulischen Reaktionen tragen zu besseren Schülerleistungen bei?“ ermöglichen würden. Ein erster Schritt in diese Richtung wird von Ehren und Visscher aufgezeigt: Sie benennen potentielle Kombinationen von Schulcharakteristika, externen Impulsen und Schulinspektionsmerkmalen und zeigen denkbare schulische Reaktionen und (Neben)Wirkungen auf (vgl. Ehren/Visscher 2006, S. 68). Bspw. erweist es sich als empfehlenswert, dass Schulen mit schwach ausgeprägten Innovationskapazitäten und wenigen externen Impulsen von den Inspektoren eine klare Stärken-Schwächen-Benennung, mögliche Begründungen sowie Entwicklungsmöglichkeiten zurückgemeldet bekommen. Hingegen sollten Schulen mit starkem Innovationspotential weniger direktiv unterstützt werden.

Die Grenzen des Modells zeigen sich in seiner Beschränkung auf die Zielgröße *Schulentwicklung*. Überdies differenziert das Modell nicht zwischen Wirkungen der Schulinspektion auf innerschulischer Ebene und der System- und Steuerungsebene. Letztere müsste mitmodelliert werden, um die Perspektive der Einzelschule zu erweitern. Zudem findet eine Berücksichtigung der institutionellen und konzeptuellen Faktoren der Schulinspektion und ihrer Variationen nur unzureichend statt. Trotz seiner Grenzen stellt das Modell eine wichtige Grundlage für Forschungsaktivitäten dar. Es fungiert bspw. in dem o. g. vom BMBF geförderten Projekt *Externe Evaluation/Schulinspektion und schulische Qualitätsentwicklung* als theoretischer Hintergrund (vgl. BMBF 2010b).

### **3.4.7. Zusammenfassung**

Vor dem Hintergrund z. T. widersprüchlicher Befunde lässt sich der Forschungsstand der bisherigen Schulinspektionsforschung in Deutschland wie folgt zusammenfassen: Insgesamt werden der Schulinspektion von den schulischen Akteuren überwiegend positive Akzeptanzbekundungen bescheinigt, obwohl ein ausgewogenes Aufwand-Nutzen-Verhältnis nicht bestätigt wird. Vorrangig stammen die Befunde zur Akzeptanz von Schulinspektionen von Schulleitungen, gefolgt von Lehrkräften. Aus der Perspektive von Eltern, Schülern und der Schulaufsicht liegen bislang nur wenige Erkenntnisse vor. Es zeigt sich deutlich, dass von den Schulleitungen tendenziell eine größere Akzeptanz im Hinblick auf die Schulinspektion festzustellen ist als von den Lehrkräften. Zudem scheint die Akzeptanz von Schulinspektionen positiv mit dem Inspektionsergebnis zusammenzuhängen, wobei die Verfahrensakzeptanz im Vergleich zur Ergebnisakzeptanz der Schulinspektion positiver ausfällt. Schließlich nimmt die Akzeptanz von Schulinspektionen insofern eine wichtige Rolle ein, als sie eine zentrale Voraussetzung für die Nutzung der Inspektionsergebnisse darstellt.

Die Befundlage zu den Wirkungen und zur Wirksamkeit von Schulinspektionen ist deutlich skeptischer zu bewerten: „Schulinspektion, so die Indizienlage, hat sich bislang eher nicht als ein Instrument der Schulentwicklung erwiesen“ (Böttcher/Keune 2011, S. 139). Bislang hat die Forschung eher Informationen zur Datennutzung als zur Wirksamkeit der Schulinspektion geliefert. Dass die Inspektionsergebnisse von den Schulen genutzt werden, zeigt sich an dem

starken Zusammenhang, welcher zwischen den Inspektionsbefunden und der Einleitung und Umsetzung von Maßnahmen besteht. Die Mehrheit der Studien operationalisiert die Wirksamkeit von Schulinspektionen mit den Indikatoren ‚Anzahl eingeleiteter und umgesetzter Maßnahmen‘. Obwohl mittlerweile belegt ist, dass die in den Zielvereinbarungen enthaltenen Ziele zum überwiegenden Anteil auf der Basis des Inspektionsberichts getroffen werden, liegen bislang keine empirisch gesicherten Befunde zur Qualität der Zielvereinbarungen sowie zu ihrem Potential hinsichtlich der Schulentwicklung vor. Auch die Bestimmung der Qualität und des Erfolgs der Maßnahmen stellt vorerst noch ein Forschungsdesiderat dar.

Die internationalen Studien zu den Auswirkungen von Schulinspektionen auf die Schülerleistungen erweisen sich als heterogen. Tendenziell zeigt sich, dass positive Auswirkungen ausbleiben; z. T. lassen sich aber auch marginale positive oder negative Effekte erkennen. Diese heterogene Befundlage abmildernd muss betont werden, dass i. d. R. nur kurze Zeitabstände zwischen der Schulinspektion und der Wirkungserfassung liegen und sowohl Schulentwicklung als auch verbesserte Schülerleistungen sich erst nach mehreren Jahren oder Jahrzehnten zeigen.

Resümieren lässt sich bezüglich der Schulinspektionsforschung in Deutschland, dass insbesondere die ersten Studien von unmittelbar beteiligten Bildungsbehörden und -instituten initiiert wurden. Methodisch basieren diese i. d. R. auf Schulleitungs- und Lehrkräftebefragungen, welche angesichts der Beteiligung der Bildungsbehörden anfällig für Verzerrungen der sozialen Erwünschtheit und der Positivdarstellung sind. Des Weiteren ist die Forschung weitestgehend auf die innerschulische Perspektive beschränkt und hat die Systemebene noch nicht hinreichend berücksichtigt. Dabei steht die Entwicklungsfunktion der Schulinspektion im Vordergrund und die weiteren Funktionen wie Rechenschaftslegung oder Wissensgewinnung wurden noch unzureichend beleuchtet. Aktuell lassen sich auch von nicht unmittelbar in die Schulinspektion involvierten, unabhängigen Institutionen Forschungsbemühungen erkennen. Von den zu Beginn des Kapitels genannten, aktuell laufenden Großprojekten zu Schulinspektionen kann erhofft werden, dass sie einen Beitrag dazu leisten, diese aufgezeigten Forschungslücken in den nächsten Jahren zu schließen.

Als methodische Desiderata zukünftiger Forschungsvorhaben fordern Dederling und Müller (2008) die multiperspektivische Einbeziehung aller beteiligten Akteure, den Einsatz von Kontrollgruppen- und quasi-experimentellen Designs, die systematische Berücksichtigung von Schulen mit unterschiedlichen Qualitätsständen sowie schließlich die Unterscheidung von kurz- und langfristigen Effekten. Die systematische Sichtung des Forschungsstands auf nationaler Ebene lässt zudem Fragen der Validität und der Effektivität der Schulinspektion erkennen. Insgesamt bleibt festzuhalten, dass durch die Schulinspektion kontinuierlich schulische Daten erzeugt werden, die in viel größerem Umfang und einer viel größeren Intensität für Forschungszwecke genutzt werden könnten.

Abschließend gilt es auch zu reflektieren, ob die Erwartungshaltung an die Schulinspektion, merkliche Fortschritte in der Schulentwicklung und eine messbare Verbesserung von Schülerleistungen herbeizuführen, überhaupt realistisch ist. Vielmehr deuten die Erwartungen und Hoffnungen auf eine Überbewertung der Wirksamkeit und Wirkungen von Schulinspektion hin. Diesen Gedanken aufgreifend, sehen auch Böttcher und Keune (2011) eine Dysbalance und Disparität zwischen den Erwartungen, welche an die Schulinspektion gestellt werden und dem, was sie tatsächlich leisten kann:

Aber die Erwartung, dass Schulinspektion genau dies bewirken könnte, nämlich einen qualitativ bedeutenden Wandel erzeugen, dürfte überzogen sein. Es ist die Politik, die Schulinspektion mit dem Anspruch überlastet, Entwicklung zu initiieren oder gar zu steuern. Wenn Schulinspektion von dieser Erwartung befreit wird, kann sie fairerweise lediglich im Hinblick auf ihre Qualität als Diagnoseinstrument bewertet werden (Böttcher/Keune 2011, S. 139).

## 4. Rezeption und Nutzung von Ergebnissen der Schulinspektion

An dieser Stelle wird von dem ersten zentralen theoretischen Thema der Arbeit *Schulinspektion* zum zweiten Hauptthema *Rezeption und Nutzung von (Schulinspektions)Ergebnissen* übergeleitet. Wenn in dieser Arbeit die Rede von der *Nutzung* der Ergebnisse ist, gilt dieser Begriff als Oberbegriff von Rezeption, Reflexion und Aktion. Auf diese Trias wird im folgenden Theorieteil eingegangen. Dabei geht es nicht nur um bereits umgesetzte, d. h. sichtbare Maßnahmen, sondern um den Prozess der Verarbeitung der Ergebnisse, der auch ihre verbindliche Planung, z. B. festgehalten in Zielvereinbarungen, beinhaltet.

Ausgehend von den vorangegangenen Ausführungen über die generellen Zielsetzungen der Schulinspektion einerseits und ihren nur moderaten Wirkungen andererseits stellt sich die Frage, ob und wie die Schulinspektionsergebnisse für Schulentwicklung genutzt werden. Gerade im Zuge der rasanten Implementierung der Schulinspektion in Deutschland und der beachtlichen Ansammlung zurückgemeldeter Ergebnisse ist die Kenntnis ihrer Nutzung und Wirkung ein logisches Desiderat. Ergebnisrückmeldungen, welche den Schulen im Anschluss an Schulinspektionen gegeben werden, stellen noch keine Interventionen dar, sondern haben einen hohen Aufforderungscharakter und sind mit der Erwartung verknüpft, dass aus ihnen unmittelbare Konsequenzen für die Schulentwicklung abgeleitet werden. Es wird davon ausgegangen, dass die Ergebnisse der Schulinspektion in die zielorientierte und handlungsrelevante Planung von Maßnahmen der Schul- und Unterrichtsentwicklung eingehen. In einem vorgelagerten Schritt müssen die Informationen von den Schulen rezipiert, reflektiert und verarbeitet werden. In diesem Kontext spricht Altrichter von einem Konzept überzeugender Plausibilität ohne eine eindeutige Datenlage, dass Rückmeldungen auch tatsächlich zur Schul- und Unterrichtsentwicklung führen (vgl. Altrichter 2010, S. 231). Eine wissenschaftliche Bestätigung der Annahme einer direkten Ergebnisnutzung von Rückmeldungen wurde noch nicht erbracht (vgl. van Ackeren/Klemm 2009, S. 177; Schneewind/Kuper 2009, S. 125f.). Gerade die Konsequenzen stellen im Rahmen der Schulinspektion „wie bei jeder externen Evaluation einen neuralgischen Punkt dar“ (Bos u.a. 2006, S. 118). Zudem vergrößern Rückmeldungen laut Baumert (2001, S. 13) eher die

Komplexität der Veränderungsmöglichkeiten und -entscheidungen, als dass sie Entscheidungshilfen aufzeigen. Im Allgemeinen werden Evaluationsergebnisse häufig ignoriert und von den Verantwortlichen in der Praxis selten zur Entscheidungsfindung herangezogen. Dabei erfolgt die Reaktion auf Evaluationsergebnisse am wenigsten in jenen Fällen, in denen diese nicht erwartungskonform sind und sie werden am ehesten dann genutzt, wenn sie bereits vorhandenes Wissen bestätigen (vgl. Patton 2008).

Der Frage nach der Ergebnisnutzung im Kontext der Schulinspektion wird im weiteren Verlauf des Theorieteils auf zwei Ebenen begegnet. Zum einen wird ein weiterer Akteur – *die Schulaufsicht* – und ein zentrales Instrument im Anschluss an die Schulinspektion – *die Zielvereinbarung zwischen Schulen und Schulaufsicht* – skizziert. Dabei trifft die Schulaufsicht mit den Schulen in einem gemeinsamen Aushandlungsprozess Vereinbarungen über künftige Maßnahmen. Für diesen Prozess im Anschluss an die Schulinspektion wurde die Zielvereinbarung als entscheidende Schnittstelle zwischen den Ergebnissen der Schulinspektion und den darauf aufbauenden Maßnahmen zur Schulentwicklung sowie als zentrales Instrument zur Initiierung zusätzlicher Impulse konstituiert. Zum anderen erfolgt die Einbettung der Nutzung von Schulinspektionsergebnissen auf der Grundlage der *Rezeptionsforschung*. Die Rezeptionsforschung wird herangezogen, da sie die Nutzung und die Effekte von Ergebnissen aus Evaluationen wissenschaftlich beleuchtet.

Im Anschluss an die Darstellung der Erkenntnisse über die Nutzung von Evaluationsergebnissen im Allgemeinen und von Ergebnissen aus Schulinspektionen im Speziellen werden gezielte Forschungsfragen aufgeworfen, welchen im empirischen Teil der Arbeit nachgegangen wird.

## 4.1. Rolle der Schulaufsicht im Zusammenhang mit der Schulinspektion

Dieses Kapitel befasst sich mit der Struktur und den zentralen Aufgaben der Schulaufsicht insbesondere im Kontext der Schulinspektion. Die grundsätzlichen Funktionen der Schulaufsicht leiten sich aus Art. 7 Abs. 1 des Grundgesetzes der Bundesrepublik Deutschland ab: „Das gesamte Schulwesen steht unter der Aufsicht des Staates“. Dieser Grundsatz spiegelt sich in allen Länderverfassungen wider, so auch in NRW (s. Verfassung des Landes NRW Art. 8 Abs. 2; s. ergänzend Schulgesetz NRW (§ 86 (1) SchulG).

Die Organisation der Schulaufsicht ist in den einzelnen Bundesländern unterschiedlich konzipiert, wobei in den meisten Bundesländern ein zweistufiges Modell überwiegt (vgl. Schnell 2010, S. 198): Neben dem Ministerium wird die Schulaufsicht auf der unteren Ebene durch das Landesschulamt/die Oberschulämter resp. den staatlichen/regionalen Schulämtern ausgeübt (vgl. ebd.; Maritzen 2008, S. 88). In NRW wird die Schulaufsicht durch das MSW NRW als oberste Schulaufsicht, durch die Schulabteilungen der fünf Bezirksregierungen Arnsberg, Detmold, Düsseldorf, Köln und Münster als obere Schulaufsichtsbehörden und durch 53 Schulämter, die als untere Schulaufsichtsbehörden bei den Kreisen bzw. kreisfreien Städten angesiedelt sind, ausgeübt (s. Abb. 5).



**Abb. 5: Struktur der Schulaufsicht in Nordrhein-Westfalen**

Zum Aufgabenspektrum der Schulaufsicht gehört eine Vielzahl unterschiedlicher Rollen- und Aufgabenwahrnehmungen, welche häufig als *Rollenmix* bezeichnet werden (vgl. Rolff/Schmidt 2002, S. 12; Brüsemeister/Newiadomsky 2008). Insgesamt ist die Schulaufsicht zuständig für innere Schulangelegenheiten, welche die Fach-, Dienst- und Rechtsaufsicht über die Schulen umfassen (vgl. van Ackeren/Klemm 2009, S. 107; KMK 2010, S. 53f.; Weitzel 2002, S. 33, S. 323). Durch das breite Aufgabenspektrum und aufgrund ihrer Position auf der „intermediären“ Ebene (Kussau/Brüsemeister 2007, S. 82f.) zwischen der übergeordneten Instanz und den Schulen befindet sich die Schulaufsicht in einem „Pflichtenzwiespalt“ oder „inneren Spagat“ (Weitzel 2002, S. 333; vgl. auch Bürger 2007): Auf der einen Seite hat sie die Aufgabe, die Schulen zu unterstützen und zu beraten, auf der anderen Seite muss sie die Einhaltung staatlicher Vorgaben kontrollieren. Weitere Probleme der traditionellen Schulaufsicht werden in der fehlenden Profession durch die große Aufgabenbandbreite, der fehlenden Akzeptanz durch die Lehrkräfte und der Überlastung mit Verwaltungs- und Routinetätigkeiten gesehen (vgl. Burkard 2001, S. 51-66). Der größte Kritikpunkt besteht jedoch in der „Wirkungslosigkeit des schulaufsichtlichen Handelns in entscheidenden Qualitätsfragen“ (Maritzen 2008, S. 88).

Die Neuorientierung des Schulwesens, mit der eine Entwicklung zu mehr Dezentralisierung und Autonomie eingeleitet wurde (vgl. Zlatkin-Troitschanskaia 2008; s. auch Kapitel 2.2 zur Steuerung) führte auch zur intensiven Diskussion um die Rollen und Aufgaben der Schulaufsicht und es erfolgte eine Neubewertung der Schulaufsicht. Bei gleichzeitiger Aufwertung der Schulleitungen wurde die Schulaufsicht abgeschwächt (vgl. Gruschka 2010; Rosenbusch 2002, S. 16). Parallel zu dieser Veränderung wurde eine wachsende Ergebnis- und Kontrollorientierung im Schulbereich und eine damit einhergehende neue Aufwertung der Schulaufsicht prognostiziert (vgl. Zlatkin-Troitschanskaia 2008, S. 18; Stryck 2000). Zurzeit ist in allen Bundesländern folgender Trend zu beobachten: „Die personenbezogene Betreuung durch die Schulaufsicht [verliert] tendenziell an Gewicht gegenüber den Aufgaben der Unterstützung und Beratung von Schulentwicklung und schulischem Qualitätsmanagement“ (KMK 2010, S. 54). Als künftige bzw. z. T. bereits aktuelle Arbeitsbereiche der Schulaufsicht können vier Schwerpunkte benannt werden (vgl. Schnell 2010, S. 198f.):

1. Budget: Stellen, Controlling, empirische Daten
2. Reduzierung der bisherigen Funktionen der Dienst- und Fachaufsicht
3. Beratung und Unterstützung der Schulen, einschließlich der Qualifizierung und Fortbildung für alle Lehrkräfte in der Region
4. Kommunikations-, Kooperations- und Koordinationsmanagement

Dass das schulaufsichtliche Aufgabenspektrum nicht statisch organisiert ist, zeigt sich im historischen Rückblick (vgl. Dubs 2009; Kluchert/Leschinsky 2001; Schnell 2010, S. 197). Auch in der Vergangenheit hat sich die Schulaufsicht durch politische und wissenschaftliche Einflüsse und mit der Entwicklung des Schulwesens gewandelt. Dubs (2009, S. 511) systematisiert die Formen der Schulaufsicht im Verlauf der Zeit und gliedert sie in vier Phasen (s. Tab. 6).

**Tab. 6: Arten der Schulaufsicht nach Dubs**

Arten der Schulaufsicht	Merkmale	Einfluss
<b>Inspektion</b> (etwa bis 1920)	Bürokratische Kontrolle, welche die Einhaltung staatlicher Vorschriften und gesellschaftlicher Regeln überprüft	Autokratische Staatsauffassung
<b>Aufsicht</b> (etwa bis 1970)	Bürokratische Überwachung der Unterrichtsführung mit dem Ziel der Verbesserung der Lernerfolge	Wissenschaftliche Erkenntnisse zum Lehren und Lernen
<b>Unterstützung</b> (bis heute)	Überwachende und konsultative Zusammenarbeit zwischen Lehrkräften und Schulaufsicht mit dem Ziel der Qualitätsverbesserung der Schule	Human Relations-Bewegung, gruppensdynamische Prozesse, Demokratisierungstendenzen
<b>Controlling</b> (Zukunft?)	Förderung der Schulentwicklung durch Überwachung der Zielerreichung (Controlling) sowie durch Beratung	Geleitete, teilautonome (selbständige) Schule (New Public Management)

Dabei lässt sich eine Tendenz von eher kontrollierenden zu eher unterstützenden und beratenden Tätigkeitsfeldern der Schulaufsicht feststellen. Die von Dubs als „Controlling“ bezeichnete zentrale Aufgabe der Schulaufsicht, d. h. die Beratung auf der Grundlage von Beobachtungen und Daten, ist heute in vielen Bundesländern Realität und gehört im Zusammenhang mit der Zielvereinbarungspraxis zu den prioritären Aufgabenfeldern der Schulaufsicht. Mit dieser Lösung, d. h. der Einheit von Controlling und Beratung, sieht Dubs die Divergenz unterschiedlicher Schulaufsichtsrollen aufgehoben.

Im Folgenden wird auf die Rolle der Schulaufsicht im Zusammenhang mit der Schulinspektion eingegangen. Bundesweit zog die Einführung der Schulinspektion „Verschiebungen im Gesamtsystem der Steuerung des Schulwesens“ (Maritzen 2008, S. 88) nach sich. Dabei wird die Schulinspektion als eine Reaktion auf eine Krise der Schulaufsicht und der alt hergebrachten Verfahren der Wissensgewinnung interpretiert, mit der ein Wandel „vom Augenschein zur Empirie, vom Vor-Gewussten zum Nachgewiesenen, vom Bescheidwissen zum Draufschauen, von kasuistischen Vorgehensweisen zu standardisierten Prozessen“ (Maritzen 2008, S. 93) vollzogen wird. Damit „erscheint Schulinspektion in Deutschland als eine Ausdifferenzierung, Professionalisierung und Verwissenschaftlichung einer Teilaufgabe [der Schulevaluation] des bisherigen Spektrums schulaufsichtlicher Tätigkeiten“ (Rürup 2008, S. 470).

Dass das neue Verhältnis von Schulaufsicht und Schulinspektion Unsicherheiten birgt, wird durch die Verwendung der unterschiedlichen Konjunktionen *und*, *oder*, *versus* und *als* zwischen Schulinspektion und Schulaufsicht deutlich. Hiermit wird auf die Überschrift eines Beitrags von Füssel (2008, S. 153) Bezug genommen. Als Begründung für diese Unsicherheit können Unterschiede auf Länderebene in der Etablierung der Schulinspektion, der Rekrutierung des Personals sowie der Rolle der Schulaufsicht im Verfahren der Schulinspektion angeführt werden. Im Vorfeld der Implementierung der Schulinspektion stimmten die Überlegungen der Bundesländer darin überein, keine bereits vorhandenen Einrichtungen zu nutzen, sondern neue Institutionen mit neuem Personal zu errichten (vgl. ebd., S. 154). Diese Überlegungen wurden allerdings in der Praxis nicht umgesetzt.

Das Verhältnis der Schulaufsicht zur Schulinspektion stellt sich exemplarisch für das Bundesland NRW wie folgt dar: Die Mitarbeiter der Qualitätsanalyse – die Qualitätsprüfer – sind mit einem eigenen Dezernat (Dezernat 4Q) in der oberen Schulaufsichtsbehörde tätig und somit eindeutig Mitarbeiter der Schulaufsicht. Ihre Aufgaben werden in demselben Paragraphen des Schulgesetzes, in dem die Aufgaben der Schulaufsicht geregelt sind, beschrieben (SchulG § 86 Schulaufsicht, Abs. 5 Qualitätsanalyse). Die Qualitätsprüfer haben ausschließlich Aufgaben der Qualitätsanalyse wahrzunehmen, während ihnen weitergehende schulaufsichtliche Aufgaben und Befugnisse nicht übertragen werden. Insbesondere ist zu betonen, dass es nicht zu den Aufgaben der

Qualitätsprüfer gehört, die Schulen zu beraten, da dies ausschließlich in den Aufgabenbereich der zuständigen Schulaufsichtsbehörden fällt.

Auf Ebene der einzelnen Bundesländer herrschen unterschiedliche institutionelle und organisatorische Anbindungsformen der Schulinspektion zur Schulaufsicht und z. T. gehört die in NRW zu unterlassende Beratungsfunktion der Qualitätsprüfer in anderen Ländern ausdrücklich zur Aufgabe der Schulinspektoren. Tabelle 7 enthält eine im Jahr 2012 aktualisierte Übersicht von Döbert, Rürup und Dederich (2008, S. 142-144) zu den Aufgaben der zuständigen Schulaufsicht im Zusammenhang mit der Schulinspektion in den Bundesländern. Neben den Unterschieden zeigen sich die bundesweiten Gemeinsamkeiten eher im Umgang mit den Inspektionsergebnissen. Die Länderregelungen sehen dabei vor, dass die Ergebnisse der Schulinspektionen an die unteren Schulaufsichtsbehörden zu übermitteln sind. Im Anschluss an die Ergebnismeldung reichen die Aufgaben der Schulaufsicht von der passiven Entgegennahme der Inspektionsergebnisse bis zur aktiven Mitwirkung an der schulinternen Ergebnisauswertung und der konkreten Planung von Handlungsabsichten. Im Rahmen des Prozesses von den Rückmeldungen zu konkreten Handlungen stellt der Abschluss von Zielvereinbarungen zwischen Schulen und Schulaufsicht ein wichtiges Steuerungsinstrument dar, mit dem die Schulaufsicht betraut wird.

Zusammenfassend wird festgehalten, dass die Schulinspektion Teil der Schulaufsicht ist, einen bestimmten Teil ihrer Aufgaben übernimmt und standardisiert ausführt. Diese Rollenteilung erfordert von beiden Akteurengruppen unterschiedliche professionelle Aufgaben – einerseits die Erhebung von Evaluationsdaten und andererseits die Beratung und Begleitung der Schulen. Damit stehen sie organisationstheoretisch in einem Interdependenzverhältnis zueinander. D. h. die Schulinspektion ist einerseits auf die institutionelle Anbindung an die Schulaufsicht angewiesen und hat andererseits die Gewissheit über die Abnahme und Weiterverarbeitung der Ergebnisse. Ohne eine Nachsorge erschiene ihre Arbeit wirkungslos. Die Schulaufsicht wiederum benötigt datengestützte Informationen über die schulische Arbeit zur Beratung und bedarfsgerechten Unterstützung der Schulen. Für die Entfaltung der Wirkung der Schulinspektion spielt die Kooperation beider Systeme sowie ein gemeinsames Verständnis der Inspektionsergebnisse eine entscheidende Rolle.

**Tab. 7: Aufgaben der zuständigen Schulaufsicht im Zusammenhang mit der Schulinspektion in den Bundesländern**

<b>Bundesland</b>	<b>Rolle der zuständigen Schulaufsicht</b>
Baden-Württemberg	Abstimmung und Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der Fremdevaluation; Unterstützung der Schulen im Prozess der Ergebnisauswertung und Maßnahmenumsetzung
Bayern	Auswahl der Evaluatoren; Organisation des Einsatzes der Evaluationsteams; Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation; Unterstützung der Schulen im Prozess der Ergebnisauswertung und Maßnahmenumsetzung; fortlaufende Beratung, Begleitung und Unterstützung
Berlin	Besprechung des Inspektionsberichts mit den Schulen; Abschluss von Zielvereinbarungen über die Weiterentwicklung der Schulqualität
Brandenburg	Auswertung der Schulvisitation und Abschluss von Zielvereinbarungen
Bremen	Auswertungsgespräch mit den Schulen; Klärung des Unterstützungsbedarfs; Abschluss von Zielvereinbarungen mit den Schulen nach der externen Evaluation
Hamburg	Besprechung der Inspektionsergebnisse und Abschluss von Ziel- und Leistungsvereinbarungen (ZLV) mit den Schulen
Hessen	Beteiligung an der Auswertungskonferenz; Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation; Begleitung der Schulen bei der Umsetzung der gemeinsamen Ziel- und Leistungsvereinbarungen
Mecklenburg-Vorpommern	Auswahl der Schulen; Abschluss von Zielvereinbarungen mit den Schulen
Niedersachsen	Teilnahme an der mündlichen Rückmeldung der Inspektionsergebnisse; Anregung schulischer Entwicklungsprozesse; Abschluss von Zielvereinbarungen mit den Schulen; Begleitung der Schulen bei der Umsetzung der Zielvereinbarungen
Nordrhein-Westfalen	Analyse des Qualitätsberichts; Beratung der Schulen bei der Maßnahmenentwicklung; Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der Qualitätsanalyse; Controlling; Bilanzgespräche
Rheinland-Pfalz	Teilnahme an der ErgebnISRückmeldung in den Schulen; Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation
Saarland	Information über die Evaluationsergebnisse der Schulen: Teilnahme am Gespräch über die Ergebnisse des Pädagogischen Tages (ggf. Kontaktaufnahme mit den Schulen); (Schulen treffen schulinterne Zielvereinbarungen auf der Grundlage der Empfehlungen der Abschlussberichte)
Sachsen	Teilnahme an der ErgebnISRückmeldung in den Schulen (auf Einladung der Schulen); Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation; Beratung und Unterstützung der Schulen im Prozess der Ergebnisauswertung und Maßnahmenumsetzung
Sachsen-Anhalt	Unterstützung der Schulen im Prozess der Ergebnisauswertung und Maßnahmenumsetzung; Abschluss von Zielvereinbarungen mit den Schulen
Schleswig-Holstein	Geschäftsführende Leitung des Verfahrens; Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation (seit dem Schuljahr 2010/11 ist EVIT (Evaluation im Team) abgeschafft)
Thüringen	Während der Evaluation: Parallele Bearbeitung des Schulfragebogens; Nachträgliche Begleitung der schulischen Veränderungsprozesse der Evaluationsbefunde; ggf. Teilnahme an der Rückmeldeveranstaltung (bei Wiederholungsbesuchen Pflicht); Abschluss von Zielvereinbarungen mit den Schulen nach Beendigung der externen Evaluation; Kontrolle der Umsetzung dieser Zielvereinbarungen

## **4.2. Zielvereinbarungen im Anschluss an die Schulinspektion**

Im vorangegangenen Kapitel wurde darauf hingewiesen, dass das Treffen von Zielvereinbarungen mit den Schulen als neue Form beruflichen Handelns in den veränderten Aufgabenbereich der Schulaufsicht fällt. Die Arbeit mit dem Instrument Zielvereinbarung spiegelt durch die erforderliche Kooperation und Konsensorientierung die neue Beziehung zwischen Schulen und Schulaufsicht wider (vgl. Thym 2009, S. 286). Zielvereinbarungen sind in den Kreis der neuen Steuerungsinstrumente im Bildungsbereich einzuordnen, da sie anstatt durch input- und prozessorientierte Detailsteuerung nun ergebnisorientiert und strategisch „auf Abstand“ steuern (vgl. Krems 2012). Im Zusammenhang mit der Schulinspektion nehmen Zielvereinbarungen durch ihre Brückenfunktion zwischen Ergebnisrückmeldung und -nutzung eine zentrale Funktion ein. Die Nutzung der Inspektionsergebnisse und ihre Überführung in konkrete Maßnahmen ist nicht zuletzt von den Potentialen der Zielvereinbarung als Steuerungsinstrument und dem Zielvereinbarungsprozess determiniert.

Bevor im Folgenden der konkrete Zielvereinbarungsprozess zwischen Schulen und Schulaufsicht am Beispiel der Qualitätsanalyse NRW beschrieben wird, erfolgt eine Darstellung der theoretischen Grundlagen und wissenschaftlichen Erkenntnisse zu Zielvereinbarungen. Dabei wird das Instrument Zielvereinbarung zunächst aus zwei Perspektiven betrachtet: Einerseits erfolgt die Beleuchtung der Zielvereinbarungen aus ökonomischer Perspektive, da ihre historischen Wurzeln im Wirtschaftsbereich liegen, andererseits werden die Funktionen von Zielen und Zielmerkmalen auf der Grundlage der Motivations- sowie Arbeits- und Organisationstheorie erörtert. Daran schließt eine Darstellung der Zielvereinbarungen im Bildungsbereich an und wird nochmals vor dem Hintergrund der Schulinspektion spezifiziert.

### **4.2.1. Definition von Ziel und Zielvereinbarung**

Der Klärung des Zielvereinbarungsbegriffs wird die Definition des Begriffs „Ziel“ vorangestellt: Ein Ziel ist als „ein in der Zukunft liegender, angestrebter Zustand, der eindeutig beschrieben ist (...) [und als] vorausgedachte[s] Ergebnis einer Handlung“ (Buhren/Rolff 2009, S. 68) zu definieren. Ziele sind

i. d. R. nicht a priori festgelegt und werden aus einer Vielzahl an Zielmöglichkeiten ausgewählt (vgl. Kleinbeck 2006, S. 256). Sie beschreiben einen erwünschten Soll-Zustand in der Zukunft, der durch bestimmte Handlungen erreicht werden soll, ohne den Weg zur Zielerreichung vorzugeben (vgl. Breisig 2006, S. 30; Krems 2012). Diesem Kerngedanken folgend haben Ziele eine wichtige Bedeutung für menschliche Handlungen:

Ziele bilden die Dreh- und Angelpunkte bei der psychischen Steuerung des menschlichen Handelns. Sie stehen im Zentrum theoretischer Konzepte (...), um herauszufinden, warum Menschen sich zu einem bestimmten Zeitpunkt für die Aufnahme, Ausführung und Beendigung einer bestimmten Handlung entscheiden. Ohne Ziele sind Handlungen undenkbar (Kleinbeck 2006, S. 255).

Den Zielbegriff zugrunde legend, stellen Zielvereinbarungen eine i. d. R. zwischen zwei hierarchischen Ebenen verbindliche Absprache und gemeinsame Festlegung von Zielen dar. Durch ihren partizipativen und durch Transparenz und Gleichberechtigung der Beteiligten gekennzeichneten Charakter unterscheiden sich Ziele von Zielvorgaben, bei welchen die alleinige Zielentscheidungssouveränität von den übergeordneten Hierarchieebenen ausgeht. Grundsätzlich können Zielvereinbarungen nicht nur von Einzelpersonen, sondern auch von organisatorischen Einheiten getroffen werden. In Zielvereinbarungen wird anhand der Kurzformel *Wer, Was, Wann, Womit* über die zu erreichenden Ergebnisse und Wirkungen (Outcome) und die zu erbringenden Leistungen (Output) eine Übereinkunft getroffen (vgl. Krems 2012). Der Zielvereinbarungsprozess endet erst nach Ablauf der festgelegten Terminierung durch einen Soll-Ist-Abgleich und der gemeinsamen Überprüfung der Zielerreichung. Entsprechend ihrem breiten Einsatzspektrum können Zielvereinbarungen Funktionen der Personalführung, Personalbeurteilung, Entgeltpolitik und leistungsbezogenen Entlohnung erfüllen (vgl. Breisig 2006, S. 34f.). Die historischen Wurzeln der Zielvereinbarungen liegen im Managementkonzept „Management by Objectives“ (MbO; Führen durch Ziele) und in der motivations- sowie arbeits- und organisationspsychologischen Forschung und werden im Folgenden präsentiert (vgl. Bungard/Kohnke 2000).

### **4.2.2. Die ökonomische Perspektive: Management by Objectives**

Der ursprünglich in den USA von Peter Drucker 1954 entwickelte Gedanke zu Zielvereinbarungen war durch seinen anwendungsorientierten Einsatz als Personalführungsinstrument gekennzeichnet, welcher wirtschaftliche Gewinne herbeiführen sollte. Entsprechend wurden Ziele in MbO-orientierten Unternehmen von der Unternehmensführung abgestuft auf die unteren Hierarchieebenen heruntergebrochen (vgl. Schmidt/Kleinbeck 2006, S. 6). Diese top-down-orientierte Managementstrategie fand in der Nachkriegszeit enormen Zuspruch, geriet aber in den 1970er/80er Jahren zunehmend in die Kritik (vgl. Bungard 2000, S. 17). Mit der neuen Funktionszuweisung als unternehmensstrategische Controlling-Instrumente wurden die Zielvereinbarungen in der Wirtschaft in den 1990er Jahren wiederentdeckt und sind seitdem weit verbreitet (vgl. Hölzle 2000). Aktuell finden sie vorrangig als Organisationsführungs- und Kontrollinstrumente dezentraler Strukturen Verwendung, während ihre ursprüngliche Funktion als Personalführungsinstrument in den Hintergrund geraten ist (vgl. Breisig 2006, S. 34f.; Kohnke 2000).

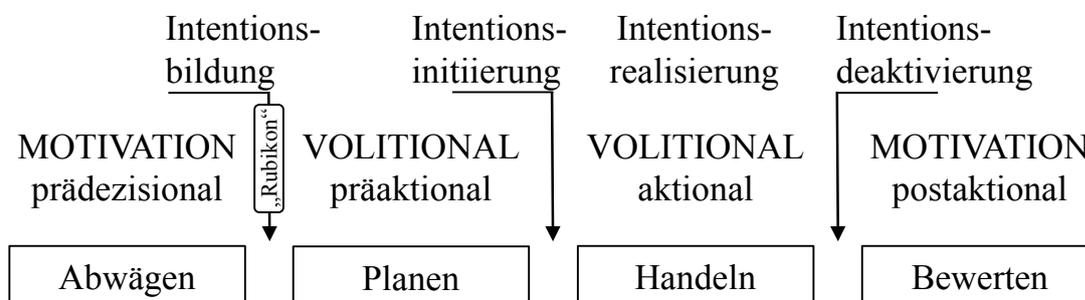
### **4.2.3. Die psychologische Perspektive: Rubikon-Modell und Zielsetzungstheorie**

Parallel zur Entdeckung und Anwendung in der Wirtschaft wurden Ziele auch Gegenstand eines regen Forschungsinteresses der Psychologie. Das Setzen von Zielen und ihre Überführung in Handlungen stellt einen komplexen Prozess dar, der in den Forschungsbereich der Motivationspsychologie fällt. Diese gilt wiederum als zentraler Bestandteil der Arbeits- und Organisationspsychologie. Die folgenden Ausführungen skizzieren die Grundlagen der Motivationspsychologie in ihren wesentlichen Aspekten, insbesondere dem Rubikon-Modell der Handlungsphasen und der Zielsetzungstheorie. Auf dieser Grundlage werden modell- und theoriegeleitete Anhaltspunkte bezüglich des Prozesses, der Funktion und der Gelingensbedingungen von Zielvereinbarungen im Bildungsbereich – insbesondere im Anschluss an die Schulinspektion – abgeleitet.

Laut Rheinberg (2006, S. 16) bezeichnet Motivation „die aktivierende Ausrichtung des momentanen Lebensvollzugs auf einen positiv bewerteten Zielzu-

stand“, wobei mit dem Begriff „Lebensvollzug“ das aktive Handeln gemeint ist (vgl. dazu auch Achtziger/Gollwitzer 2006, S. 281). Nach Heckhausen und Heckhausen (2006, S. 1-3) sind die beiden universellen Charakteristika motivierten Handelns das Streben nach Wirksamkeit und die Organisation von Zielengagement sowie Zieldistanzierung. Mit dem Streben nach Wirksamkeit ist der Wunsch nach einer direkten oder primären Kontrolle der physischen und sozialen Umwelt gemeint. Zielengagement und Zieldistanzierung sind gleichsam bedeutend, da sie dazu beitragen, dass ein Ziel erreicht oder ein unerreichtes Ziel aufgegeben werden kann (vgl. ebd., S. 2). Durch das bewusste Vorgehen grenzt sich damit motiviertes Verhalten von Automatismen wie Reflexen und Routinen ab (vgl. Achtziger/Gollwitzer 2006, S. 277).

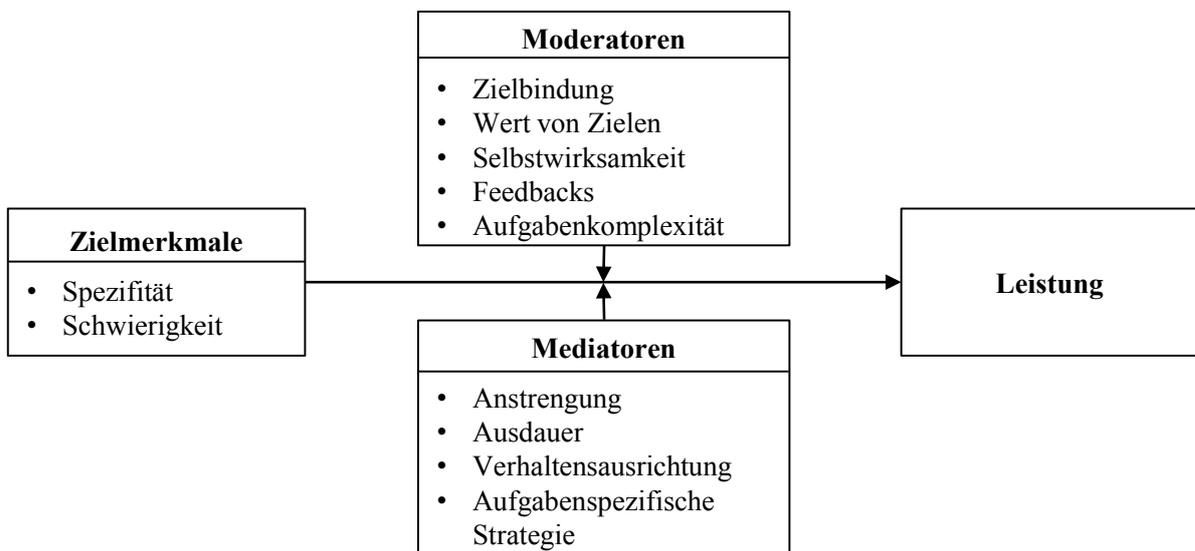
Die Motivationspsychologie befasst sich mit der Untersuchung der Richtung, Ausdauer und Intensität zielgerichteten Verhaltens unter Berücksichtigung personen- und situationsbezogener Faktoren (vgl. Heckhausen/Heckhausen 2006, S. 3; Rheinberg 2006, S. 16). Zur Erklärung der Zielsetzung und der Realisierung von gesetzten Zielen wurde das Rubikon-Modell der Handlungsphasen entwickelt (vgl. Heckhausen/Gollwitzer 1987). Das innovative Element des Modells liegt in der Zusammenführung und in den Übergängen zwischen motivationalen und volitionalen Handlungsphasen. Volitionen sind „eigene regulative Prozesse, die entscheiden, welche Motivationstendenzen (...) realisiert werden sollen“ (Heckhausen/Heckhausen 2006, S. 7). Insgesamt beschreibt das Modell einen Handlungsverlauf zielorientierten Verhaltens in einem zeitlichen Prozess und differenziert vier eigenständige Phänomene zielorientierten Verhaltens: Abwägen, Planen, Handeln und Bewerten (s. Abb. 6).



**Abb. 6: Das Rubikon-Modell der Handlungsphasen**

Diese Phänomene werden jeweils in den prädezisionalen, präaktionalen, aktionalen und postaktionalen Handlungsphasen umgesetzt (s. Achtziger/Gollwitzer 2006, S. 278-281). In der prädezisionalen Handlungsphase werden Handlungsoptionen mitsamt ihren potentiellen Konsequenzen eingeschätzt. Das Abwägungsergebnis mündet in Zielintentionen, d. h. erwünschten Endzuständen von Handlungen. Mit der Bildung einer verbindlichen Intention und der empfundenen Verpflichtung, das Ziel auch zu erreichen, wird metaphorisch der *Rubikon* überschritten und ein Wunsch in ein verbindliches Ziel umgewandelt. In der nächsten Phase, der präaktionalen Phase, werden konkrete Strategien geplant, die zu einer Realisierung des verbindlich festgelegten Ziels führen. Anschließend erfolgt in der aktionalen Phase die Handlungsinitiierung, d. h. die Realisierung der zuvor geplanten Ziele. Schließlich wird in der postaktionalen Handlungsphase das Handlungsergebnis bewertet (vgl. ebd., S. 278-281). Insgesamt trägt das Rubikon-Modell der Handlungsphasen dazu bei, den Ablauf zielorientierten Verhaltens differenziert und sequentiell darzustellen.

Im Rahmen der psychologischen Zielsetzungstheorie (Goal-Setting-Theory) wurden bestimmte Zielmerkmale identifiziert, welche die Zielrealisation beeinflussen.



**Abb. 7: Zielsetzungstheorie**

Die Zielsetzungstheorie wird seit mittlerweile ca. 40 Jahren intensiv beforscht und zählt inzwischen zu einer der am besten abgesicherten Ansätze der Psychologie (vgl. Lee/Locke/Latham 1989; Locke/Latham 2002). Als ihre zentrale Forschungserkenntnis gilt, dass schwierige und spezifische Ziele zu Leistungssteigerungen führen (vgl. Lee u.a. 1989). Die wesentlichen Bestandteile der Zielsetzungstheorie sind in einer verkürzten Darstellung des *High Performance Cycle* veranschaulicht (s. Abb. 7). Ausgangspunkt des Modells sind Zielmerkmale, die nach Spezifität und Schwierigkeit unterteilt sind. Die Umsetzung von Zielen in Leistung erfolgt unter Berücksichtigung von Mediatoren und Moderatoren. Im Modell sind vier Mediatoren aufgeführt, über die Ziele direkt auf die Leistung bzw. Leistungssteigerung wirken: Verhaltensausrichtung, Anstrengung, Ausdauer und aufgabenspezifische Strategien. Daneben existieren folgende Moderatoren, die den Effekt zwischen Zielmerkmalen und Leistung beeinflussen: Zielbindung, Wert von Zielen, Selbstwirksamkeit, Feedback und Aufgabenkomplexität (vgl. Locke/Latham 2002).

#### 4.2.4. Zielmerkmale

Als wesentlicher Verdienst der motivations- sowie arbeits- und organisationspsychologischen Forschung gilt die Identifikation einer Reihe von Dimensionen, durch welche sich Ziele voneinander unterscheiden lassen und welche das Handeln unterschiedlich beeinflussen. Die zentralen Dimensionen wurden bereits im Zusammenhang mit der Zielsetzungstheorie benannt und werden im Folgenden erläutert. Ziele sind durch folgende beiden Merkmale gekennzeichnet: *Inhalt* und *Intensität*. Kennzeichnend für das Zielmerkmal *Intensität* ist insbesondere die Zielbindung (*Commitment*), d. h. das Ausmaß der Identifikation und der gefühlten Verpflichtung, das Ziel auch bei möglichen Schwierigkeiten und Rückschlägen zu erreichen. Zu der Kategorie *Zielinhalt* lassen sich Aspekte der Schwierigkeit (*goal difficulty*) und Spezifität (*goal specificity*) zuordnen. Bezüglich der Bestimmung der subjektiven und objektiven Schwierigkeitsgrade von Zielen können Referenzmaße vorheriger Leistungen einer Person oder einer Gruppe herangezogen werden (vgl. Kohnke 2000, S. 39). Forschungsergebnisse belegen einen positiven Zusammenhang zwischen der Zielschwierigkeit und der Attraktivität des Ziels (vgl. Locke/Latham 2002). Die Leistung stagniert allerdings bei zunehmender Schwierigkeit, wenn die Leis-

tungsfähigkeitsgrenze erreicht wird. Damit hat der positive Zusammenhang zwischen Zielschwierigkeit und Leistung nur bei einfachen und ergebnisbezogenen Zielen, nicht jedoch bei komplexen Aufgaben und Lernzielen Bestand.

Als zweites Zielinhaltsmerkmal wird Spezifität definiert als „das Ausmaß, in dem das Ziel bzgl. Zielinhalt, Zielausmaß und zeitlichem Bezug beschrieben ist“ (Kohnke 2000, S. 40). Dabei können Ziele bezüglich ihrer zeitlichen Nähe oder Ferne sowie ihrer Komplexität variieren und qualitativer oder quantitativer Art sein. Die Effektgrößen für den Zusammenhang zwischen spezifischen Zielen und einer Leistungssteigerung liegen in einer Metaanalyse zwischen  $d = ,42$  und  $d = ,80$  (vgl. Locke/Latham 2002, S. 706). Als Ursache für die Variabilität der Effektstärken können unterschiedliche beeinflussende Faktoren der Leistungswirkungen verantwortlich gemacht werden (s. Befunde der Zielsetzungstheorie und Abb. 7). Im Gegensatz zu vagen oder abstrakten Zielen liegen die Vorteile spezifischer Ziele auch in der Operationalisierung und der Möglichkeit eines Soll-Ist-Abgleichs (vgl. Kohnke 2000, S. 40).

Zur Formulierung von Zielen und Zielvereinbarungen wird die sog. SMART-Regel herangezogen, welche eher als eine Orientierung und weniger als eine allgemeingültige Regel fungiert. Insgesamt vermittelt die SMART-Regel den Appell, dass die Zielsetzung und -überprüfung aktiv gesteuert werden kann. Entsprechend sollten Ziele unter Berücksichtigung der in Tab. 8 aufgeführten Merkmale formuliert werden (in Anlehnung an Schmidt/Kleinbeck 2006, S. 41). Die Dechiffrierung der Akronyme „smarter“ Ziele in Tab. 8 macht deutlich, dass entscheidende Zielmerkmale der Zielsetzungstheorie aufgegriffen werden. Durch die Berücksichtigung „smarter“ Ziele ergeben sich insbesondere Möglichkeiten der Messung und Überprüfung der Zielerreichung und sie fungieren als Mediatoren zwischen Zielsetzung und Leistung.

**Tab. 8: SMART-Regel für Ziele**

<b>S</b>	spezifisch	genaue Beschreibung des angestrebten Leistungsergebnisses in leicht verständlicher Form; Vermeidung von Standardfloskeln
<b>M</b>	messbar	Bereitstellung von überprüfbaren Messwerten bzw. Parametern
<b>A</b>	anspruchsvoll (attraktiv)	angemessen fördernde und fordernde Ziele wählen
<b>R</b>	realistisch	herausfordernde, aber erreichbare Zielschwierigkeit
<b>T</b>	terminiert	Festlegung eines für die Zielerreichung verbindlichen Termins

#### 4.2.5. Zielvereinbarungen im Bildungsbereich

Im Zusammenhang mit der Neuen Steuerung kann im Bildungsbereich eine zunehmende Verbreitung von Zielvereinbarungen, welche auch als Verträge oder Kontrakte bezeichnet werden, beobachtet werden (vgl. Wottawa 2001; Mittag 2006). Im schulischen Bereich hat sich die „neue Kultur des Vereinbarens“ (Füssel/Kretschmann 2005, S. 60) in Form von Schul-, Erziehungs- und Lernverträgen auf unterschiedlichen Ebenen etabliert: Sie können zwischen Ministerium und Schulaufsicht, zwischen Schulaufsicht und Einzelschule, zwischen Schulleitung und Lehrkraft/Lehrkräften und zwischen Lehrkraft und Schüler/n verabredet werden (vgl. Brückner/Tarazona 2010, S. 101; Reiners-Woch 2002, S. 127). Sie lassen sich differenzieren in verwaltungsinterne Formen, d. h. Vereinbarungen zwischen der Schulaufsicht und einzelnen Schulen, und verwaltungsexterne Formen, gekennzeichnet durch Ziel- oder Bildungsvereinbarungen zwischen Schulen und einzelnen Schülern oder Eltern (vgl. Thym 2009, S. 280; Füssel/Kretschmann 2005, S. 61-66; Buhren/Rolff 2009, S. 70). Bei den verwaltungsexternen Formen steht meist die Vereinbarung von Verhaltensanforderungen und -standards zwischen den schulischen Akteuren, insbesondere Schülern, Eltern und Lehrkräften, im Vordergrund. Diese bereits existierenden Vereinbarungsformen haben durch die Schulinspektion mit dem Instrument *Zielvereinbarung* eine Erweiterung erfahren: In diesem Kontext können Zielvereinbarungen datengestützt auf der Grundlage extern gemessener schulischer Qualitätsmerkmale getroffen werden.

Zur Bewertung von Zielvereinbarungen im Bildungsbereich sowie ihrer Chancen und Grenzen sind besondere Kennzeichen der Vertragsform hervorzuheben. Zielvereinbarungen sind durch einen vermeintlichen Gegensatz gekennzeichnet: Zum einen finden innerhalb eines offiziellen Vorgangs Verhandlungen und Einigungen zwischen den Vertragspartnern statt. Zum anderen stellen sie informelle Verträge dar, da sie weder juristisch einklagbar noch rechtlich bindend sind (vgl. Füssel/Kretschmann 2006, S. 58). Sie gelten als eher weiche und unselbständige Verträge, indem sie als Ergänzung bzw. „Scharnier zwischen den rechtlich greifbaren Einzelinstrumenten der Schulaufsicht“ (Thym 2009, S. 282) fungieren. Dies bedeutet aber nicht, dass Zielvereinbarungen durch die fehlende juristische Absicherung a priori dazu verurteilt sind,

wirkungslos zu sein. Vielmehr tragen „die moralische Bindung und der Wille der Gemeinsamkeit“ (Füssel/Kretschmann 2005, S. 58) sowie die Gründung von sog. „Verantwortungsgemeinschaften“ (Brückner/Tarazona 2010, S. 102) zur bindenden Wirkung von Zielvereinbarungen bei. Die Auslegung der Informalität kann jedoch auch in eine andere Richtung gehen. Gerade diese und die fehlende juristische Durchsetzbarkeit können die Funktion von Zielvereinbarungen befördern (vgl. Thym 2009, S. 282): Dadurch, dass sich Zielvereinbarungen von allgemeingültigen Gesetzen und Richtlinien abgrenzen, fungieren sie als individuelles und „maßgeschneidertes“ Steuerungswerkzeug für die Einzelschulen (Füssel/Kretschmann 2005, S. 66).

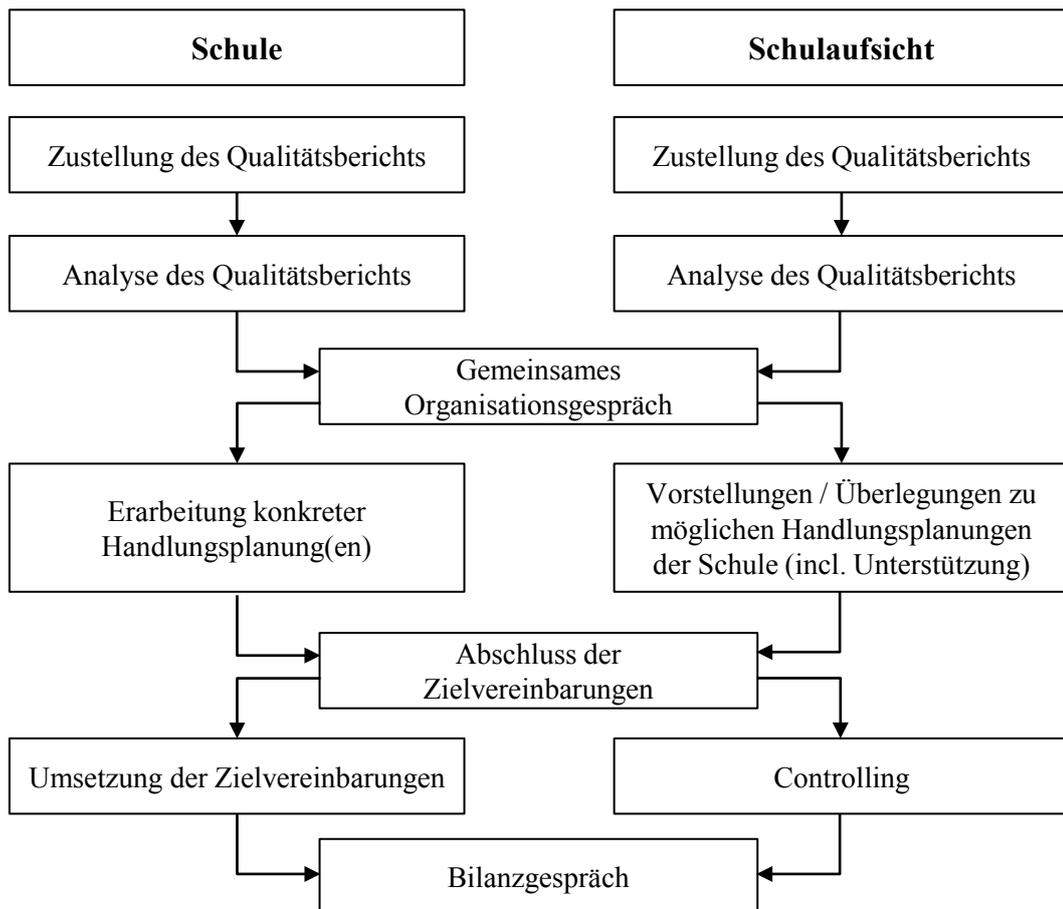
#### **4.2.6. Zielvereinbarungen nach der Schulinspektion – am Beispiel der Qualitätsanalyse NRW**

In Tabelle 7 (Aufgaben der Schulaufsicht in Zusammenhang mit der Schulinspektion) ist deutlich geworden, dass in sämtlichen Bundesländern nach der Berichtslegung als obligatorische Folge der Schulinspektion Zielvereinbarungen zwischen Schulen und Schulaufsicht getroffen werden (vgl. Döbert u.a. 2008, S. 144; Hovestadt 2009, S. 34-37). In diesem Zusammenhang schaffen Zielvereinbarungen als strategisches Steuerungsinstrument nach der Ergebnismeldung einen regelnden Rahmen für operative Handlungen (vgl. Brückner/Tarazona 2010, S. 101f.; Wendt/Klein 2007, S. 26). Dabei sind die Schulen zum Abschluss einer Zielvereinbarung verpflichtet, können jedoch über die Zielinhalte und die Umsetzung der Ziele frei entscheiden bzw. mitbestimmen. Der politisch-administrativen Ebene liefern Zielvereinbarungen Wissen darüber, welche Rückschlüsse die Schulen aus den Inspektionsergebnissen ziehen, welche Handlungsfelder sie als wesentlich erachten und wie sie darauf in Form von Handlungsabsichten reagieren. Damit dienen Zielvereinbarungen dazu, der Systemebene Steuerungswissen und Handlungsoptionen aufzuzeigen, die auf Ergebnissen der Schulinspektionen basieren. Die Herausforderung besteht für die politisch-administrative Ebene darin, die Zielvereinbarungen in einer Gesamtschau auszuwerten, Gemeinsamkeiten und Unterschiede herauszustellen und daraus Maßnahmen einzuleiten, bspw. durch die Bereitstellung spezifischer Fortbildungsangebote.

Im Folgenden wird die Funktion sowie der idealtypische Prozess von Zielvereinbarungen am Beispiel der Qualitätsanalyse NRW dargestellt (vgl. Homeier 2008; MSW NRW 2009a; Müller 2008). Sobald die Aufgaben der Schulinspektion mit der Berichtslegung enden, beginnt für die Schulen und die Schulaufsicht der sechsmonatige Prozess, der zum Abschluss einer Zielvereinbarung führt. Für das Zustandekommen von Zielvereinbarungen sind die Institutionen Schulaufsicht und Schulen unter Mitwirkung der Schulkonferenz verantwortlich. Dabei liegt bei den Schulen die Initiative über die Erarbeitung der Handlungsvorschläge, während die Schulaufsicht die Schulen dabei ggf. berät und unterstützt. Zur landesweit einheitlichen Vorgehensweise des Zielvereinbarungsprozesses im Anschluss an die Qualitätsanalyse hat die oberste Schulaufsichtsbehörde in der Handreichung zum institutionellen Zielvereinbarungsprozess zwischen Schulen und Schulaufsicht in NRW eine Formularvorlage bereitgestellt (s. MSW NRW 2011a). Diese gibt die Struktur vor, die bei der Verschriftlichung einer Zielvereinbarung zu berücksichtigen ist und schafft einen standardisierten und verbindlichen Rahmen. Den Kern der Zielvereinbarung stellt die Beschreibung von Zielen und Teilzielen dar. Die inhaltliche Zuordnung der Ziele soll am Qualitätstableau vorgenommen werden. Gefordert wird zudem die Benennung der Indikatoren und Maßnahmen mitsamt der Angabe des Datums der Zielerreichung und des Bilanzgesprächs.

Zur Beschreibung des idealtypischen Prozesses für die Entwicklung einer Zielvereinbarung greift Homeier (2008, S. 133f.) den Ablaufplan der obersten Schulaufsichtsbehörde auf (s. Abb. 8; verkürzte Darstellung; vollständig s. MSW NRW 2009a, S. 66) und verknüpft ihn mit den Prozessschritten *Kooperative Rezeption* (1), *Kooperative Zielreflexion* (2) und *Kooperative Handlungsplanung* (3). In einem ersten Schritt werden die Ergebnisse des Qualitätsberichts unabhängig voneinander von der Schule und Schulaufsicht rezipiert. Ziel ist es, dass die Schule ein gemeinsames Verständnis entwickelt, in welchen Bereichen ihre Stärken und Schwächen liegen. In der kooperativen Zielreflexion sollte eine eindeutige Prioritätensetzung und eine gemeinsame Zielvorstellung vereinbart werden, die in einer konkreten Handlungsplanung mündet. Es muss sich dabei nicht zwangsläufig um Schwächen der Schulen handeln, sondern es können auch Stärken weiter ausgebaut werden. Auch die zuständigen Schulaufsichtsmitarbeiter müssen unabhängig von den Schulen die Ergebnisse

der Qualitätsanalyse rezipieren, reflektieren und Überlegungen zu möglichen Handlungsplanungen anstellen. Erst dann laufen die unabhängig voneinander abgeschlossenen Prozessschritte im Abschluss einer Zielvereinbarung zusammen. Bis zum vereinbarten Bilanzgespräch setzt die Schule die Zielvereinbarung um, während der Schulaufsicht die Aufgabe des Controllings zukommt.



**Abb. 8: Idealtypischer Ablauf des Zielvereinbarungsprozesses zwischen Schule und Schulaufsicht nach erfolgter Qualitätsanalyse**

Der beschriebene Zielvereinbarungsprozess nach der Schulinspektion stellt sich aus der Perspektive der motivationspsychologischen Erkenntnisse wie folgt dar: Zunächst lässt sich das Rubikon-Modell der Handlungsphasen auf den Zielvereinbarungsprozess nach Schulinspektionen übertragen. Mit der Verschriftlichung einer Zielvereinbarung nach der Schulinspektion wird der „Rubikon“ überschritten, d. h. mehrere Handlungsoptionen werden in ein bestimmtes verbindliches Ziel umgewandelt. Dieser Prozess setzt die Rezeption und Reflexion

der rückgemeldeten Ergebnisse voraus. In der präaktionalen Phase werden konkrete Schritte und Strategien geplant, die zur Zielerreichung führen. Es folgt die aktionale und postaktionale Handlungsphase in Form von Maßnahmenumsetzungen und Bilanzgesprächen. Neben dem Rubikon-Modell werden im Zielvereinbarungsprozess auch die zentralen Erkenntnisse der Zielsetzungstheorie einbezogen. Durch die Berücksichtigung der SMART-Regel und der Zielvereinbarungsvorlage erfüllen Zielvereinbarungen eine Reihe der Zielmerkmale, welche im Rahmen der Zielsetzungstheorie identifiziert wurden und welche einen positiven Einfluss auf die Handlungen ausüben.

Der theoretisch möglichen Übertragbarkeit der motivationspsychologischen Erkenntnisse auf den Zielvereinbarungsprozess nach Schulinspektionen stehen allerdings in der Praxis Einschränkungen gegenüber: Viele Studien der Zielsetzungstheorie sind in kontrollierten Laborbedingungen und mit einfachen Arbeitsaufgaben durchgeführt worden (vgl. Bungard 2000, S. 23). Damit ist die Implikation auf den Bildungsbereich unter den Bedingungen der komplexen schulischen Arbeit nur begrenzt möglich. Hinzu kommt, dass sich die Studien der Zielsetzungstheorie überwiegend auf quantitative Ziele beschränken und auf qualitative Ziele nur bedingt zutreffen. Im Bildungsbereich sind Ziele hingegen i. d. R. qualitativer Art und erschweren damit eine Spezifizierung und Operationalisierung der Zielerreichung. Zusammenfassend sind die Kernaussagen der motivationspsychologischen Theorie nützlich für die theoretische Einbettung von Zielvereinbarungen nach Schulinspektionen, jedoch sind der praktischen Anwendung gewisse Grenzen der Übertragbarkeit gesetzt.

Grundsätzlich stehen Zielvereinbarungen im Anschluss an die Schulinspektion in einem Spannungsfeld: Einerseits sollen sie frei von Anordnungen und Zielvorgaben sowie flexibel einsetzbar sein, andererseits sollten sie sich auf die Inspektionsergebnisse beziehen. Dadurch, dass der Abschluss von Zielvereinbarungen von der Systemebene vorgegeben wird und die eigenverantwortlichen Schulen nicht gänzlich frei in der Zielwahl sind, herrscht ein Ungleichgewicht. Zudem kann die Prämisse der gleichberechtigten Vertragspartner in Frage gestellt werden. Aufgrund ihrer hierarchischen Stellungen können Schulen und die Schulaufsicht grundsätzlich nicht gleichberechtigt sein, insbesondere nicht angesichts der Doppelrolle der Schulaufsicht. Die oberste Schulaufsicht fungiert im Zielvereinbarungsprozess als Auftraggeber, während die Aufgaben der

oberen und unteren Schulaufsicht neben der Beratung und Unterstützung des Zielvereinbarungsprozesses auch die Kontrolle der Zielerreichung umfassen.

Am schwierigsten ist die Frage zu beantworten, ob die Zielvereinbarungen ein wirksames Instrument der Schulentwicklung darstellen, wenn die Nicht-Einhaltung von Zielvorhaben nicht sanktioniert wird und sie rechtlich nicht bindend sind. Gerade die erfolgreiche Umsetzung der Ziele in Handlungen sowie die Initiierung von Schulentwicklungsmaßnahmen und nicht die Verschriftlichung der Ziele im vorgegebenen Formular stellt den Erfolgsmaßstab und das ausschlaggebende Kriterium für die Wirksamkeit des Instruments dar. Ziele und Maßnahmen stellen zwar eine Einheit dar, jedoch dürfen Maßnahmen nicht als Ziele verstanden werden, auch wenn sie einfachere Überprüfungsmöglichkeiten bieten (vgl. Reimers 2008). In diesem Zusammenhang besteht auch die Gefahr, dass nicht die Ziele, sondern eher die Zielerreichung fokussiert wird. Dies verleitet zu „kurzfristigem Aktionismus“ und zur Vernachlässigung „längerfristiger Zieldimensionen“ (Sprenger 2000, S. 111). Die Wirksamkeit und die bindende Wirkung des Instruments Zielvereinbarung wird auch entscheidend von der Akzeptanz der Beteiligten abhängen: Diese ist dann nicht gegeben, wenn Misstrauen mitschwingt, die betroffenen Schulleiter oder Lehrkräfte sie als Mittel der Fremdkontrolle wahrnehmen und Vorbehalte haben. Deshalb sollten partizipative Ziele und keine verdeckt zugewiesenen Ziele vereinbart werden. Ob und inwiefern die Schulen und die Schulaufsicht von den Zielvereinbarungen profitieren bzw. ob der verbindliche Entwicklungsprozess nach erfolgter Schulinspektion als hilfreich empfunden wird, wird sich erst nach einigen Jahren der Praxis und entsprechender Evaluation zeigen.

Im folgenden Kapitel wird auf der Grundlage der Rezeptionsforschung konkretisiert, auf welche Art und Weise die Rezeption von Rückmeldungen im Bildungsbereich erfolgt und wovon sie beeinflusst wird.

### 4.3. Rezeptionsforschung

Im Zuge der ersten Ergebnisse der internationalen Schulleistungsuntersuchungen TIMSS 1995 und PISA 2000 machte Terhart (2002, S. 108) auf das Ungleichgewicht zwischen der Messung von schulischen Leistungsdaten und der Kenntnis ihrer Nutzungsmöglichkeiten aufmerksam. Die Schieflage manifestiert sich dadurch, dass die Frage, wie die Ergebnisse von Leistungs- und Evaluationsstudien für die Verbesserung von Schule und Unterricht genutzt werden können, von der Bildungspolitik und Wissenschaft zunächst nachrangig behandelt wurde. Auch mit dem rasanten Anstieg externer Evaluationsformen seit TIMSS und PISA konnte diese Dysbalance nicht erkennbar behoben werden. Die Gründe dafür können einerseits in der Implementierung und Konsolidierung nationaler und regionaler Leistungsstudien, welche einen Großteil der bildungspolitischen Anstrengungen bündelten, gesehen werden. Andererseits kann als Erklärung das Wissensdefizit bezüglich der Ergebnisrezeption, -nutzung und -wirkung der externen Evaluationsformen, welche „die Erziehungswissenschaft relativ unvorbereitet“ (Stamm 2003, S. 184) getroffen haben, angeführt werden. Da sich aber gerade die Kenntnis der Nutzungsmöglichkeiten als zentral für die Legitimierung der zahlreichen externen Evaluationsformen erweist, wurde in diesem Zusammenhang ein spezifischer Forschungstypus angelehnt an die Verwendungs-, Interventions- und Innovationsforschung gefordert (vgl. Terhart 2002, S. 104-107). Mittlerweile hat sich dieser neue Forschungstypus – die *Rezeptionsforschung* – etabliert und geht der Frage nach,

wie Informationen über Ergebnisse von Evaluationsstudien an die Adressaten gelangen, wie sie dort aufgenommen, bewertet und genutzt werden, zu welchen Maßnahmen und Aktivitäten sie führen und welche Effekte diese dann haben (Schrader/Helmke 2004, S. 141f.).

Als weiterer Schwerpunkt der Rezeptionsforschung gilt die Rückmeldegestaltung, d. h. die Frage,

wie die Resultate externer Evaluationen in möglichst informativer, verständlicher und nutzbringender Form an die beteiligten Schulen und Lehrkräfte zurückgemeldet werden können (Kohler 2009, S. 83).

Analog zu den externen Evaluations- und Steuerungsinstrumenten kann die Rezeptionsforschung insbesondere im angloamerikanischen Raum auf eine längere Forschungstradition zurückblicken und entsprechend zahlreiche empirische Studien vorweisen, während sie in Deutschland erst seit etwa einer Dekade existiert (vgl. Dederling 2011, S. 65). Die im angloamerikanischen Raum vorliegenden Erkenntnisse erweisen sich durch die starke Betonung der Rechenschaftslegung externer Steuerungsinstrumente nur eingeschränkt auf die deutsche Situation übertragbar, so dass die verstärkte Durchführung von Rezeptionsstudien auch in Deutschland notwendig geworden ist.

Das von Terhart konstatierte Ungleichgewicht ist zwar mit den aktuellen Bemühungen der Rezeptionsforschung noch nicht ausbalanciert, jedoch konnten sich auch hierzulande die Forschungsanstrengungen in die geforderte Richtung potenzieren und bedeutende Erkenntnisse liefern. Dabei kann vorgegriffen werden, dass weder eine Linearität noch ein Automatismus zwischen der Bereitstellung von Ergebnismeldungen, ihrer unmittelbaren Nutzung und der Schul- und Unterrichtsentwicklung zugrunde gelegt werden kann (vgl. Stamm 2003, S. 196; Kohler/Schrader 2004, S. 6; EMSE 2009, S. 390). Rolff (2002, S. 83) merkt hierzu an:

Rückmeldungen von Daten und politische wie innerschulische Nutzung sind zweierlei Welten, die jeweils eigenen Regeln gehorchen. Die Daten stammen aus der Welt der Wissenschaft, die Nutzung passiert in der Welt der politischen und pädagogischen Praxis. Die Handelnden in der erstgenannten Welt wissen mehr als sie können, die Handelnden in der letztgenannten können mehr als sie wissen.

Weiterhin ergänzt er: „Beim Verhältnis von Datenrückmeldung und Nutzung handelt es sich um ein Juxtapositionsverhältnis, also um eine Nebeneinanderstellung und nicht um ein unmittelbares Ineinandergreifen“ (ebd.).

Die Frage, ob das allgemein angenommene Nutzungspostulat der Evaluation tatsächlich Bestand hat oder ihm eine Illusion inhärent ist, wird von Stamm (2003) gestellt: So zeigen sich auf der Grundlage der Befunde der PISA-Studie unterschiedliche Nutzungsverläufe und -typen in Schulen, die von der engagierten (Innovation) über die diffuse bis zur ausbleibenden Nutzung (Blockade) der Evaluationsergebnisse reichen. Resümierend stellt Stamm fest, dass zwar die

„Evaluation eindeutig vom möglichen Vorwurf der Folgenlosigkeit freigesprochen werden“ (Stamm 2003, S. 195) kann, die Frage nach der Berechtigung der Wirkungshoffnungen vorerst jedoch mit einem „Plädoyer für mehr Distanz zu Evaluationen“ (ebd., S. 197) beantwortet wird.

Damit müsste der bei externen Evaluationen häufig intendierte Impulscharakter von Rückmeldungen auf Schul- und Unterrichtsentwicklung relativiert werden: Die Impulsgebung bezieht sich mutmaßlich weniger auf die Initiierung von Handlungen, als vielmehr auf reflexive Prozesse, welche aber durchaus auch in Handlungen münden können. Dies belegt auch folgendes Zitat von Schneewind und Kuper (2009, S. 125f.): „Ergebnisrückmeldungen regen in erster Linie zu Reflexion an, die aber nicht automatisch in verändertes Handeln mündet“. Zentral ist, dass durch die Rückmeldungen im Sinne der mechanischen Impuls-Metapher etwas in Bewegung gesetzt wird und eine ausbleibende Reaktion somit ausgeschlossen ist. Aber vielleicht verleitet auch gerade dieser Metapherngebrauch bezüglich der Konzeption der externen Evaluationsformen dazu, die Stärke des Bewegungsanstoßes zu überschätzen.

Neben der Vielzahl an Einzelstudien zur Rezeption von Ergebnisrückmeldungen bestimmter Evaluationsformen liegen zwei berücksichtigungswerte Bestandsaufnahmen von Rezeptionsstudien vor: Eine aktuelle Bestandsaufnahme von Dederling (2011)<sup>24</sup> sowie eine frühe von Bonsen und von der Gathen (2004)<sup>25</sup>. Zudem haben sich zwei vollständige Themenhefte der Zeitschrift *Empirische Pädagogik* in den Jahren 2004 und 2007 den Themen Ergebnisrückmeldung, Rezeption, Nutzung und Nutzen von Evaluationsstudien in Schule und Unterricht angenommen. Schließlich wurde von Bohl und Kiper (2009) ein facettenreiches Herausgeberwerk mit dem Titel „Lernen aus Evaluationsergebnissen“ publiziert. Diese in den letzten Jahren verstärkte Präsenz der Rezep-

---

<sup>24</sup> In Dederings (2011) Übersicht der Rezeptionsstudien im deutschsprachigen Raum werden die internationalen Leistungsvergleichsstudien PISA 2000, IGLU und DESI, die Leistungsvergleichsstudien LAU, MARKUS, QuaSUM, die kontinuierlichen Lernstandserhebungen aus zehn Bundesländern sowie vier sonstige Studien (BeLesen, Klassencockpit, Check 5, Überprüfung der österreichischen Bildungsstandards) einbezogen.

<sup>25</sup> Bonsen und Gathen (2004) berücksichtigten in ihrer Bestandsaufnahme der Erfahrungen mit der innerschulischen Verarbeitung von Leistungsrückmeldungen folgende Studien: LAU, TIMSS, PISA 2000, IGLU, QuaSUM und MARKUS/WALZER.

tions- und Nutzungsfragen ist ein deutlicher Hinweis auf ihre Bedeutung und Notwendigkeit.

Dederings (2011) Beitrag ist als systematische Aufarbeitung der empirischen Erkenntnislage der Rezeptionsforschung im deutschsprachigen Raum konzipiert. Die Autorin gibt einen Überblick über die Forschungslage von Rezeptionsstudien im Hinblick auf Evaluationen der fachlichen und überfachlichen Schülerleistungen. Dazu sind insbesondere zentrale Lernstandserhebungen zu zählen, welche durch wiederholte fachbezogene und standardisierte Leistungstests in bestimmten Jahrgangsstufen den Leistungsstand der Schüler eines Bundeslandes erfassen (vgl. Vock/Pant 2008). Die diesbezüglichen Rezeptionsstudien befassen sich fast ausschließlich mit unterrichtsmethodischen Auswirkungen von Rückmeldedaten auf der Basis von Lehrereinschätzungen. Rezeptionsstudien, welche Evaluationsdaten im Allgemeinen, intern erhobene Daten und die Rezeption durch „unbeteiligte Akteure“ wie die Schulaufsicht oder Eltern zum Gegenstand haben, bleiben in Dederings Forschungsüberblick unberücksichtigt (vgl. Dederings 2011, S. 69). Die Grundlage ihrer aktuellen Bestandsaufnahme bilden somit 15 Studien, welche in mindestens 40 Publikationen besprochen worden sind und in denen die Datenerhebung überwiegend über (teil-)standardisierte, schriftliche Befragungen direkt beteiligter Lehrkräfte erfolgt (vgl. ebd., S. 69f.).

Bevor eine Vorstellung der zentralen Befunde der Rezeptionsforschung vorgenommen wird, werden zwei prominente Modelle der Ergebnisnutzung skizziert, wovon das eine für die empirischen Untersuchungen der vorliegenden Arbeit als forschungsleitend gilt. Anschließend erfolgt eine Darstellung zentraler Rezeptionsstudien zu unterschiedlichen Evaluationsformen im schulischen Bereich und ihrer Erkenntnisse: Zur Schulleistungsuntersuchung TIMSS (vgl. Kohler 2004), zu den Lernstandserhebungen MARKUS<sup>26</sup> (vgl. Schrader/Helmke 2003; 2004) und Lernstand 9 (vgl. Hartung-Beck 2009) sowie einer Längsschnittuntersuchung zur Lesekompetenzentwicklung (vgl. Schneewind 2006). Mit der Auswahl soll ein möglichst breites Spektrum vorliegender Rezeptionsstudien abgedeckt werden, um eine systematische Befundlage deutscher Rezeptionsstudien wiederzugeben.

---

<sup>26</sup> **Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterricht, Schulkontext**

### 4.3.1. Modelle zur Rezeption und Nutzung von Ergebnisrückmeldungen

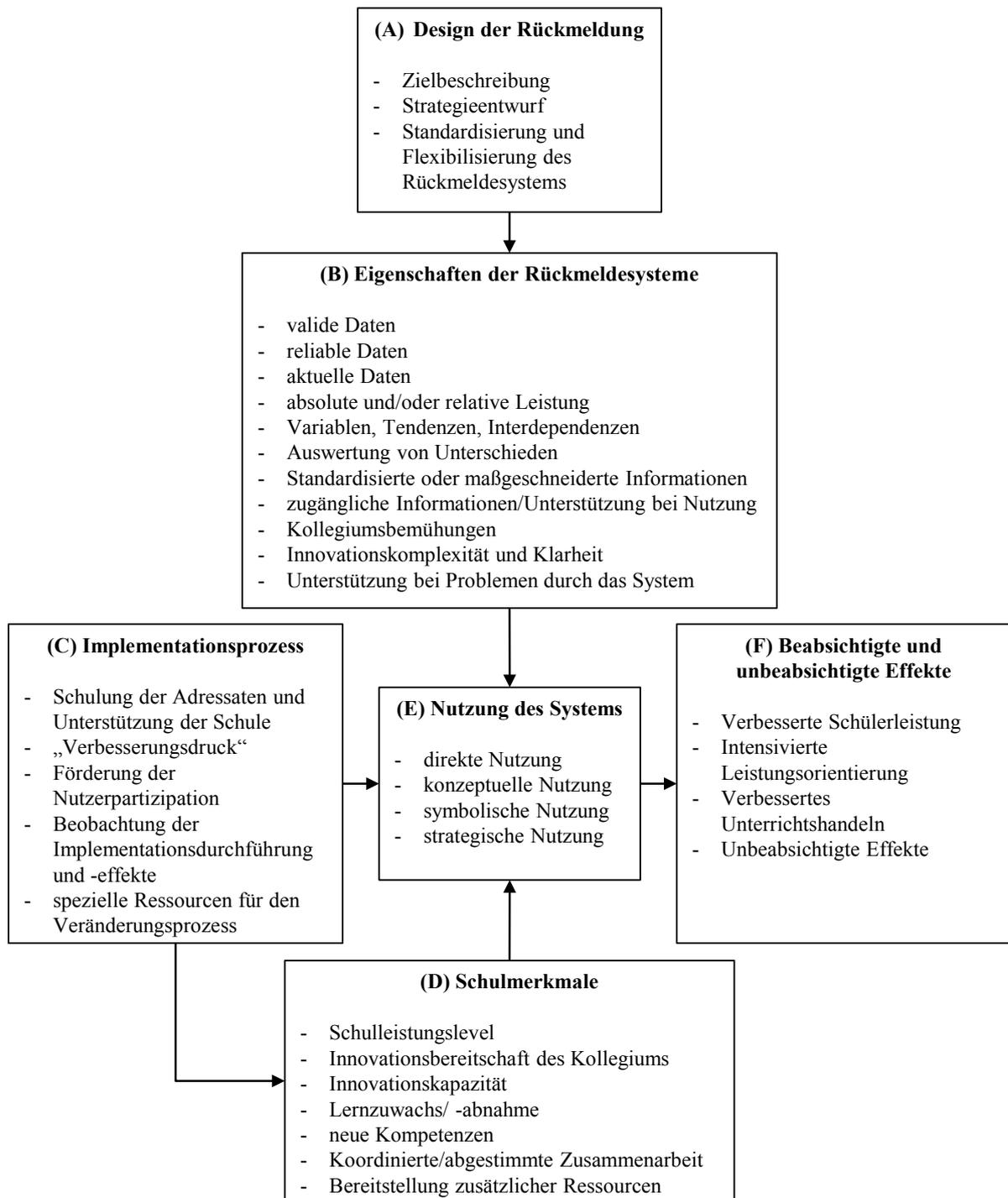
Da die genaue Kenntnis über die Ergebnisnutzung und -wirkung von Evaluationen im schulischen Bereich noch ein Desiderat darstellt, sind heuristische Modelle entwickelt worden, welche dem Prozess von der Rückmeldung bis zur Nutzung dieser Ergebnisse unter der Berücksichtigung verschiedener beeinflussender Faktoren Form verleihen. Sie sind einerseits im Rahmen modellgeleiteter Untersuchungen, andererseits auch in der Praxis als Richtschnur bzw. Idealverlauf der prozessgeleiteten Ergebnisnutzung anwendbar.

In Anlehnung an Visschers und Coes (2002) *School Performance Feedback System* (SPFS) haben Bosen und von der Gathen (2004, S. 249) ein *Rahmenmodell für die Rückmeldung von Leistungsdaten* entwickelt, welches die inner-schulische Nutzung vor dem Hintergrund zahlreicher Faktoren skizziert (s. Abb. 9). Dabei nehmen das Design der Rückmeldung (A), die Eigenschaft der Rückmeldesysteme (B), der Implementierungsprozess (C) sowie die Schulmerkmale (D) Einfluss auf die Nutzung (E) und die beabsichtigten und unbeabsichtigten Effekte (F) von Leistungsrückmeldungen. Die Nutzung von Datenrückmeldungen kann in Anlehnung an Rossi und Freeman (1993) differenziert erfolgen: Instrumentell (direkt), konzeptionell, d. h. indirekt bspw. über einen Einstellungswandel und symbolisch. Bei der symbolischen Nutzung werden die Rückmeldungen für bestimmte taktische Absichten, z. B. zur (nachträglichen) Legitimierung von Entscheidungen<sup>27</sup> oder als Argumentationshilfe eingesetzt. Als zusätzliche Nutzungsform wurde die strategische Nutzung in das Modell mitaufgenommen, welche eine missbräuchliche Variante der Datennutzung darstellt, z. B. „teaching to the test“.

Das Modell zeichnet sich durch seinen Bezug zur Schulentwicklungsforschung aus, lässt allerdings die Modellierung einer Mehrebenensicht und Hinweise auf Zusammenhänge zwischen den einzelnen Modellkomponenten unberücksichtigt.

---

<sup>27</sup> Die symbolische Nutzungsform wurde besonders häufig bei der Rezeption und Verarbeitung der PISA-Studie und ihrer Ergebnisse von den Schulministerien beobachtet (vgl. Tillmann/Dedering/Kneuper/Kuhlmann/Nessel 2008).



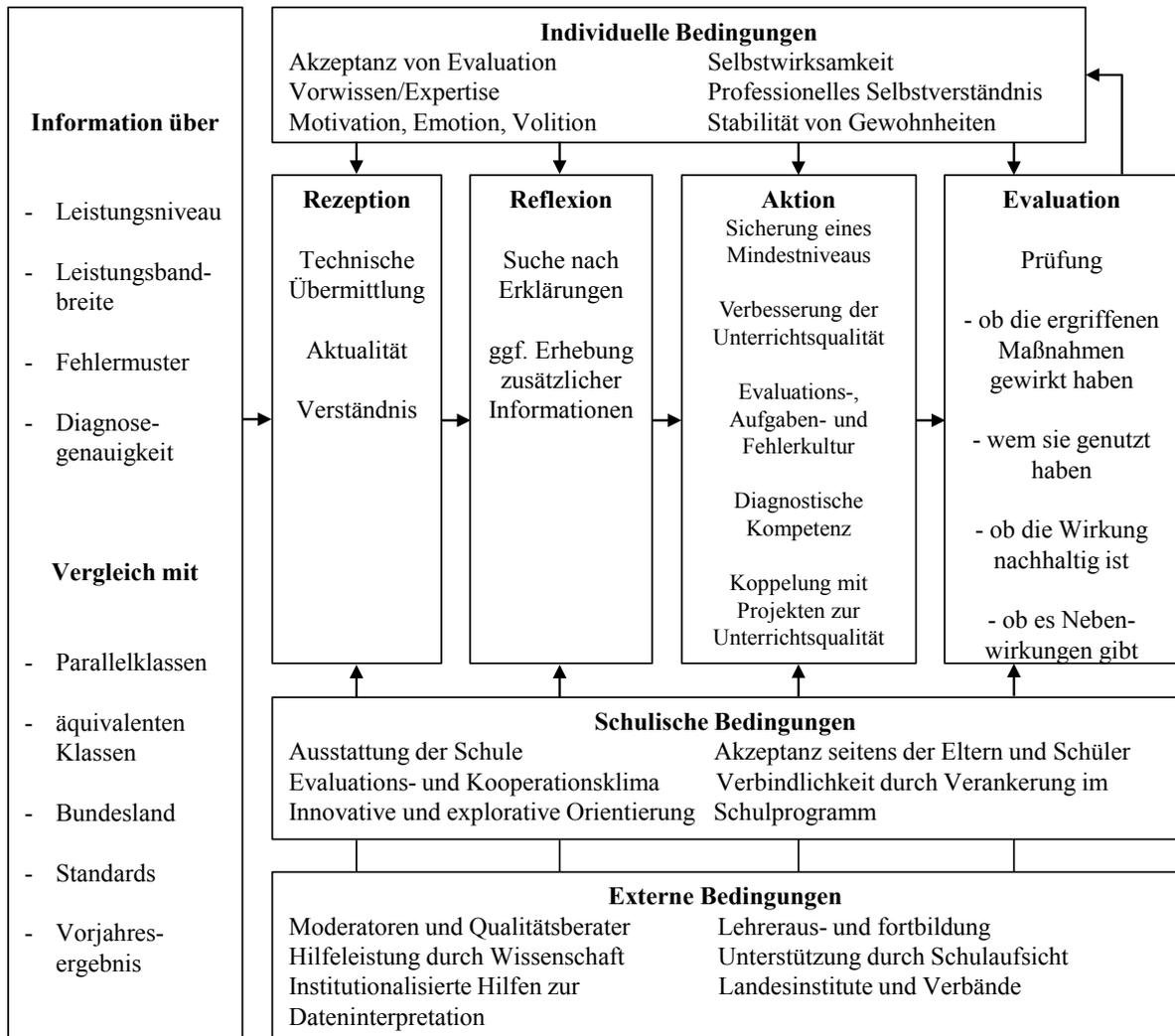
**Abb. 9: Rahmenmodell für die Rückmeldung von Leistungsdaten (Bonsen/von der Gathen 2004)**

Auch Helmkes (2004) Variante der idealtypischen Verwendung von Datenrückmeldungen ist lediglich aus der Perspektive einer Lehrkraft oder Einzelschule modelliert. Sein Rahmenmodell zur pädagogischen Nutzung von Lernstandserhebungen ähnelt in seiner Grundstruktur dem – wiederum von ihm selbst entwickelten – Angebots-Nutzungs-Modell zur Unterrichtswirksamkeit (vgl. Helmke 2009, S. 71-78; S. 100) und existiert auch in anderen Varianten, bspw. als Sequenzmodell der Unterrichtsentwicklung und ihrer Bedingungen (vgl. ebd., S. 308-320). Es wurde für den Umgang mit Ergebnisrückmeldungen aus Lernstandserhebungen entwickelt, kann aber auch im Rahmen anderer Evaluationsformen eingesetzt werden (vgl. z. B. Schneewind 2007a, b; Hosenfeld 2010; Sommer 2011).

Als Charakteristikum des Rahmenmodells gilt der lineare bzw. zyklische Verlauf und die analytische Differenzierung des komplexen Nutzungsvorgangs in die vier Phasen *Rezeption*, *Reflexion*, *Aktion* und *Evaluation*. Die jeweiligen Phasen bedingen den erfolgreichen Abschluss der vorangegangenen Phase(n) und schließen ein Überspringen aus. Die Übermittlung, Auswahl und das Verständnis der Informationen bilden inhaltlich die Phase *Rezeption* ab. D. h. die Daten müssen, bevor sie in Aktionen münden, physikalisch in den Schulen ankommen, wahrgenommen und verstanden werden. Dazu sollten die Ergebnisrückmeldungen so zusammengestellt sein, dass sie alle Informationen in eindeutiger und unmissverständlicher Aufbereitung enthalten. Helmke zufolge stellt die *Reflexion* der Ergebnisse die entscheidende Phase im gesamten Prozess dar und schließt die Frage nach möglichen Ursachen für die Ergebnisse und die Erhebung zusätzlicher Informationen mit ein. In der Phase *Aktion* ist das Ergreifen spezifischer Maßnahmen, Handlungen und Initiativen vorgesehen. Den Schlusspunkt des Prozessdurchlaufs und Ausgangspunkt eines neuen Zyklus bildet schließlich die Bewertung der durchgeführten Maßnahmen (*Evaluation*), welche sich durch eine wiederholte Durchführung einer Lernstandserhebung überprüfen ließe. Das Gelingen der pädagogischen Nutzung von Ergebnissen aus Lernstandserhebungen wird zudem von verschiedenen individuellen, externen und schulischen Faktoren bedingt (s. Abb. 10).

Resümierend ist das Helmke-Modell als heuristische Bezugsgröße stimmig aufgebaut und enthält (implizite) Annahmen über Wirkungszusammenhänge und Prozessbedingungen der Ergebnisrezeption und -nutzung. Der Wert dieses

Modells liegt in seinem idealtypischen Verarbeitungsprozess und in seiner Eignung, die Rezeption und Nutzung von Rückmeldungen der zahlreichen Evaluationsformen systematisch zu analysieren.



**Abb. 10: Der Zyklus Rezeption - Reflexion - Aktion - Evaluation im Kontext von Lernstandserhebungen (Helmke 2004)**

Sowohl das Modell von Bensen und von der Gathen (2004) als auch von Helmke (2004) belegen, dass lediglich die Bereitstellung von Rückmeldungen nicht ausreicht, damit sie handlungsleitend genutzt werden. Vielmehr wird der komplexe Prozess von der Rezeption bis zur Nutzung von Rückmeldungen von verschiedenen Bedingungen auf der Personen- und Organisationsebene beeinflusst, welche notwendige, aber keine hinreichenden Größen darstellen.

### 4.3.2. Überblick über ausgewählte Rezeptionsstudien

Zur Untersuchung der Rezeption der TIMSS-Befunde führte Kohler (2004; 2005; 2009) in den Jahren 2000 und 2001 eine Fragebogenerhebung mit Lehrkräften (N = 201), Eltern (N = 194) sowie Schulaufsichtsmitarbeitern durch. Berücksichtigt wurden 13 allgemeinbildende Schulen in Baden-Württemberg, welche selbst resp. das gesamte Bundesland nicht an der TIMS-Studie teilgenommen haben. Bei den Lehrkräfte- und Elternstichproben wurde eine nahezu 100 %-ige Rücklaufquote erreicht, während sie bei den Schulaufsichtsmitarbeitern bei weit über 50 % lag. Untersuchungsleitend waren die Fragestellungen, wie die einzelnen Befragungsgruppen die Ergebnisse der TIMS-Studie wahrnehmen und worauf sie die eher moderaten Ergebnisse der Schülerleistungen in Deutschland zurückführen. Etwa die Hälfte der befragten Lehrkräfte und nahezu alle Eltern waren im Vorfeld der Rezeptionsstudie über die TIMS-Studie nicht informiert. Im Rahmen der Befragung wurden die Lehrkräfte und Eltern über kurze Informationseinheiten, welche Artikeln in Tageszeitungen entsprechen, über die TIMSS-Befunde instruiert. Die Lehrkräfte und Eltern hielten die Ergebnisse der TIMS-Studie in einem mittleren Maß für bedeutend, während die Zustimmung der Schulaufsichtsmitarbeiter in einem signifikant höheren Bereich lag. Im Hinblick auf die Attribution bezüglich der moderaten TIMSS-Ergebnisse machten alle drei Befragungsgruppen einvernehmlich in erster Linie die Schüler verantwortlich, während der jeweiligen eigenen Personengruppe die Verantwortung am wenigsten unterstellt wurde. Einschränkungen der Rezeptionsstudie bestehen bzgl. der Verallgemeinerung der Befunde, da in der Untersuchung nicht unmittelbar an TIMSS beteiligte Akteure befragt wurden, es sich nicht um von ihnen verantwortete Ergebnisse handelte und sie auch als keine direkten Adressaten der Rückmeldungen fungierten. Hinzu kommt, dass TIMSS als erste groß angelegte Large-Scale-Untersuchung ein Novum für die (Schul)Öffentlichkeit darstellte und ihre Ergebnisse primär an die Systemebene adressiert waren. In weiteren, an dieser Stelle nicht explizit vorgestellten Rezeptionsstudien auf der Basis von Systemmonitoring-Untersuchungen zeigt sich für diesbezügliche Rückmeldungen generell, dass sie von den Schulen nur eingeschränkt für die Schul- und Unterrichtsentwicklung genutzt werden (vgl. Terhart 2002, S. 93; Watermann/Stanat 2004, S. 40f.).

Mit der Implementierung von Lernstandserhebungen entstand ein Forschungsfeld, das sich mit der Akzeptanz, Rezeption und Nutzung von Rückmeldedaten durch Lehrkräfte und Schulen beschäftigt. Im Rahmen der Rezeptionsstudie WALZER<sup>28</sup> wurde von 2001 bis 2002 eine Lehrkräfte- (n = 81) und Schulleitungsbefragung (n = 52) in Rheinland-Pfalz durchgeführt, anhand derer untersucht wurde, wie Schulen die Evaluationsstudie MARKUS rezipiert haben und wie sie ihre Ergebnisse für die Initiierung von Schulentwicklungsmaßnahmen genutzt haben (vgl. Schrader/Helmke 2003; 2004). Der Studie lag ein Rahmenmodell zur Rezeption zugrunde (vergleichbar zu Helmke; s. Kapitel 4.3.1.). Es handelt sich um nicht-repräsentative Stichproben, da bei den Lehrkräften lediglich eine Rücklaufquote von 4,3 % und bei den Schulleitungen eine von 8,3 % erreicht werden konnte und somit von verzerrten, d. h. aktiven, innovationsfreudigen und pflichtbewussten Stichproben auszugehen ist. Daher ist die vorliegende Befundlage nur eingeschränkt verallgemeinerbar. Die Schulleitungsbefragung hat ergeben, dass die Schulen ausgehend von den Ergebnisrückmeldungen durchaus die Notwendigkeit für Änderungen und Aktivitäten bezüglich des Qualitätsmanagements erkannt und diese eingeleitet oder ausgebaut hätten. Diese Erkenntnisse interpretieren Schrader und Helmke (2003, S. 101) als einen durch die Evaluation ausgelösten Änderungsdruck in den Schulen. Sowohl zwischen der Einstellung der Schulleitung zur externen Evaluation als auch ihrem pädagogischen Interesse konnte ein positiver Zusammenhang zum Umfang der eingeleiteten Maßnahmen festgestellt werden. Hingegen gab der Großteil der befragten Lehrkräfte an, keine Konsequenzen aus der MARKUS-Ergebnisrückmeldung gezogen zu haben. Dabei schätzten sie ihre Bereitschaft, den eigenen Unterricht zu verändern und ihren Fortbildungsbedarf umso größer ein, je positiver die externe Evaluation wahrgenommen wurde und je höher ihre Selbstwirksamkeit ausgeprägt war. Lag eine generelle Unterstützung durch den schulischen Kontext vor, desto eher wurden Maßnahmen der Unterrichtsgestaltung geplant und umgesetzt. Insgesamt spielen die Ergebnisse für die Initiierung von Schulentwicklungsmaßnahmen eine eher marginale Rolle: Änderungen im Bereich des Qualitätsmanagements als Folge von MARKUS wurden überwiegend in einer Minderheit der Klassen (5 % bis 10 %) realisiert.

---

<sup>28</sup> **Wirkungsanalyse der Leistungsevaluation: Zielerreichung, Ertrag für die Bildungsqualität der Schule und die Rückmeldung von Evaluationsergebnissen**

Die Autoren folgern aus diesen ernüchternden Befunden (Schrader/Helmke 2004, S. 157):

Das Ergebnis kann dazu beitragen, ein unrealistisch hohes Erwartungsniveau, was die Umsetzung von Evaluation in Innovation angeht, nach unten hin zu korrigieren und sich mehr (und früher) Gedanken darüber zu machen, wie man evaluationsbasierte Schulentwicklung organisatorisch und sozialpsychologisch so vorbereitet, stützt und begleitet, dass der Erfolg in einem angemessenen Verhältnis zum investierten Aufwand steht.

Eine weitere, auf einer umfangreichen Stichprobe (N = 1510 Lehrkräfte) basierende Rezeptionsstudie wurde zum Projekt VERA (Vergleichsarbeiten in der Grundschule) im Jahr 2005 durchgeführt (vgl. Koch/Groß Ophoff/Hosenfeld/Helmke 2006; Groß Ophoff/Hosenfeld/Koch 2007). Die Forschungsfragen wurden am Rahmenmodell von Helmke (2004) entwickelt und sollten Hinweise zur Auseinandersetzung mit den VERA-Ergebnissen und ihrer Wirkung auf die Unterrichtsentwicklung geben. Die Autoren kommen zu dem Ergebnis, dass „die Schritte von der Rezeption über die Reflexion bis hin zur Aktion grundsätzlich gelungen sind“ (ebd., S. 196). Die Ergebnisse im Einzelnen zeigen, dass die Verständlichkeit und Nützlichkeit der Rückmeldung von den Lehrkräften zustimmend beurteilt wurden und die Auswertung kaum Probleme bereitete: Aus den Ergebnissen wurden zwar geeignete Fördermaßnahmen für einzelne Schüler abgeleitet, aber sie wurden nur in geringem Maße für die Reflexion des eigenen Unterrichts und eine ergebnisorientierte Unterrichtsentwicklung herangezogen. Zwei Drittel der befragten Lehrkräfte gaben an, sich nach der Beteiligung an VERA intensiver mit Bildungsstandards und Curricula auseinandergesetzt zu haben. Lediglich 6,7 % der Lehrkräfte offenbarten, in Folge von VERA keine Veränderungen des eigenen Unterrichts eingeleitet zu haben. VERA stellte für die Mehrheit der Lehrkräfte einen Anlass dar, Maßnahmen im Bereich „Unterrichtsinhalte“ und „Unterrichtsgestaltung“ zu realisieren. Trotz der generell positiven Befunde zeigte sich, dass die Lehrkräfte der Schwierigkeit gegenüberstanden, die Rückmeldungen handlungsleitend für die Unterrichtsentwicklung zu nutzen. Sie äußerten vielfach den Wunsch, praxisgerechte Anleitungen zu bestimmten Maßnahmen zu erhalten. Dass die eigenständige

Auseinandersetzung mit den Ergebnissen noch verbesserungswürdig erscheint, zeigt sich auch daran, dass die von den Lehrkräften abgeleiteten Maßnahmen eher routinierter und weniger innovativer Art waren. Ähnlich wie Stamm (2003) sowie Rossi und Freeman (1993) von verschiedenen Nutzungstypen von Evaluationsergebnissen berichten, so konnten auch Groß Ophoff, Hosenfeld und Koch (2007) drei unterschiedliche Rezeptionsprofile identifizieren: Dabei setzt sich der Rezeptionstyp 1 intensiv mit den VERA-Rückmeldungen auseinander und entwickelt daraus Lern- und Testmaterialien. Der Rezeptionstyp 2 fällt durch eine weniger intensive Auseinandersetzung mit den Ergebnissen auf, bewertet diese gleichzeitig als nützlich und weniger verständlich, während der Rezeptionstyp 3 die VERA-Rückmeldungen kritisch einschätzt und seltener Maßnahmen im Bereich der Unterrichtsentwicklung ableitet. Die Autoren erachten die identifizierten Rezeptionstypen insofern als bedeutend, als sie sich auch hinsichtlich ihrer realisierten Unterrichtsentwicklungsmaßnahmen unterscheiden und somit auf der Typisierungsgrundlage den Schulen gezielte und spezifische Unterstützungsangebote zur Verfügung gestellt werden sollten.

Im Rahmen des Projektes BeLesen (Berliner Längsschnittstudie zur Lesekompetenzentwicklung von Grundschulkindern) wurden in einer Rezeptionsstudie weitere Ergebnisse zum Umgang von Lehrkräften mit Ergebnisrückmeldungen erarbeitet (vgl. Schneewind 2007a, b). Das Forscherteam der Rezeptionsstudie, welches nicht zum Projektteam von BeLesen gehörte, führte zu fünf Messzeitpunkten schriftliche Befragungen von Lehrkräften ( $n = 56$ ) und Interviews mit einer Teilstichprobe von zehn Lehrkräften durch. Erörtert wurde, welche Einstellungen Lehrkräfte bezüglich Ergebnisrückmeldungen aufweisen, ob sie die zurückgemeldeten Informationen verstehen, akzeptieren und inwiefern sie die Bereitschaft bekunden, die Informationen zu reflektieren und in Handlungen umzusetzen. Die Anlage der Studie als Längsschnittuntersuchung ermöglichte es, die Einstellungen im zeitlichen Verlauf und vor dem Hintergrund möglicher Einstellungswandel zu beleuchten. Wie bei den vorangegangenen zwei gen. Untersuchungen WALZER und VERA liegt auch dieser Studie ein theoretisches Rahmenmodell zum Umgang mit Ergebnisrückmeldungen zugrunde (vgl. Helmke/Hosenfeld 2005), welches um die Aspekte, die den Ergebnisrückmeldungen vorausgehen, erweitert wurde. Als zentrales Ergebnis der Fragebogenstudie ist festzuhalten, dass die Einstellung der Lehrkräfte gegen-

über Rückmeldungen im Allgemeinen eher positiv ausgeprägt war und die Ankündigung der Ergebnismrückmeldungen speziell im Projekt BeLesen als Motivator zur Teilnahme an der Untersuchung fungierte. Sowohl die Verständlichkeit als auch der Nutzen der Rückmeldung wurden positiv bewertet. Die Übereinstimmung zwischen den eigenen Einschätzungen zur Leistung bestimmter Schüler und den zurückgemeldeten Informationen wurde in 75 % der Fälle bestätigt. Ein weiterer Befund zeigt den Zusammenhang zwischen der Spezifität der Information und ihrer Rezeption auf: Je spezifischer die Rückmeldung zu den Leistungen eines Schülers gestaltet war, desto intensiver wurden die Ergebnisse rezipiert. Die Interviewstudie konnte die Ergebnisse der schriftlichen Befragung bestätigen und weiterführende Befunde bereitstellen: Für die befragten Lehrkräfte stellte insbesondere die Übertragung der Informationen auf ihre Schüler die größte Herausforderung dar (vgl. Schneewind 2007b, S. 377). Ein Hindernis bei der Ergebnisnutzung sahen die Lehrkräfte in der fehlenden Zeit, die Ergebnisse gründlich zu reflektieren und konkrete Maßnahmen abzuleiten.

Als eine der wenigen qualitativen Rezeptionsstudien ist Hartung-Becks (2009) Untersuchung zur Rezeption und Nutzung von Rückmeldungen aus Lernstandserhebungen (Lernstand 9) in NRW hervorzuheben. Das Forschungsziel einer Modell- und Typenbildung zum Umgang mit den zurückgemeldeten Daten wurde in professions- und organisationssoziologische Theorien eingebettet. Auf der Organisationsebene interessierte Hartung-Beck (2009), „wie die Lernstandserhebung als Medium der Organisation bzw. als organisatorische Ressource in den schulischen Organisationsstrukturen etabliert wird, und [auf] der Professionsebene, wie die Nutzung der Daten durch professionelle (epistemologische) Überzeugungen strukturiert wird“ (Hartung-Beck 2009, S. 232). Auf der Grundlage eines Fallstudiendesigns wurden an zwei nordrhein-westfälischen Gesamtschulen Interviews mit 19 Lehrkräften durchgeführt und relevante Dokumente ausgewertet. Die Einzelfälle konnten den vier verschiedenen Nutzungstypen auf der Basis professioneller Überzeugungen zur Datennutzung zugeordnet werden: Während bspw. der Professionstyp A eine intensive Rezeption der Ergebnisse aufweist, ist der Typ D durch Ablehnung und Widerstand gegenüber Lernstandserhebungen und ihren Ergebnissen gekennzeichnet. Auch bezüglich der Nutzungstypen auf der Basis der organisationalen Überzeugungen der befragten Lehrkräfte konnten vier verschiedene Ausprä-

gungen differenziert werden: Während der Organisationstyp 1 die Rückmeldungen der Lernstandserhebungen als ein Instrument der autonomen professionellen Kontrolle der eigenen und kollektiven Arbeit nutzt, beschränkt sich der Organisationstyp 2 auf die individuelle Nutzung. Die weiteren Typen 3 und 4 empfinden die Lernstandserhebungen als von außen vorgegeben und zeichnen sich einerseits durch eine Auseinandersetzung mit den Ergebnissen auf schulorganisatorischer Ebene oder durch Widerstand und Ablehnung gegenüber Lernstandserhebungen aus. In einem weiteren Schritt wurden schulspezifische Nutzungsmuster auf der Basis des Zusammenhangs zwischen den unterschiedlichen professionellen und organisatorischen Überzeugungen der Lehrkräfte analysiert. Die befragten Lehrkräfte ließen sich elf der 16 Mustermöglichkeiten zuordnen und zeigten bestimmte Muster innerhalb der jeweiligen Gesamtschule auf, welche mit dem Führungsstil der Schulleitung erklärt werden konnten: Die Gesamtschule A ist durch einen dezentralen Führungsstil und eine ausgeprägte Absprache- und Kommunikationskultur bei der Auseinandersetzung mit den Ergebnissen der Lernstandserhebung gekennzeichnet. Hingegen wird bei der Gesamtschule B ein zentraler Führungsstil sowie eine geringer ausgeprägte Bereitschaft und Intensität der Ergebnisrezeption und -reflexion deutlich. Insgesamt zeigen die Ergebnisse, dass auch nicht-intendierte Formen der Verarbeitung zu beobachten sind und die Rückmeldungen aus Lernstandserhebungen auf unterschiedliche Nutzungstypen treffen, welche durch bestimmte professionelle und organisatorische Überzeugungen gekennzeichnet sind. Die Gegenüberstellung der Erkenntnisse der Typenbildung ohne und mit Schulspezifizierung zeigt deutlich, dass schulspezifische Faktoren einen großen Einfluss auf die Datennutzung haben. Hartung-Beck (2009) resümiert die erzielten Ergebnisse wie folgt:

Die Lernstandserhebungen können nur dann zu einer organisatorischen Ressource werden, wenn sowohl die schulischen Kapazitäten und Strukturen für eine Auseinandersetzung geschaffen, als auch eine direkte inhaltliche Anbindung an die Profession der Lehrkräfte hergestellt wird (S. 248).

Eine weitere berücksichtigenswerte Rezeptionsstudie wurde auf der Grundlage einer Lehrkräftebefragung (n = 1136) in Baden-Württemberg und Thüringen zu

den jeweiligen Lernstandserhebungen durchgeführt (vgl. Maier 2008a, b; 2009). Durch den Ländervergleich lieferte sie Erkenntnisse zur Rezeption und Nutzung von Leistungsrückmeldungen in Abhängigkeit verschiedener Systeme von Lernstandserhebungen: Die Kompetenztests in Thüringen unterscheiden sich insofern von den baden-württembergischen Vergleichsarbeiten, als die Lehrkräfte kriteriale und faire Vergleiche als Rückmeldung erhalten. Es zeigte sich, dass Lehrkräfte aus Thüringen den Lernstandserhebungen deutlich positivere Akzeptanzbekundungen bescheinigen als ihre baden-württembergischen Kollegen: Insbesondere die allgemeine Akzeptanz, die curriculare Validität, die Hinweise auf zukünftige Wiederholungen der Lernstandserhebungen wurden in Thüringen positiver eingeschätzt und die Belastung durch die Lernstandserhebungen als geringer wahrgenommen. In Baden-Württemberg dienten die Rückmeldungen aus Lernstandserhebungen hingegen eher als Hinweise für die Notengebung. In beiden Ländern wurde die Akzeptanz und die Nutzung der Lernstandserhebungen durch das getestete Schulfach beeinflusst: Zentrale Tests im Fach Mathematik erzielten dabei positivere Bewertungen als im Fach Deutsch. Die nachgewiesenen Unterschiede in der Wahrnehmung der Lehrkräfte führt Maier (2008a, S. 471f.) auf Merkmale der Test- und Rückmeldesysteme zurück. Das Test- und Rückmeldesystem in Thüringen ist an wissenschaftlichen Standards orientiert. Hingegen wurde den baden-württembergischen Lehrkräften lediglich der soziale Vergleichsmaßstab mit einer Pilotierungsstichprobe zur Verfügung gestellt, während faire Vergleichsdaten fehlten und sich dies negativ auf die Akzeptanz- und Nutzbarkeitseinschätzungen auswirkte. Damit vermag die Art des Systems der jeweiligen Lernstandserhebung das Rezeptions- und Nutzungsverhalten von Lehrkräften eindeutig zu beeinflussen. Die Durchführung von Interviews mit Lehrkräften an Regelschulen und Gymnasien in Thüringen ( $n = 18$ ) lieferte zusätzliche Befunde zur Nutzung der Rückmeldungen: Die diagnostische Funktion zentraler Lernstandserhebungen zur gezielten Individualförderung und weniger die Reflexion ihres Unterrichts stand für die befragten Lehrkräfte bei der Nutzung der Rückmeldungen im Vordergrund. Das Nutzungsverhalten war unterschiedlich ausgeprägt und konnte dabei den drei Kategorien instrumentelle, strategische und keine Nutzung zugeordnet werden.

### 4.3.3. Zusammenfassung

Zusammenfassend liefern die Befunde der Rezeptionsforschung eindeutige Hinweise darauf, dass auf der Grundlage der zurückgemeldeten Ergebnisse an Schulen nicht unweigerlich davon ausgegangen werden kann, dass hieraus unmittelbar Maßnahmen zur Schul- und Unterrichtsentwicklung resultieren.

Die wissenschaftliche Auseinandersetzung mit der Nutzung von Ergebnismeldungen durch Schulen setzte in der deutschsprachigen Forschung erst vergleichsweise spät vor etwa zehn Jahren ein, vollzieht sich aber seitdem in relativ zügigem Tempo. Als bedeutender Progress zählt die Modellierung von verschiedenen, aufeinander aufbauenden Prozessschritten, anhand derer sich die Fragen der Rezeptionsforschung modellgeleitet beantworten lassen (vgl. Modelle von Helmke sowie Bensen und von der Gathen). Bei der gründlichen Aufarbeitung des Forschungsstands zeigt sich, dass die zweite Fragestellung der Rezeptionsforschung zur Rückmeldegestaltung hinter den Fragen zum Prozess des Umgangs mit den Rückmeldungen zurückbleibt. Es herrscht weitestgehend immer noch eine Unsicherheit, wie eine ideale, adressatenspezifische Rückmeldung aussehen sollte. Des Weiteren werden in Rezeptionsstudien hauptsächlich Einstellungen der schulischen Akteure gegenüber Evaluationen erhoben und nicht der Transformationsprozess vom evidenzbasierten Wissen zu den Handlungen untersucht. Als problematisch ist dabei die vielfach vorliegende Übereinstimmung zwischen den Personen(gruppen), welche die jeweilige Evaluationsuntersuchung konzipieren und anschließend in einer Rezeptionsstudie überprüfen, anzusehen.

Die bisherige Forschungslage resümierend ist zu konstatieren, dass eine hohe Rezeptionsbereitschaft der schulischen Akteure vorliegt und dass Rückmeldungen aus Evaluationen von den befragten schulischen Akteuren überwiegend in zufriedenstellendem Ausmaß rezipiert werden. Sowohl die Verständlichkeit der Rückmeldungen als auch die allgemeine Nützlichkeit der Ergebnisse lassen positive Ausprägungen erkennen. Tendenziell zeigt sich eine umso positiver wahrgenommene Nützlichkeit von Rückmeldungen, je konkreter diese gestaltet sind und je unterrichtsferner die befragten Akteure verortet werden.

Diesen weitestgehend positiven Befunden zur Rezeption ist das begrenzte Nutzungsausmaß der Rückmeldungen entgegenzustellen. Rückmeldungen ha-

ben durch eine eher geringe veränderte Gestaltung des Unterrichts einen indirekten Effekt auf die Nutzung und werden im Rahmen von Lernstandserhebungen eher zur Leistungsdiagnostik herangezogen. Plakativ auf den Punkt gebracht, findet Rezeption in hohem Ausmaß statt, während die Nutzung sich als begrenzt erweist und keine stringente Folge der Rückmeldungen darstellt.

Insgesamt lässt sich bei der Rezeption als auch Nutzung ein unterschiedlicher und nicht immer intendierter Umgang mit Ergebnismrückmeldungen erkennen. Studien zeigen, dass Schulen und Lehrkräfte unterschiedlich auf bildungspolitische Reformen reagieren und sich auch die Rezeption und Nutzung unterschiedlich und nicht immer konform zu der intendierten Variante verhält (z. B. Stamm 2003; Hartung-Beck 2009; Maier 2008). Die Befunde verdeutlichen, dass der Umgang mit Ergebnismrückmeldungen von länderspezifischen resp. evaluationsinstrumentenabhängigen Faktoren als auch persönlichkeits-, professions- und organisationsabhängigen Variablen neben einer Vielzahl an noch ungeklärten Faktoren determiniert ist. Eine zentrale Voraussetzung der Ergebnismnutzung stellt das Verständnis und die Übersetzung der extern gemessenen Ergebnisse dar, wofür es spezifischer Kompetenzen, z. B. der Datenkompetenz sowie eines Unterstützungssystems bedarf.

Auch wenn im Bereich der Rezeptionsforschung zahlreiche Aktivitäten und Befunde zu verzeichnen sind, wodurch die Wissensbasis kontinuierlich erweitert werden konnte, sind gleichzeitig viele Fragen offen geblieben, was durch Altrichter (2010, S. 253) wie folgt herausgestellt wird: „Es scheint sich eine Routine der Rezeptionsstudien zu etablieren, die an ihre Grenzen stößt“. Damit wird auf die nahezu ausschließliche Befragung von Lehrkräften und Schulleitungen verwiesen, die auf der Basis von Selbstberichten keine Anhaltspunkte zu den Nutzungsprozessen zu liefern vermögen. Entsprechend sollten bspw. Interventionsstudien, Längsschnittstudien, Beobachtungen der Verarbeitungsprozesse oder weitere Perspektiven wie die der Schulaufsicht Berücksichtigung finden (vgl. ebd.).

Insgesamt wird die Rezeption und Nutzung von Ergebnismrückmeldungen im Bildungsbereich von einer Vielzahl an Einflussfaktoren, Bedingungen und Prozessen begleitet und bedarf weiterer Forschung. Diesem Desiderat versucht die vorliegende Arbeit zu entgegnen, indem im Folgenden die Rezeption und Nutzung von Daten der Schulinspektion empirisch beleuchtet werden.

## **5. Entwicklung eines theoretischen Rahmenmodells zur Erfassung der Rezeption und Nutzung von Schulinspektionsergebnissen**

In Kapitel 4.3.1. wurden die für die Fragestellung dieser Arbeit relevantesten Modelle zur Erfassung der Ergebnisrezeption und -nutzung im Bildungsbereich präsentiert und diskutiert. Auf die Vorteile des Rückgriffs auf Modelle in empirischen Untersuchungen macht Ditton (2007, S. 83) aufmerksam:

Modelle versuchen die Wirklichkeit in einer noch überschaubaren, aber dennoch genügend differenzierten Form abzubilden. Sie benennen dazu Elemente und deren Beziehungen, was sowohl der Forschung als auch der Praxis hilft, systematisch und strukturiert an (Teil-)Fragestellungen zu arbeiten.

Das Ziel dieses Kapitels ist es, das Rahmenmodell zur pädagogischen Nutzung von Lernstandserhebungen (vgl. Helmke 2004) auf die Schulinspektion zu übertragen und weiterzuentwickeln, damit es als Ausgangspunkt der empirischen Untersuchungen in dieser Arbeit dienen kann.

### **5.1. Verwendung des Helmke-Modells (2004) in Untersuchungen zum Umgang mit Ergebnismeldungen aus Lernstandserhebungen**

Für die Entwicklung eines Modells zur Bestimmung der Ergebnisrezeption und -nutzung von Schulinspektionsergebnissen wird aufgrund analoger Zielsetzungen mit den zentralen Lernstandserhebungen eine Adaption des Rahmenmodells zur pädagogischen Nutzung von Vergleichsarbeiten (Helmke 2004) auf die Schulinspektion vorgenommen.<sup>29</sup> Die externen Evaluationsinstrumente *Lernstandserhebung* und *Schulinspektion* korrespondieren dahingehend, als

---

<sup>29</sup> Das bereits bestehende Modell zu den Auswirkungen von Schulinspektionen von Ehren und Visscher (2006; s. Abschnitt 3.4.6./Abb. 4) wurde bei der Entwicklung eines eigenen Rahmenmodells zur Erfassung der Rezeption und Reflexion von Schulinspektionsergebnissen nur in Ansätzen herangezogen, da es nicht den Rezeptions-Nutzungsprozess, sondern vielmehr eine grobe Wirkungsstruktur der Schulinspektion modelliert.

dass beide Informationen liefern, welche innerschulisch zur Unterrichts- und Schulentwicklung genutzt werden sollen. Hingegen unterscheiden sie sich im Hinblick auf ihr Instrumentarium und Vorgehen. Während Schulinspektionen die Schulqualität mittels standardisierter Verfahren evaluieren, sind Lernstandserhebungen wissenschaftlich konzipierte Testverfahren zur Erhebung des schulfachlichen Kompetenzstands der Schüler.

Die Verwendung des Helmke-Modells bietet neben den Übertragungsmöglichkeiten auf die Schulinspektion und seines Potentials der Weiterentwicklung den Vorteil, die Ergebnismutzung mehrerer Instrumente der Bildungsreform zu vergleichen. Mittlerweile sind mehrere empirische Arbeiten zur Rezeption und Nutzung von Lernstandserhebungen in Anlehnung an das *Rahmenmodell zur pädagogischen Nutzung von Vergleichsarbeiten* von Helmke durchgeführt worden (zuletzt Koch 2011 und Hosenfeld 2010). In diesen Untersuchungen bleibt das Grundgerüst des „Master-Modells“ bestehen, d. h. die einzelnen Phasen (Rezeption – Reflexion – Aktion – Evaluation) sowie die Rahmenbedingungen (individuelle, schulische und externe) werden beibehalten, während die einzelnen Bestandteile des Modells in Abhängigkeit der jeweiligen Forschungsfragen unterschiedlich operationalisiert werden. So wird bspw. die Phase *Rezeption* bei Hosenfeld (2010, S. 68-74) mit den Variablen „Häufigkeit/Intensität des Anschauens der Filme und schriftlichen Rückmeldungen“ messbar gemacht, während Koch diese Phase anhand der Variable „Datenkompetenz“ operationalisiert (vgl. Koch 2011, S. 108-113). Die Intensität der Auseinandersetzung mit den Daten ordnet Koch (2011) bereits der Phase *Reflexion* zu. Des Weiteren zeigt sich insbesondere in den Rahmenbedingungen die Berücksichtigung unterschiedlicher, den Rezeptions- und Nutzungsprozess bedingender Faktoren. Die Operationalisierungen der einzelnen Phasen und Rahmenbedingungen in Anlehnung an Helmkes Rahmenmodell in den Untersuchungen von Hosenfeld (2010) und Koch (2011) sind in Tabelle 9 dargestellt. Die originären Hemkeschen Operationalisierungen der einzelnen Modellbestandteile sind in Abbildung 10 enthalten.

Die Gegenüberstellung der inhaltlich unterschiedlich operationalisierten Phasen in Tabelle 9 soll aufzeigen, dass sich Helmkes Rahmenmodell als Grundgerüst für vielfältige Fragestellungen eignet, grundsätzlich Modifikationen zulässt und dennoch die Vergleichbarkeit der Ergebnisse ermöglicht.

**Tab. 9: Verwendung des theoretischen Rahmenmodells zur Nutzung von Vergleichsarbeiten (Helmke 2004)**

Phase	<b>Hosenfeld: Wirkung von videografischer und schriftlicher Rückmeldung</b> (2010, S. 68-74)	<b>Koch: Nutzung von Ergebnisrückmeldungen aus Vergleichsarbeiten</b> (2011, S. 108-113)
<b>Information</b>	Informationen über Unterricht - Videografierter Unterricht - Schriftliche Rückmeldung (Leistungsentwicklung, Schülerbefragung, Lehrerbefragung, Auswertung der Filme)	Rückmeldungen aus Lernstandserhebungen (Evaluationsbefragungen zu den VERA-Arbeiten der Jahre 2004 bis 2008)
<b>Rezeption</b>	Häufigkeit/Intensität des Anschauens der Filme und der schriftlichen Rückmeldung	Verständnis: Daten abrufen, wahrnehmen und verstehen; Operationalisierung: empirisches Verständnis (Datenkompetenz)
<b>Reflexion</b>	Nachdenken über den eigenen Unterricht	Beschäftigung mit den Daten; - Intensität der Auseinandersetzung (quantitative Komponente) - erlebte Nützlichkeit (qualitative Komponente)
<b>Aktion</b>	(Bereitschaft zu/r) - Veränderung des eigenen Unterrichts - Veränderungen in der Selbstwahrnehmung	Mögliche Aktionen/Abgeleitete Maßnahmen in den Schwerpunkten: - Weiterbildung - Unterrichtsentwicklung - Kooperation - Schulentwicklung
<b>Evaluation</b>	nicht untersucht	nicht untersucht
<b>Rahmenbedingungen</b>	<b>Individuelle Bedingungen:</b> (variable und stabile Lehrermerkmale)  <i>Kognitive/motivationale Variablen</i> (u.a. Vorerfahrung, Interesse, Erwartungen)  <i>„Stabile Lehrermerkmale“</i> (Berufszufriedenheit, Fähigkeit zur Selbstreflexion, Kooperationsbereitschaft, Selbstwirksamkeitserwartung)	<b>Individuelle Bedingungen:</b> - Position des Fachleiters - Akzeptanz gegenüber Schulleistungsstudien/Vergleichsarbeiten - Bisherige Unterrichtsmethoden <b>Schulische Bedingungen:</b> - Kooperation innerhalb der Schule - Kooperation bei der VERA-Durchführung (Durchführungskooperation) - Bisherige Maßnahmen zur Schulentwicklung - Verankerung der externen Evaluation im Schulprogramm <b>Externe Unterstützung:</b> - Beratung/externe Anleitung - Fortbildung

## 5.2. Übertragung des Helmke-Modells (2004) auf das Instrument *Schulinspektion*

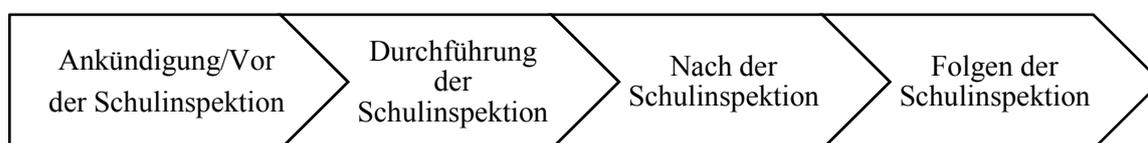
Das Helmkesche Rahmenmodell wurde bereits im Rahmen einer empirischen Arbeit modifiziert und inhaltlich an das Instrument Schulinspektion angepasst: Sommer (2011, S. 150-156) entwickelte ein Inspektionswirkungsmodell sowie ein erweitertes Handlungsphasenmodell der Schulentwicklung nach einer Inspektion und wendete es auf die niedersächsische Schulinspektion an. Dabei wurden die einzelnen Phasen wie folgt operationalisiert:

**Tab. 10: Aspekte des Inspektionswirkungsmodells nach Sommer (2011)**

<i>Information</i>	<ul style="list-style-type: none"> <li>- Inspektionsergebnisse zu 15 Qualitätskriterien, ca. 100 Teilkriterien</li> <li>- mündliche Rückmeldung und Bericht</li> </ul>
<i>Rezeption</i>	<ul style="list-style-type: none"> <li>- Weitergabe des Berichts an die schulischen Gruppen</li> <li>- Lesen des Berichts</li> <li>- Individuelle Auseinandersetzung mit Kritik</li> </ul>
<i>Reflexion</i>	<ul style="list-style-type: none"> <li>- Konferenz/Dienstbesprechung</li> <li>- Vergleich mit Selbsteinschätzung</li> <li>- Suche nach Erklärungen</li> <li>- Beschlussfassung/Zielvereinbarungen</li> <li>- Priorisierung von Maßnahmen</li> </ul>
<i>Aktion</i>	<ul style="list-style-type: none"> <li>- Bilden von Arbeitsgruppen</li> <li>- Umsetzung der Planung</li> <li>- Konzentration auf Schwächen/Stärken</li> <li>- Einholen externer Beratung</li> <li>- Maßnahmen z. B. zur Verbesserung des Schullebens, des Unterrichts etc.</li> </ul>
<i>Evaluation</i>	<ul style="list-style-type: none"> <li>- Evaluation der Maßnahmengestaltung entsprechend den Zielen</li> <li>- Effekte nachweisbar und nachhaltig?</li> </ul>
<i>Individuelle Bedingungen</i>	<ul style="list-style-type: none"> <li>- Selbstwirksamkeitserwartung/Attribuierung von Ergebnissen</li> <li>- Empfundene Belastung durch tägliche Arbeit/Schulreformen</li> <li>- Akzeptanz der Inspektion</li> <li>- Sorge vor der Inspektion</li> <li>- Umfang der Vorarbeiten für die Inspektion</li> <li>- Erwartungen an das Inspektionsergebnis</li> </ul>
<i>Schulspezifische Bedingungen</i>	<ul style="list-style-type: none"> <li>- Personalsituation</li> <li>- Gebäudezustand</li> <li>- Konkurrenz zu anderen Schulen</li> <li>- Führungsstärke der Schulleitung</li> <li>- Kommunikationsstrukturen im Kollegium</li> <li>- belastete vs. aktive Schule (PISA)</li> <li>- Validität der Inspektionssituation</li> <li>- Erfahrungen mit Schulentwicklung/Selbstevaluation/Schulprogrammarbeit</li> <li>- Reputation der Schule im Umfeld</li> </ul>
<i>Externe Bedingungen</i>	<ul style="list-style-type: none"> <li>- Aktivität der Schulaufsicht</li> <li>- Aktivität des Schulträgers</li> <li>- Unterstützungssysteme (Schul-/Unterrichtsentwicklungsberatung)</li> <li>- öffentliche, politische und wissenschaftliche Diskussion zum Bildungssystem</li> </ul>

Das von Sommer erstellte Modell wurde wiederum von Gärtner (2010; 2011) herangezogen, um die Schulinspektion als Rückmeldesystem mit einem Fokus auf schulische Prozesse zu entwickeln. Während Gärtner die prozesshaften Grundelemente sowie die Rahmenbedingungen in Anlehnung an Helmke bzw. Sommer übernimmt, liegt der innovative Charakter seines Rückmeldesystems in den den Ergebnisrückmeldungen vorgelagerten Prozessschritten. Gärtners Modell beginnt somit nicht erst mit der Rückmeldung der Ergebnisse, sondern bereits mit den Reaktionen der Schule auf die Ankündigung der Inspektion. Die der Rückmeldung vorangehenden Phasen haben – so die Hypothese Gärtners – Einfluss auf die Nutzung der Ergebnisse. Dabei konnte Gärtner feststellen, dass sich Wirkungen von Schulinspektionen auf die Schulentwicklung schon in der Vorbereitung auf die Schulinspektion einstellen (vgl. Gärtner 2011, S. 158; Gärtner/Pant 2011, S. 26f.).

Im Folgenden wird das eigens entwickelte Rezeptions-Nutzungs-Modell von Schulinspektionsergebnissen beschrieben (s. Abb. 12). Es bildet in Anlehnung an Helmkes Modell zur pädagogischen Nutzung von Vergleichsarbeiten (2004) sowie Sommers (2011) und Gärtners (2010; 2011) Erweiterungen die Grundlage für die anschließenden empirischen Untersuchungen. Dabei wird der vierstellige Prozess nach der Rückmeldung der Inspektionsergebnisse (Rezeption – Reflexion – Aktion – Evaluation) aufgegriffen und anhand verschiedener Modifikationen ergänzt und weiter ausdifferenziert. Die Phasen *Ankündigung der Schulinspektion* („Vor“) und *Durchführung der Schulinspektion* („Während“) stellen eine Erweiterung des ursprünglichen Modells von Helmke dar (s. Abb. 11 in Anlehnung an Scheerens u.a. 2003, S. 326).



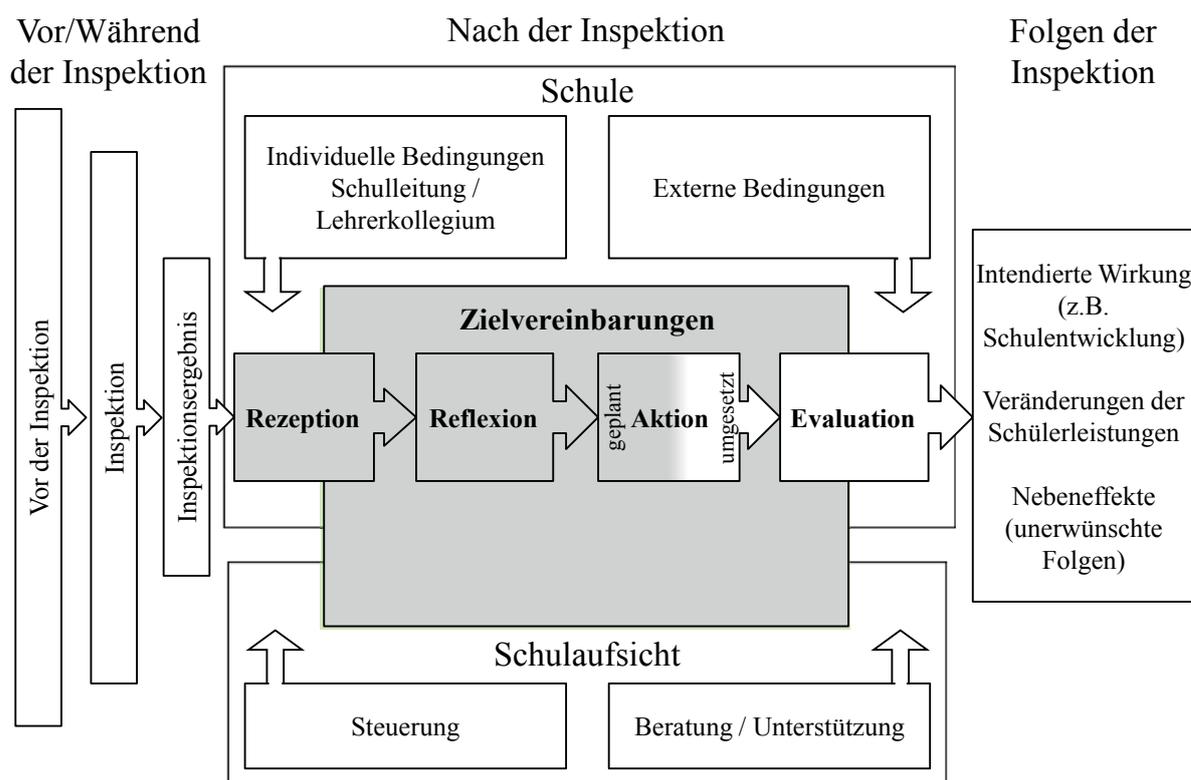
**Abb. 11: Prozessphasen der Schulinspektion**

Das Modell von Helmke und Sommers Anpassung des Modells an die Schulinspektion konzentrieren sich primär auf die Phase nach der Ergebnisrückmeldung. Dies ist insofern folgerichtig und plausibel, als dass die pädagogische Nutzung der Ergebnisse formal erst mit der Ergebnisrückmeldung einsetzt. In

diesen Modellen bleiben jedoch die der Rückmeldung vorgelagerten Phasen, z. B. spezifisch für die Schulinspektion die Reaktionen der Schulen auf die Ankündigung (übernommen von Gärtner 2010), entsprechende Vorbereitungsaktivitäten oder die Durchführung der Schulinspektion, unberücksichtigt. Zudem erfährt das Modell eine Ergänzung um die Phase *Folgen der Schulinspektion*, welche aus dem Rahmenmodell zur Wirkung von Schulinspektionen nach Ehren und Visscher (2006) adaptiert ist. Diese Phase ist insofern bedeutend, als dass die Schulinspektion per se nicht mit der Phase *Evaluation* endet, sondern aus ihr sowohl intendierte Effekte als auch unerwünschte Nebeneffekte resultieren.

Die Phase der *Aktion* erfährt im Vergleich zum ursprünglichen Helmke-Modell die größte Modifikation, indem sie in zwei getrennte Phasenabschnitte untergliedert wird. Die rezipierten und reflektierten Ergebnisse münden zunächst in *geplante* und schließlich in tatsächlich *umgesetzte Maßnahmen*. Diese Unterteilung wird dadurch legitimiert, dass die in den Zielvereinbarungen fixierten Maßnahmenplanungen als erste Schritte der Nutzung interpretiert werden, da sie durch ihre schriftliche Konkretisierung die Auseinandersetzung mit den Ergebnissen und bereits abgeschlossene Reflexionsprozesse voraussetzen. Im Gegensatz zu Sommers Einordnung werden Zielvereinbarungen im eigenen Modell nicht der Reflexionsphase zugeordnet.

Eine weitere Modifizierung stellt die Ergänzung eines neuen Akteurs – der Schulaufsicht – dar. Dadurch wird die Beschränkung auf den Akteur Schule in den vorherigen Modellen aufgebrochen. Dies ist insofern gefordert, als dass die Inspektionsergebnisse beiden Akteuren zur Verfügung gestellt werden und der Abschluss von Zielvereinbarungen durch eine kooperative Zusammenarbeit gekennzeichnet ist. Es ist nicht möglich, von einer gemeinsamen Rezeption und Reflexion auszugehen, da die jeweiligen Institutionen vorwiegend unabhängig voneinander die Ergebnismeldungen rezipieren, reflektieren und nutzen. Der Akteur Schulaufsicht wird zwar im originären Rahmenmodell von Helmke (2004) sowie im Modell von Sommer (2011) mitberücksichtigt, allerdings wird die „Unterstützung durch die Schulaufsicht“ den externen schulischen Rahmenbedingungen zugeordnet. Bei der Entwicklung eines eigenen Modells wird die Notwendigkeit gesehen, die Schulaufsicht als zweite aktive Protagonistin aufzunehmen.



**Abb. 12: Eigenes Rezeption-Nutzungs-Modell von Schulinspektionsergebnissen**

Dadurch, dass zwei unterschiedliche Perspektiven – Schulen und Schulaufsicht – modelliert werden, ist der vierschriftige Prozess im Anschluss an die Ergebnisrückmeldung differenziert zu betrachten. Die Zielvereinbarungen stellen das verbindende Gelenk zwischen der Schule und der Schulaufsicht dar und umschließen den vierschriftigen Prozess. Ergänzt wird der Prozess um Rahmenbedingungen aus schulischer und schulaufsichtlicher Perspektive. Die Identifizierung dieser Bedingungen stellt noch ein Desiderat dar und soll im empirischen Teil dieser Arbeit eruiert werden.

Abschließend erfolgt die kritische Hinterfragung des herangezogenen Helmkeschen Modells. Das Rahmenmodell zeichnet sich durch den Vorteil aus, variabel für vielfältige Fragestellungen eingesetzt werden zu können. Gleichzeitig birgt es jedoch auch einige diskussionswürdige Schwachstellen, die es bei seiner Weiterentwicklung und Erweiterung zu berücksichtigen gilt: Es projiziert ein abstraktes, idealisiertes Abbild der Realität und erfasst aufgrund des begrenzten Wissens über den analytischen Prozess und die Wirkmechanismen der Rahmenbedingungen auf die einzelnen Phasen nur selektive und beschränkte Merkmale des latenten Rezeptions-Nutzungsprozesses. Die Vorbereitungs- und

Durchführungsphase der Schulinspektion, die Einfluss auf die Rezeption und Nutzung der Ergebnisse nehmen können, werden in diesem Prozess nicht berücksichtigt. Auch welche Bedeutung die einzelnen Phasen und Rahmenbedingungen im Gesamtprozess spielen, bleibt offen. Das Modell trennt zudem nicht zwischen Prozess und Ergebnis. Im Vordergrund steht die Abfolge aufeinander bezogener Prozesse, deren alleinige Betrachtung nicht ausreicht (vgl. Böttger-Beer/Vaccaro/Koch 2010, S. 333): „Es muss genauso untersucht werden, unter welchen Bedingungen ein bestimmter Zwischenprozess *das* Zwischenergebnis hervorbringt, das für die Fortsetzung der Wirkung notwendig ist.“ Zudem wird eine deutliche Abgrenzung der einzelnen Phasen nicht erkennbar: Wann genau gilt eine Phase als abgeschlossen und wann beginnt die nächste?

Diese Schwächen des Modells einkalkulierend, wird der Gewinn der Modellverwendung darin gesehen, dass es zu erklären versucht, wie komplex und strukturiert der Vorgang von der Ergebnisrückmeldung zur Nutzung verläuft und durch welche Komponenten die Nutzung der Inspektionsergebnisse gekennzeichnet ist. Das Modell bildet ein Gerüst für die Erkenntnisgewinnung im Hinblick auf die Rezeption und Nutzung von Schulinspektionsergebnissen, das es durch empirische Untersuchungen auszubauen gilt.

## 6. Ableitung von Forschungsfragen

Das Ziel des vorangegangenen Theoriekapitels war die umfassende Darstellung der Schulinspektion und ihre Einbettung in das Schulqualitätsverständnis, das neue Steuerungsparadigma und die verschiedenen Evaluationsformen. Zudem wurden Schwerpunkte und empirische Befunde der aktuellen Schulinspektionsforschung in Deutschland aufgezeigt. Deutlich gemacht wurde des Weiteren, dass die Auseinandersetzung mit der Schulinspektion zwangsläufig in Fragen der Rezeption und Rezeptionsforschung mündet, da nur eine gelungene Rezeption der zurückgemeldeten Ergebnisse die intendierten Impulse zur Schulentwicklung zu geben vermag. Auf der Grundlage des Rezeptionsforschungsstands wurde deutlich, dass im Allgemeinen die Rückmeldung von extern erhobenen schulischen Leistungsdaten nicht automatisch zu Schul- und Unterrichtsentwicklungsprozessen führt. Spezifisch für die Schulinspektion erweist sich die empirische Forschungslage zur Rezeption und Nutzung zurückgemeldeter Schulinspektionsergebnisse als defizitär.

Vor dem Hintergrund der identifizierten Forschungslücken werden vier Forschungsfragen formuliert, die auf der Grundlage des eigens entwickelten Rezeption-Nutzungs-Modells von Schulinspektionsergebnissen (s. Kapitel 5) geprüft werden sollen. Die übergeordnete Fragestellung zur Rezeption und Nutzung von Schulinspektionsergebnissen wird folgendermaßen präzisiert:

1. Wie rezipieren und reflektieren Schulen die Ergebnismeldungen der Schulinspektion?
2. In welcher Weise spiegeln sich diese Reflexionen in den erarbeiteten Zielvereinbarungen wider, d.h. wie viele und welche Entwicklungsmaßnahmen werden geplant und in welchem Zusammenhang stehen sie zu den Ergebnissen der Schulinspektion?
3. Wie stellt sich der Umgang mit den Ergebnissen der Schulinspektion aus der Perspektive der Schulaufsicht dar und in welcher Weise unterstützt und berät sie die Schulen bei der Maßnahmenplanung?
4. Lässt sich aus den verschiedenen Perspektiven auf den Rückmeldeprozess ein allgemeines Modell der Rezeption und Nutzung von Ergebnissen aus Schulinspektionen ableiten?



### III EMPIRISCHER TEIL

---

#### 7. Methodisches Vorgehen: Triangulation

Die empirische Untersuchung des Forschungsgegenstands soll auf der Grundlage der Qualitätsanalyse Nordrhein-Westfalen bearbeitet werden. Dazu eignet sich methodisch der Einsatz eines triangulativen Verfahrens, d. h. der kombinierten Verwendung unterschiedlicher Datenquellen und Auswertungsmethoden. Grundsätzlich eröffnen sich durch den Einsatz unterschiedlicher Forschungsmethoden, Datenquellen etc. vielfältige Verknüpfungsmöglichkeiten. Zu den aktuelleren Entwicklungslinien in der Erziehungswissenschaft und der Schulentwicklungsforschung zählt die Kombination qualitativer und quantitativer Verfahren (vgl. Gläser-Zikuda 2010, S. 154; Krüger/Pfaff 2008, S. 157). Insbesondere die Triangulationsforschung der vergangenen Jahre hat wesentlich dazu beigetragen, die Kombination beider Verfahren sowohl aus praktischen als auch forschungsmethodischen Erwägungen zunehmend einzufordern und progressiv zu praktizieren. Die Verknüpfung verschiedener Methoden wird oft mithilfe der Triangulationsmetapher beschrieben oder auch als Methodenintegration, Methodenkombination oder Methodenmix bezeichnet (vgl. Kelle/Erzberger 2009, S. 302). Rost (2007, S. 10) fasst die hinter diesen Bezeichnungen liegende Konzeption als „banale Erkenntnis [zusammen], dass man gut beraten ist, möglichst mehrere Methoden/Ansätze/Vorgehensweisen/ Theorien etc. miteinander zu kombinieren, will man besonders aussagekräftige Resultate erhalten“. Der Begriff *Triangulation* – einer ursprünglich aus der Navigation und Landvermessung entlehnten Messmethode – wird in der eigenen Untersuchung in Anlehnung an die Definition Flicks (2010, S. 281) verstanden:

Triangulation beinhaltet die Einnahme unterschiedlicher Perspektiven auf einen untersuchten Gegenstand oder allgemeiner: bei der Beantwortung von Forschungsfragen. Diese Perspektiven können in unterschiedlichen Methoden, die angewandt werden und/oder unterschiedlich gewählten theoretischen Zugängen konkretisiert werden, wobei beides wiederum miteinander in Zusammenhang steht bzw. verknüpft werden sollte.

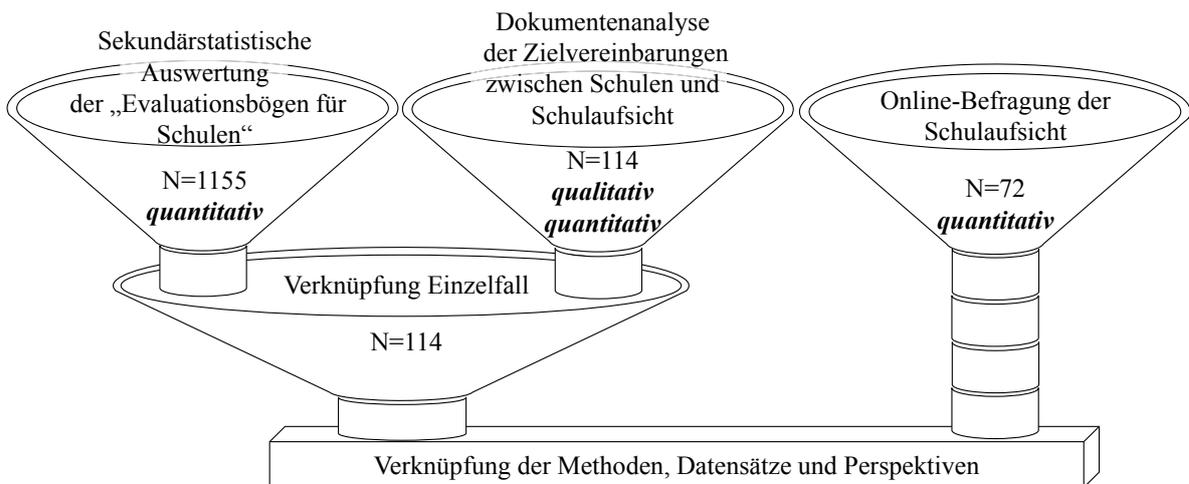
Der Gewinn eines triangulativen Untersuchungsdesigns liegt in der „Herstellung eines kaleidoskopartigen Gesamtbildes von einem Gegenstand“ (Krüger/Pfaff 2008, S. 160). Diese komplementäre Begründungsargumentation strebt vordergründig eine Erkenntniserweiterung durch eine mehrperspektivische und multimethodische Erfassung eines Gegenstandes an. Insbesondere in der neueren Literatur zur Triangulation wird überwiegend der Stellenwert der Komplementaritätsmodelle akzentuiert (vgl. Flick 2008, S. 112; Kelle/Erzberger 2009, S. 304; Krüger/Pfaff 2008, S. 174).

Es existieren mehrere Verknüpfungsmöglichkeiten der Methoden, Daten und Ergebnisse in Triangulationsuntersuchungen, die auf unterschiedliche Weise parallel oder sequentiell eingesetzt werden können (vgl. Flick 2008, S. 80-84). Gemäß der am häufigsten zitierten und diskutierten Konzeption nach Denzin lassen sich vier verschiedene Formen der Triangulation unterscheiden (vgl. Denzin 1970; 1989; Flick 2008, S. 13-17):

1. *Daten-Triangulation*: Einbeziehung unterschiedlicher Datenquellen, die mit unterschiedlichen Personen, zu unterschiedlichen Zeitpunkten oder verschiedenen Orten erhoben wurden.
2. *Investigator- (Forscher-)Triangulation*: Einsatz unterschiedlicher Beobachter oder Interviewer zur Dezimierung von Verzerrungen durch die Forscherperson.
3. *Theorien-Triangulation*: Einbeziehung von verschiedenen Theorien und Hypothesen zur Erklärung eines Phänomens.
4. *Methoden-Triangulation*: Denzin unterscheidet zwischen „within-method“, d. h. Triangulation innerhalb einer Methode und „between-method“, d. h. Triangulation zwischen verschiedenen Methoden.

Insgesamt können die einzelnen Methoden und Daten unabhängig voneinander angewandt und erhoben und erst auf der Ebene der Ergebnisse trianguliert werden. Konsequenter an die Idee der Triangulation angelehnt zeigt sich die verknüpfende Auswertung unterschiedlicher Datensorten und methodischer Zugänge am Einzelfall.

Bezogen auf die Datenquellen und Methoden der vorliegenden Untersuchung wird ein Triangulationsmodell konstruiert, welches grafisch in einem Trichtermodell abgebildet ist (s. Abb. 13).



**Abb. 13: Untersuchungsdesign „Trichtermodell Triangulation“ zur Rezeption und Nutzung von Rückmeldungen der Qualitätsanalyse NRW**

Es werden unterschiedliche Triangulationstypen mit einer Fokussierung der Daten- und Methodenkombination herangezogen. Wie im Trichtermodell veranschaulicht, kommen drei unterschiedliche Datenquellen zum Einsatz: Die schulischen Evaluationsbögen, die Zielvereinbarungen und die Befragungsergebnisse der Schulaufsicht. Die Triangulation bezieht sich zudem auf die Verbindung verschiedener Forschungsmethoden: Die Evaluationsbögen werden sekundärstatistisch ausgewertet, bei der Analyse der Zielvereinbarungen wird die quantitativ-qualitative Inhaltsanalyse eingesetzt und schließlich wird die Schulaufsicht schriftlich befragt. Die Einbeziehung der „*between-method-Triangulation*“ dient in dieser Untersuchung zur Ausschöpfung verschiedener Auswertungsstrategien, zur Überwindung begrenzter Erkenntnismöglichkeiten, zur Minimierung der jeweiligen Schwächen und Verzerrungspotentiale der Einzelmethoden sowie zur Prüfung von Zusammenhängen. Das Prinzip der *Forscher-Triangulation* wird im Rahmen der Zielvereinbarungsanalyse verwendet, um bei der Codierung die Einflussnahme eines einzigen Forschers zu verringern. Die Nutzung der *Theorien-Triangulation* findet im theoretischen Teil dieser Arbeit ihre Begründung und mündet in das modifizierte Modell (s. Kapitel 5). Im vorliegenden Untersuchungsdesign setzt die Triangulation einerseits am Einzelfall an, indem von den Schulen, deren Zielvereinbarungsdokumente vorliegen, ebenfalls die entsprechenden Evaluationsbögen herangezogen werden können. Andererseits setzt durch den Einbezug der Schulaufsichtsbefragung die Triangulation erst an der Schnittstelle der Ergebnisse an.

Im Anschluss an die methodische Einbettung der Untersuchung sollen an dieser Stelle die drei Teiluntersuchungen separat skizziert werden, bevor sie jeweils eingehender im empirischen Teil der Arbeit dargestellt werden. Das Erkenntnisinteresse der folgenden Studien liegt in der Identifikation und Verknüpfung der auf die Schulinspektion übertragenen Rezeptions- und Nutzungsphasen in Anlehnung an das Rahmenmodell von Helmke (2004). Es wird die Frage aufgegriffen, in welcher Weise die Schulen und die Schulaufsicht die Ergebnisse der Schulinspektion rezipieren und wie sie darauf handelnd reagieren. Die unterschiedlichen Studien werden wie folgt konzipiert:

**Studie I:** *Sekundärstatistische Analyse der schulischen Evaluationsbögen (n = 1155)*

Studie I widmet sich mit einem sekundärstatistischen Zugang der schulischen Rezeption und Reflexion der Inspektionsergebnisse. Die Datengrundlage zur Erfassung der schulischen Rezeptions- und Nutzungsprozesse der Ergebnisse aus Qualitätsanalysen stellen ministeriell administrierte Evaluationsbögen dar. Neben der Identifikation einzelner Phasen des in Kapitel 5 entwickelten Rezeptions-Nutzungs-Modells sollen insbesondere die Beziehungen zwischen den einzelnen Phasen überprüft werden.

**Studie II:** *Inhaltsanalyse der Zielvereinbarungen (n = 114)*

Studie II beinhaltet eine inhaltsanalytische Untersuchung der Zielvereinbarungen. Mit der Auswertung der Zielvereinbarungen sollen Aussagen zur geplanten Nutzung der Ergebnisse aus Qualitätsanalysen aus schulischer und schulaufsichtlicher Perspektive getroffen werden. Dabei sollen anhand eines deduktiv erzeugten Kategoriensystems formale und inhaltliche Merkmale sowie die Zielqualität der Dokumente bestimmt werden. Zusätzlich soll der Zusammenhang zwischen den Zielinhalten und den festgestellten Schwächen der Schulen überprüft werden.

**Studie III:** *Online-Befragung der Schulaufsicht (n = 72)*

Studie III konzentriert sich auf eine quantitative Befragung der Schulaufsicht zur Rezeption, Reflexion und Nutzung der Schulinspektionsergebnisse. Das Ziel einer standardisierten Befragung der schulfachlichen Deponenten der oberen und unteren Schulaufsicht liegt zudem in einer Ergänzung der schulischen Sicht.

### **Triangulation der Studien I, II und III**

Jede dieser Studien akzentuiert ein spezifisches Segment der Fragestellungen und ist forschungsorganisatorisch auf die anderen Studien abgestimmt, um die verschiedenen Datenquellen in der übergreifenden Auswertung zusammenführen zu können. Keine Studie allein könnte auf das Forschungsdesiderat, die Rezeption und Nutzung der Ergebnisse umfassend abzubilden, angemessen reagieren. Jede dieser Teiluntersuchungen liefert einen eigenständigen Beitrag dazu, dass die einzelnen Phasen sowie Perspektiven (Schule und Schulaufsicht) des modifizierten Helmke-Modells in einer triangulativen Gesamtdarstellung inhaltlich angereichert werden können.

Im Folgenden werden die einzelnen Teiluntersuchungen zunächst separat hinsichtlich ihrer Methode, ihrer Durchführung und ihrer Ergebnisse dargestellt.



## **8. Studie I: Sekundäranalyse der Evaluationsbögen der Qualitätsanalyse**

Im Rahmen der Studie I wird eine Sekundäranalyse der Evaluationsbögen der Qualitätsanalyse NRW durchgeführt. Die Daten aus der Primärerhebung zur Evaluation der Qualitätsanalyse NRW aus schulischer Perspektive werden unter einer modellgeleiteten Fragestellung in Anlehnung an das erweiterte Modell nach Helmke (2004; s. Kapitel 5) sekundärstatistisch ausgewertet. Im Zentrum steht die Frage, ob sich die einzelnen Phasen des Rahmenmodells durch statistische Analysen identifizieren und unterscheiden lassen und – unter der Annahme, dass dies gelingt – welche Ausprägungen und Zusammenhänge sie aufweisen. Zusätzlich wird danach gefragt, ob sich weitere und nicht im Modell enthaltene Phasen und Rahmenbedingungen aufdecken lassen.

### **8.1. Datengrundlage und methodisches Vorgehen**

Die Datengrundlage der Sekundäranalyse bilden Evaluationsbögen der Qualitätsanalyse NRW (s. zum Evaluationsbogen: MSW NRW 2008e)<sup>30</sup>. Die Bereitstellung der Daten war freundlicherweise nach Genehmigung des MSW NRW als Dateninhaber realisierbar. Die Sekundäranalyse zeichnet sich gegenüber einem klassischen Forschungsprozess durch die Entkoppelung der Datenauswertung und -interpretation von der Datenerhebung aus (vgl. Kromrey 2002, S. 526; Klingemann/Mochmann 1975, S. 178; Friedrichs 1990, S. 359; Schnell u.a. 2008, S. 251). Der Einsatz von Sekundäranalysen ist somit aus forschungsökonomischen, darunter finanziellen und zeitlichen Gründen vorteilhaft.

In der Erziehungswissenschaft eignen sich neben den Leistungsvergleichsstudien in vielen Fällen auch amtliche Daten zur erneuten Auswertung (vgl. Böhm-Kasper/Weishaupt 2008, S. 93). Amtliche Daten zeichnen sich i. d. R. durch einen großen Stichprobenumfang aus, weisen einen geringen Stichprobenfehler auf und sind nur in geringem Ausmaß von Unit-Non-Response betroffen, da häufig eine Teilnahmepflicht besteht und die Daten oftmals auf

---

<sup>30</sup> Der Evaluationsbogen ist auch einsehbar unter: <https://www.schulministerium.nrw.de/docs/Schulentwicklung/Qualitaetsanalyse/Stichworte/Evaluationsbogen/Evaluationsbogen.doc> (18.03.2015).

Vollerhebungen beruhen (vgl. Wirth/Müller 2006, S. 97f.). Zudem sind sie i. d. R. durch eine hohe Kontinuität gekennzeichnet, da sie anhand von identischen Instrumenten über meist längere Zeiträume gesammelt werden. Neben den genannten Vorteilen, die im Zusammenhang mit der Nutzung bereits vorhandener Daten zu sehen sind, ist in der heutigen Zeit im Hinblick auf die Konjunktur der Forschungsthemen *Schule* und *Evaluation* ein umsichtiger Umgang mit Möglichkeiten der Datenerhebung gefordert. Im vorliegenden Fall erschien es sinnvoll, auf vorhandene Daten zurückzugreifen, da die schulischen Zielgruppen so vor der Überfragung, welche in den letzten Jahren vermehrt von Seiten der Bildungsinstitutionen beklagt wird, und der daraus resultierenden Forschungsmüdigkeit und -resistenz, verschont bleiben.

Als amtliche Daten sind die Evaluationsbögen im Auftrag des MSW NRW als offizieller Institution primär für politisch-administrative Zwecke erhoben worden. Für die Sekundäranalyse war neben den forschungsökonomischen Vorteilen insbesondere das hohe Potential der Datengrundlage ausschlaggebend: Die Evaluationsbögen bieten eine reichhaltige Quelle an relevantem Forschungsmaterial für die Fragestellungen dieser Arbeit, da die Stichprobe durch einen großen Umfang und durch eine zeitnahe Beantwortung der Fragebögen gekennzeichnet ist. Dies hätte mit einer eigenen Befragung nicht erreicht werden können, da für einige Schulen die Qualitätsanalyse bereits einige Jahre zurückliegt und somit die Qualität der Daten durch eine länger zurückliegende retrospektive Einschätzung der Qualitätsanalysen nicht gewährleistet wäre.

### **8.1.1. Beschreibung und Charakterisierung der Datengrundlage der Sekundäranalyse**

Zur Klärung der Nutzbarkeit der vorliegenden Sekundärdaten wird die Bestimmung der Zielsetzung der Primäranalyse, ihrer Inhalte, Methoden sowie der Aktualität und Qualität der erhobenen Daten vorgenommen (vgl. die leitfragengestützte Anleitung und die Empfehlung in Medjedović 2010, S. 311; Kromrey 2002, S. 526; Friedrichs 1990, S. 355; Schnell u.a. 2008, S. 251).

Zur Evaluation der Qualitätsanalyse setzt das MSW NRW seit dem Schuljahr 2006/07 einen *Evaluationsbogen für Schulen* ein, der gemeinsam mit dem Institut für Schulentwicklungsforschung (IFS an der Technischen Universität Dortmund) entwickelt worden ist. Ziel des Evaluationsbogens ist es, aus der

Perspektive der besuchten Schulen Rückmeldungen zur Qualitätsanalyse zu erhalten, um der Administration Aufschlüsse zur Systemüberprüfung und -steuerung sowie für die Weiterentwicklung und Optimierung des Verfahrens und der Instrumente zu geben. Zudem dient sie den Qualitätsprüfern, welche neben dem MSW NRW als Adressaten der Evaluationsbögen anzusehen sind, als Feedback. Alle an der Qualitätsanalyse beteiligten Schulen werden im Anschluss an die Qualitätsanalyse zeitnah anhand eines Fragebogens schriftlich zu ihren Einschätzungen bezüglich der Prozesse der Qualitätsanalyse – Vorgespräch, Schulbesuch und Bericht – befragt. Es ist vorgesehen, dass der Fragebogen von der Schulleitung in Absprache mit den schulischen Gremien ausgefüllt wird. Die Evaluationsbögen wurden vom IFS an die beteiligten Schulen versandt und nach der Rücksendung kontinuierlich in das Statistikprogramm SPSS eingegeben und ausgewertet. Der Fragebogen befindet sich seit dem Schuljahr 2006/07 im Einsatz und wurde seitdem an einigen Stellen modifiziert und optimiert, was zwar zur Verbesserung des Instruments beiträgt, allerdings eine einheitliche Auswertung – sowohl in einer Primär- als auch Sekundäranalyse – erschwert. Die dieser Untersuchung zugrunde liegenden Daten wurden in den Schuljahren 2006/07 bis 2010/11<sup>31</sup> im Rahmen der Evaluation der Qualitätsanalyse NRW mittels Fragebögen erhoben. Insgesamt enthält der Rohdatensatz 77 Variablen: 60 geschlossene Fragen und 17 String-Variablen mit offenen Antwortmöglichkeiten. Zudem wird Raum für Stellungnahmen der Lehrer-, Schüler-, und Elternvertretungen durch eigens dafür vorgesehene Kommentarfelder gegeben. Der aktuellste Stand (22.06.2010) des Fragebogens enthält folgende, am Ablauf der Qualitätsanalyse orientierte Themenkomplexe:

- I. Angaben zur Schule, zum Schulbesuch, dem Qualitätsteam und der evaluationsbogenausfüllenden Person/Personengruppe
- II. Vorgespräch und Vorbereitung des Schulbesuchs
- III. Durchführung der Schulbesuchstage
- IV. Bericht
- V. Ertrag der Qualitätsanalyse
- VI. Anmerkungen oder Kommentare der Schulleitung, Schulpflegschaft, Lehrerkonferenz, Schülervertretung

---

<sup>31</sup> Im Schuljahr 2010/11 wurde nur das erste Schulhalbjahr berücksichtigt; Stand: 02/2011.

Die Evaluationsbögen wurden zum überwiegenden Teil von der Schulleitung (87 %), in jeweils 6 % von der Steuergruppe und dem gesamten Kollegium und lediglich in 1 % der Fälle von einzelnen Lehrkräften ausgefüllt. Der Rohdatensatz beinhaltet Daten von 1155 Schulen, darunter den Schulformen Grundschule, Förderschule, Hauptschule, Realschule, Gesamtschule, Gymnasium und Berufskolleg. Der Anteil der Stichprobe an der Grundgesamtheit<sup>32</sup> beträgt 19 %. Ein Überblick über die Verteilung der Schulformen in der Stichprobe und der Grundgesamtheit wird in Tabelle 11 gegeben. Die Schulform Grundschule stellt die im Rohdatensatz, aber auch in der Grundgesamtheit häufigste Schulform dar. Vergleichsweise gering ist jeweils der Anteil der Gesamtschulen und Berufskollegs. Die heterogene Verteilung der Schulen zu den einzelnen Schulformen entspricht in etwa der Verteilung der Schulformen in der Population.

**Tab. 11: Verteilung der Schulformen in der Stichprobe der Sekundäranalyse und der Grundgesamtheit in der Amtlichen Schulstatistik NRW**

Stichprobe Sekundäranalyse					Grundgesamtheit nordrhein-westfälischer Schulen*	
Schulform	absolut	prozentual	Erwartete Anzahl	Standardisierte Residuen <sup>a</sup>	absolut	prozentual
<b>GS</b>	446	38,6 %	587,8	<b>-5,8</b>	3228	53,2 %
<b>HS</b>	143	12,4 %	134,2	0,8	696	11,5 %
<b>RS</b>	139	12,0 %	104,0	<b>3,4</b>	511	8,4 %
<b>GE</b>	57	4,9 %	41,4	2,4	202	3,3 %
<b>GY</b>	119	10,3 %	101,9	1,7	518	8,5 %
<b>FS</b>	160	13,9 %	129,1	2,7	647	10,8 %
<b>BK</b>	91	7,9 %	56,5	<b>4,6</b>	262	4,3 %
<b>Gesamt</b>	1155	100,0 %	1155,0		6064	100,0 %

\*Amtliche Schuldaten des Schuljahres 2008/09 (N = 6064): Grundgesamtheit der öffentlichen Schulen in NRW; Quelle: MSW NRW 2009b, S. 11.

a: Die fett markierten standardisierten Residuen sind auf einem Signifikanzniveau von  $p = ,001$  signifikant.

GS = Grundschule; HS = Hauptschule; RS = Realschule; GE = Gesamtschule; GY = Gymnasium; FS = Förderschule; BK = Berufskolleg

<sup>32</sup> Die Grundgesamtheit setzt sich zusammen aus allen öffentlichen Schulen des Landes NRW im Schuljahr 2008/09. Dieses Schuljahr wurde herangezogen, da es intermediär die Schuljahre 2006/07 bis 2010/11 in zwei Hälften teilt.

Eine Prüfung der statistischen Repräsentanz der Stichprobe zu der Grundgesamtheit in der amtlichen Schulstatistik des Landes NRW ergab hinsichtlich des Merkmals *Schulform* einen negativen Befund. Der Anteil der Schulform Grundschule liegt in der Stichprobe niedriger als erwartet, während alle anderen Schulformen häufiger als erwartet vorkommen (s. Tab. 11). Der Chi-Quadrat-Test nach Pearson liefert ein signifikantes Ergebnis ( $p = ,000$ ) und die standardisierten Residuen zeigen signifikante Abweichungen ( $p < ,001$ ) in den Schulformen Grundschule, Realschule und Berufskolleg auf.

Auf der Grundlage der Primärdaten sind bereits mehrere Veröffentlichungen<sup>33</sup> entstanden. Dabei liegt der Datenumfang der Sekundäranalyse ( $n = 1155$ ) um das Dreifache über der Anzahl bislang ausgewerteter Evaluationsbögen der Primärerhebung ( $n = 396$ ) und stellt die folgenden Berechnungen und Ergebnisse auf eine solide Basis. Während in der Primärerhebung die Daten in erster Linie deskriptiv ausgewertet wurden, soll bei der Sekundäranalyse das Instrumentarium der multivariaten Statistik zur Anwendung kommen.

Zusammenfassend ist hinsichtlich der Beschreibung der Datengrundlage festzuhalten, dass der Datensatz – mit der Einschränkung der Kontinuität – durch eine hohe Datenqualität sowie -aktualität gekennzeichnet ist, sich allerdings hinsichtlich der Schulformverteilung als nicht repräsentativ erweist. Die Reflexion, inwiefern die Untersuchungsziele der Primär- und Sekundäranalyse übereinstimmen, führt zu einem positiven Befund. Dabei bildet die Sekundäranalyse durch eine modellgeleitete Untersuchung eine Ergänzung und Erweiterung der Untersuchungsintention der Evaluationsbefragung und stellt damit keine widersprüchliche oder thematisch gänzlich ungeeignete Herangehensweise an die Daten dar. Insgesamt ist die Nähe der Fragestellungen zwischen Primär- und Sekundärerhebung eine gute Bedingung für die geplanten Berechnungen und trägt dazu bei, die vorhandenen Daten im Rahmen einer Sekundäranalyse mit einem höheren Grad an Detailliertheit auszuwerten. Die inhaltliche Ausrichtung der Evaluationsbögen lässt analoge Themenbereiche des zugrunde gelegten Modells in Anlehnung an Helmke (2004) und eine Extraktion inhaltlich passender Komponenten erwarten.

---

<sup>33</sup> Ergebnisse auf der Grundlage der Evaluationsbögen der Qualitätsanalyse NRW finden sich exemplarisch in folgenden Publikationen: Müller 2008a ( $n = 35$  Schulen); MSW NRW 2009a, S. 17-21 und Müller 2010 (beide:  $n = 396$  Schulen; Frühjahr 2009).

### **8.1.2. Fragestellungen**

Die Studie I geht der folgenden zentralen Fragestellung nach:

- Lassen sich auf der Datengrundlage der Evaluationsbögen die einzelnen Phasen des weiterentwickelten Rahmenmodells faktorenanalytisch als unterscheidbare Dimensionen identifizieren?

Unter der Annahme des Gelingens werden folgende daran anknüpfende Fragen aufgeworfen:

- Wie lassen sich die identifizierten Faktoren inhaltlich beschreiben?
- Welche deskriptiven Kennwerte weisen die identifizierten Faktoren auf?
- Welche Zusammenhänge bestehen zwischen den einzelnen Faktoren?
- Lassen sich die postulierten Beziehungen des Modells überprüfen?

### **8.1.3. Methodisches Vorgehen und Analyseverfahren**

Zur Beantwortung der Fragestellungen wird ein hybrides Vorgehen gewählt: In einem ersten Schritt wird eine explorative Faktorenanalyse (EFA) durchgeführt und die Passung der identifizierten Faktorenstruktur zu den empirischen Daten in einem zweiten Schritt konfirmatorisch (CFA) überprüft.

Im Rahmen der Datenbereinigung wurden 27 Variablen durch einen zu hohen Anteil fehlender Werte (zwischen 17 % und 93 %) anlässlich der Modifikation des Evaluationsbogens für die weiterführenden Analysen ausgeschlossen, während vier inhaltlich redundante Variablen mit einem ebenfalls hohen Anteil fehlender Werte zu zwei Variablen zusammengefasst werden konnten. Um einer Reduktion der Daten und des daraus resultierenden Informationsverlustes vorzubeugen, wurden die fehlenden Werte unter der Zehnprozentgrenze der verbliebenen (quasi)metrischen Variablen anhand eines Imputationsverfahrens (EM-Algorithmus) ersetzt (vgl. Bühner 2011, S. 241; Schendera 2007; Baltès-Götz 2008; Reinecke 2005, S. 288). Die dafür erforderliche MCAR-Bedingung (Missing completely at Random) wurde mittels MCAR-Tests nach Little überprüft ( $p > ,001$ ) und konnte bestätigt werden (vgl. Baltès-Götz 2008, S. 6f.).

Zur Ermittlung der Faktorenstruktur in den Daten wurde eine explorative Faktorenanalyse (EFA) durchgeführt (vgl. Rost 2007, S. 199). Dabei wurde mit 32 Variablen eine Hauptkomponentenanalyse (Principal Component Analysis, PCA) mit anschließender orthogonaler Rotation (Varimax) durchgeführt (vgl.

Bühner 2011, S. 309-318, 349, 372; Rost 2007, S. 205). Zuvor wurde die Eignung des vorliegenden Datensatzes für die Berechnung einer Faktorenanalyse mit positiven Resultaten geprüft.<sup>34</sup> Zur Bestimmung der Faktorenanzahl wurde das Eigenwertkriterium größer eins herangezogen (vgl. Bühner 2011, S. 321f.). Die Zuordnung der Variablen zu den extrahierten Faktoren erfolgt auf der Basis der Ladungshöhe der Variablen ( $\geq ,50$ ; vgl. Backhaus u.a. 2010, S. 362).

Zum Ausschluss möglicher zufälliger Effekte und zur Konstruktvalidierung wurde in einem weiteren Schritt die konfirmatorische Faktorenanalyse angeschlossen. Die konfirmatorische Faktorenanalyse dient dazu, „theoretisch oder empirisch gut fundierte Modelle auf ihre empirische Passung mit den Daten hin zu testen oder mit alternativen Modellen zu vergleichen“ (Bühner 2011, S. 380). Es existieren verschiedene Kriterien, sog. Fitindizes, zur Beurteilung der Güte der Passung der Daten mit einem Modell (vgl. Geiser 2010, S. 60-62; Schermelleh-Engel/Moosbrugger/Müller 2003). Die Fit-Indizes sind bei der Beurteilung der Modellgüte neben theoretischen Abwägungen als ungefähre Daumenregeln zu interpretieren (vgl. Schermelleh-Engel u.a. 2003, S. 52f.). Um die beiden faktorenanalytischen Prüfungen nicht anhand derselben Stichprobe durchzuführen, wodurch die Ergebnisse repliziert würden, wurden durch das zufällige Teilen der Gesamtstichprobe zwei äquivalente Stichprobenhälften generiert (vgl. Fabrigar/Wegener/MacCallum/Strahan 1999, S. 277).

Ausgehend von den Befunden der explorativen und konfirmatorischen Analyse werden Korrelationen und Regressionsanalysen berechnet, um Zusammenhänge zwischen einzelnen Dimensionen zu prüfen.

Die statistischen Analysen wurden mit der Predictive Analytics Software (PASW) 18 sowie IBM SPSS Statistics 20<sup>35</sup> durchgeführt. Für die Berechnung der konfirmatorischen Faktorenanalyse wurde das Programm Mplus 6 (Muthén/Muthén 1998-2010) verwendet.

---

<sup>34</sup> Der Datensatz verfügt über eine ausreichend große Stichprobe, ein günstiges Verhältnis von Variablenzahl zur Stichprobengröße, substanzielle Korrelationen in der Korrelationsmatrix zwischen den Items (errechneter Kaiser-Meyer-Olkin (KMO)-Koeffizient von ,95; signifikanter Bartlett-Test  $p < ,001$ ) (vgl. Bühner 2011, S. 344-347; Rost 2007, S. 206; Backhaus u.a. 2011, S. 341-343).

<sup>35</sup> Die kurzzeitig im Jahr 2009 in PASW-Statistics 18 umbenannte Software lief vorher unter dem Namen Software Statistical Package for the Social Sciences (SPSS) und heißt seit 2010 IBM SPSS Statistics. Im Folgenden wird die Bezeichnung SPSS gewählt.

## 8.2. Ergebnisse

### 8.2.1. Ergebnisse der explorativen Faktorenanalyse (EFA)

Auf der Datengrundlage einer Stichprobenhälfte (N = 577) wird nach dem Eigenwertkriterium größer eins als Obergrenze der sinnvoll zu interpretierenden Komponenten eine sechsfaktorielle Lösung angezeigt, welche zusammen etwa 60 % der Gesamtvarianz der Variablen erklärt. Tabelle 12 ist das Ergebnis der explorativen Hauptkomponentenanalyse zu entnehmen.

**Tab. 12: Ergebnisse der explorativen Faktorenanalyse (PCA) der Evaluationsbögen**

		Rotierte Komponentenmatrix <sup>a</sup>					
		Komponente					
Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung		1	2	3	4	5	6
iv_11	Die Textteile des Berichts sind hilfreich für die weitere Arbeit unserer Schule.	<b>,76</b>	,12	,13	,13	,26	,14
v_02	Die Qualitätsanalyse hat unserer Schule wichtige Impulse für die weitere Entwicklung geliefert.	<b>,72</b>	,10	,17	,11	,02	,18
iv_12	Die Datenteile des Berichts sind hilfreich für die weitere Arbeit unserer Schule.	<b>,71</b>	,17	,20	,11	,10	,18
iv_04_1a	Der Bericht ist so detailliert, dass die Schule daraus gut Schwerpunkte und Maßnahmen für die Weiterentwicklung ableiten kann.	<b>,71</b>	,13	,15	,18	,26	,06
iii_11a	Die mündliche Erstrückmeldung hat der Schule geholfen, wichtige Impulse für Handlungsfelder der weiteren Schulentwicklung zu identifizieren.	<b>,60</b>	,40	,21	,14	,23	,04
iv_03_1a	Die Bewertungen im Bericht sind hinreichend begründet.	<b>,58</b>	,20	,22	,13	,16	,04
iii_12	Die Schule ist insgesamt zufrieden mit der Art und Weise, wie die Schulbesuchstage durchgeführt wurden.	,50	,47	,34	,25	,28	-,02
v_01	Die Schule stimmt insgesamt mit den Ergebnissen der Qualitätsanalyse überein.	,50	,44	,31	,11	,28	-,06
iii_02	Der Zeitrahmen für die Interviews war insgesamt angemessen.	,42	,19	,40	,20	,07	,02
iii_09	Der Schulbesuch hat die Abläufe in unserer Schule kaum beeinträchtigt.	,13	<b>,65</b>	,12	,20	,01	,31
iii_09_1a	Der Unterricht wurde durch die Anwesenheit der Qualitätsprüferinnen und -prüfer nicht beeinträchtigt.	,12	<b>,64</b>	,08	,10	,13	,20
iii_01	Die Stimmung im Kollegium bezüglich der Qualitätsanalyse war während der Schulbesuchstage insgesamt positiv.	,18	<b>,59</b>	,23	,00	,00	,41

<b>Rotierte Komponentenmatrix<sup>a</sup></b>							
Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung		Komponente					
		1	2	3	4	5	6
iii_06a	Das Qualitätsteam hat sich bei den Unterrichtsbesuchen und bei den Interviews freundlich und respektvoll verhalten.	,36	<b>,57</b>	,02	,17	,06	-,22
v_04	Die Stimmung im Kollegium bezüglich der Qualitätsanalyse ist nach den Schulbesuchstagen insgesamt positiv.	,49	,54	,21	,11	,20	,16
iii_10	Die Schule hat sich in der mündlichen Erstrückmeldung am Ende der Schulbesuchstage gut wiedererkannt.	,44	,53	,35	,14	,28	-,09
iii_08	Das Qualitätsteam hat während des Schulbesuchs den normalen Schul- und Unterrichtsalltag unserer Schule kennengelernt.	,08	,49	,44	,14	,16	,10
iii_09_2a	Die Besonderheiten unserer Schule wurden angemessen berücksichtigt.	,38	,48	,37	,16	,24	-,07
iii_03a	Die Anzahl der Unterrichtsbesuche war insgesamt angemessen.	,18	,15	<b>,77</b>	,10	,03	,17
iii_03_1a	Die Auswahl der Unterrichtsstunden bildet den Unterricht unserer Schule gut ab.	,24	,16	<b>,73</b>	,07	,08	,09
iii_01_1a	Der Zeitrahmen für den Schulbesuch war insgesamt angemessen.	,32	,17	<b>,71</b>	,14	,04	,11
ii_06	Die Abstimmung mit dem Qualitätsteam verlief reibungslos.	,08	,10	,19	<b>,73</b>	,21	,08
ii_03	Eventuelle Unsicherheiten und Bedenken konnten geklärt werden.	,24	,21	,02	<b>,69</b>	,00	,21
ii_01	Die Schule erhielt vor den Schulbesuchstagen alle Informationen, die notwendig waren.	,28	,13	-,04	<b>,63</b>	,05	-,03
ii_05	Die Planung des Gesamtablaufs des Schulbesuchs verlief reibungslos.	-,02	,11	,29	<b>,59</b>	,23	,21
iii_04	Der geplante Ablauf wurde eingehalten.	,07	,13	,33	,42	,23	-,19
iv_04	Die Bewertungen stimmen mit den Aussagen der mündlichen Bilanz nach dem Schulbesuch überein.	,15	,12	,09	,19	<b>,69</b>	-,01
iv_03	Der Bericht ist prägnant und verständlich.	,40	,00	,17	,09	<b>,67</b>	,16
iv_01a	Der Bericht wurde uns zeitnah (zum verabredeten Zeitpunkt) zugesandt.	,05	,11	-,07	,05	<b>,63</b>	,06
iii_05	Das Qualitätsteam hatte gute Kenntnisse über die eingereichten Unterlagen aus dem Schulportfolio.	,35	,14	,18	,20	,54	-,15
ii_04rec	Die Stimmung im Kollegium bezüglich der QA war vor dem Schulbesuch insgesamt positiv (rec).	-,06	-,24	-,16	-,17	-,05	<b>-,67</b>
ii_07rec	Der Aufwand für die Zusammenstellung der Daten und Dokumente für das Schulportfolio war vertretbar (rec).	-,33	-,10	-,04	,06	-,02	<b>-,65</b>
v_03rec	Aufwand und Ertrag der Qualitätsanalyse stehen in einem angemessenen Verhältnis (rec).	-,45	-,22	-,20	-,05	-,09	<b>-,60</b>

a. Die Rotation ist in 9 Iterationen konvergiert.

Nach der Rotation mit der Varimax-Methode erklärt der erste Faktor 17 %, der zweite Faktor 11 %, der dritte Faktor 10 %, der vierte Faktor 8 %, der fünfte Faktor 7 % und der sechste Faktor schließlich 6 % der gesamten Variablenvarianz. Die Werte in den Komponentenspalten beinhalten die Ladungen, d. h. für jede Variable die Korrelationen der Variable mit den Komponenten eins bis sechs. Die ersten neun Variablen laden am stärksten auf der ersten Komponente und weisen Faktorladungen zwischen ,76 bis ,42 auf. Allerdings erfüllen drei Variablen die geforderte Ladungshöhe von  $\geq ,50$  nicht (iii\_02) oder weisen hohe Nebenladungen auf den weiteren Komponenten auf (iii\_12, v\_01) und werden aus diesen Gründen eliminiert. Der Ausschluss vier weiterer Variablen, welche auf Komponente 2 hoch laden, ist durch hohe Nebenladungen auf der ersten und dritten Komponente begründet. Auch bei den Komponenten 3 bis 6 werden solche nichteindeutigen Variablen aus der Reliabilitätsanalyse ausgeschlossen. Der Ausschluss einer Variablen (iii\_05) erfolgt aus inhaltlichen Gründen, obwohl die Ladungshöhe auf der Komponente 5 mit ,54 akzeptabel ist: Sie berücksichtigt jedoch im Gegensatz zu den übrigen Variablen, die auf derselben Komponente hoch laden, das Schulportfolio und nicht den Qualitätsbericht. Insgesamt liegen durch die Eliminierung von neun der insgesamt 32 in die Faktorenanalyse eingegangenen Variablen den folgenden Berechnungen 23 Variablen zugrunde.

### **8.2.2. Ergebnisse der konfirmatorischen Faktorenanalyse (CFA)**

Mit der zweiten Stichprobenhälfte ( $N = 578$ ) wurde eine konfirmatorische Faktorenanalyse gerechnet. Die Zuordnung der Variablen zu den Faktoren orientiert sich an den Ergebnissen der explorativen Faktorenanalyse. Es zeigt sich, dass die explorativ ermittelte Faktorenstruktur über akzeptable Gütemaße verfügt ( $\chi^2_{df=215} = 999,038$ ; RMSEA = ,079 (90/ CI: ,074 - ,084; CFI = ,885; TLI = ,865; SRMR = ,052) und auf eine gute Übereinstimmung mit den empirischen Daten hindeutet. Das Modell ist in Abbildung 14 graphisch mitsamt den Faktorladungen und -korrelationen dargestellt. Auf den Pfeilen sind die Ladungen abgetragen.

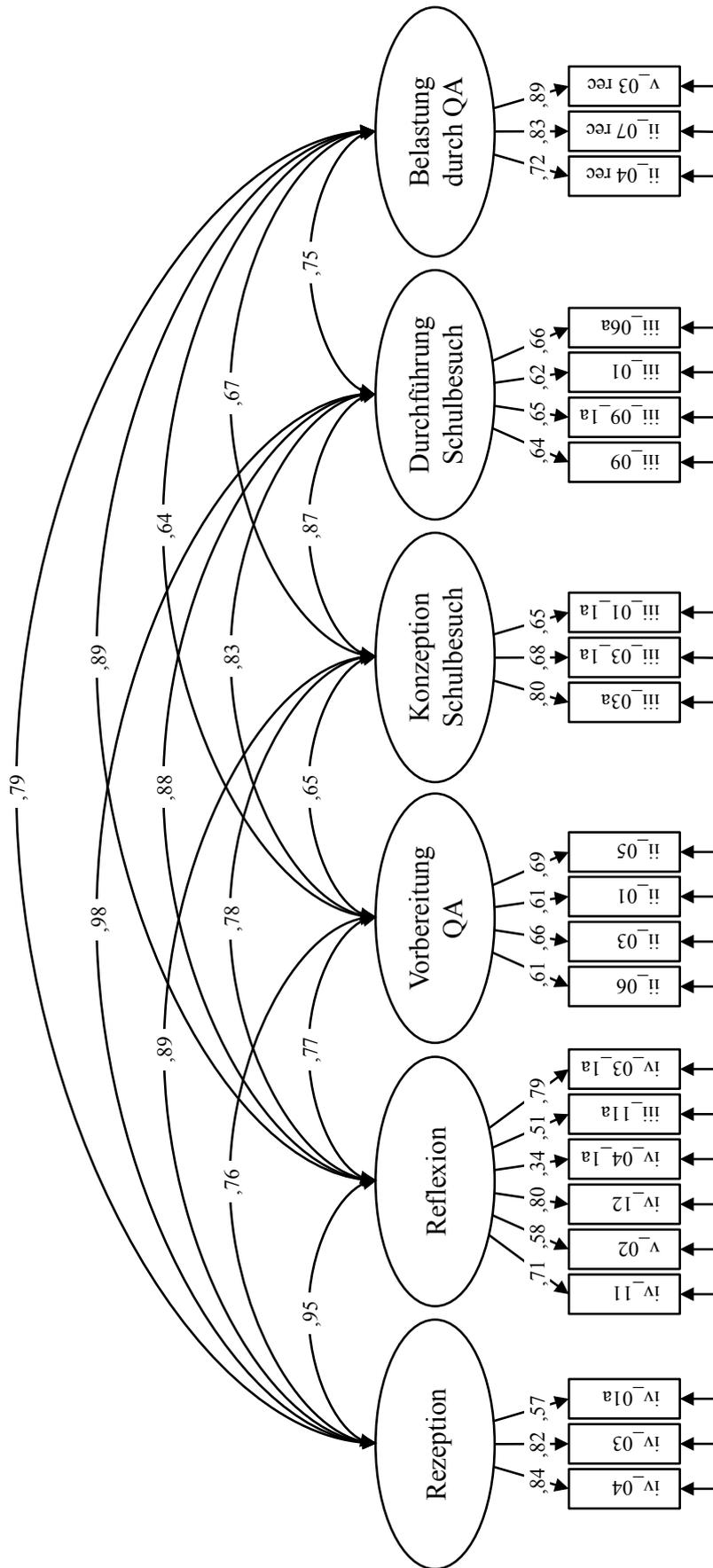


Abb. 14: Ergebnisse der konfirmatorischen Faktorenanalyse der Evaluationsbögen (N = 578)

Der signifikante Chi-Quadrat-Wert ( $p < ,05$ ) erweist sich zwar als unbefriedigend, war allerdings angesichts der umfangreichen Stichprobengröße von  $N = 578$  so zu erwarten. Auch wenn die übrigen, stichprobenunabhängigen Kriterien RMSEA, CFI und TLI eine adäquate bis gute Passung der Faktorenstruktur zu den Daten aufweisen, gilt es zusätzlich die Faktorladungen und -korrelationen zu untersuchen:

Den Parameterschätzungen der Faktorladungen kann entnommen werden, dass alle Werte mit ,34 bis ,89 im mittleren bis hohen Bereich liegen. Das Item iii\_04\_1a weist die geringste Ladungshöhe von ,34 auf, während es bei der explorativen Faktorenanalyse als Markieritem durch eine hohe Ladung von ,71 gekennzeichnet ist. Alle übrigen Faktorladungen liegen über der von Backhaus u.a. (2010, S. 362) geforderten Grenze von ,50, so dass hier von einem hohen Zusammenhang und einer starken Homogenität zwischen den Faktoren und den Variablen ausgegangen werden kann. Bezüglich der Faktorkorrelationen zeigt sich, dass alle sechs Faktoren sehr hoch miteinander korrelieren (,64 bis ,98).

Insgesamt ist mit den Ergebnissen der konfirmatorischen Faktorenanalyse wahrscheinlichkeits-theoretisch die Kompatibilität der empirischen Daten zum explorativ identifizierten Modell bestätigt worden. Allerdings impliziert der konvergierende Befund zwischen der explorativen und konfirmatorischen Faktorenanalyse nicht, dass es nicht noch ein anderes Modell gibt, welches die empirischen Daten noch besser erklären könnte.

### 8.2.3. Skalenbeschreibungen

Auf der Grundlage der Faktorenextraktion ist die Beantwortung der Frage gefordert, inwiefern die identifizierte Struktur der Daten inhaltlich dem zugrunde gelegten Modell zugeordnet werden kann. Es zeigt sich, dass die einzelnen Skalen den Phasen *Vor der Schulinspektion*, *Während der Schulinspektion* und *Nach der Schulinspektion* zugewiesen werden können. Damit stellen sie die geforderte Parallele zu dem postulierten Modell dar. Die sechs extrahierten Komponenten können inhaltlich sinnvoll und in Anlehnung an das postulierte Modell interpretiert werden und ergeben die folgenden Skalen: *Rezeption der Ergebnisse*, *Reflexion der Ergebnisse*, *Vorbereitung der Qualitätsanalyse*, *Konzeption des Schulbesuchs*, *Durchführung des Schulbesuchs* und *Empfundene Belastung durch die Qualitätsanalyse*. Dabei entsprechen die erstgenannten

beiden Skalen eindeutig dem ursprünglichen Rahmenmodell von Helmke (2004), während die übrigen vier Skalen neue Erkenntnisse zu weiteren Phasen und Rahmenbedingungen liefern. Die extrahierten Faktoren sind inhaltlich wie folgt zu charakterisieren:

*Vor der Schulinspektion:*

- *Vorbereitung der Qualitätsanalyse:* Diese Skala beinhaltet Fragen der Abstimmung, Planung und Kommunikation zwischen der Schule und dem Qualitätsteam im Vorfeld der Schulbesuchstage. Die Phase *Vorbereitung der Schulinspektion* wurde bislang im Kontext der Wirkungsdiskussion der Schulinspektion unberechtigterweise unterschätzt, da laut neusten Erkenntnissen von Gärtner (2011, S. 158; Gärtner/Pant 2011, S. 26) bereits die schulischen Vorbereitungsaktivitäten auf die Schulinspektion Wirkungen entfachen. Beispiel-Item: *Die Abstimmung mit dem Qualitätsteam verlief reibungslos.*
- *Empfundene Belastung durch die Qualitätsanalyse:* Da im Evaluationsbogen keine negativ formulierten Items eingesetzt werden, wurden entsprechende Items umgepolt und erfassen somit die gegenteilige Aussage. Diese Skala beleuchtet damit negativ wahrgenommene Aspekte der Qualitätsanalyse insbesondere im Vorfeld des Schulbesuchs. Unter anderem beinhaltet die Skala den Aufwand für die Zusammenstellung der Daten und Dokumente für das Schulportfolio und die Stimmung im Kollegium im Vorfeld des Schulbesuchs. Zudem bildet die Einschätzung des Verhältnisses von Aufwand und Ertrag der Qualitätsanalyse (Effizienz) als globale Wirkungsfrage einen Bestandteil der Skala. Beispiel-Item: *Die Stimmung im Kollegium bezüglich der QA war vor dem Schulbesuch insgesamt positiv (rec).*

*Während der Schulinspektion:*

- *Konzeption des Schulbesuchs:* Unter diese Skala wird aus schulischer Perspektive die Angemessenheit der Auswahl und Anzahl der Unterrichtsbesuche sowie der Zeitrahmen des Schulbesuchs gefasst. Anhand dieser Skala findet eine Einschätzung der logistischen und organisatorischen Aspekte des Schulbesuchs statt. Beispiel-Item: *Die Anzahl der Unterrichtsbesuche war insgesamt angemessen.*

- *Durchführung des Schulbesuchs*: Die Stimmung im Kollegium und das wahrgenommene Verhalten des Qualitätsteams bei den Unterrichtsbesuchen und den Interviews bilden neben einem störungsfreien Ablauf des Schulbesuchs die unterschiedlichen Bestandteile der Skala. Beispiel-Item: *Der Schulbesuch hat die Abläufe in unserer Schule kaum beeinträchtigt.*

*Nach der Schulinspektion:*

- *Rezeption der Ergebnisse*: Die Rezeption der Ergebnisse der Qualitätsanalysen erfolgt bereits mit der mündlichen Rückmeldung am letzten Besuchstag der Qualitätsanalyse und daran anschließend mit der Zusendung des schriftlichen Qualitätsberichts. Anhand dieser Skala werden diejenigen Merkmale der Qualitätsberichte erfasst, welche zu einer erfolgreichen Rezeption der Ergebnisse beitragen. Dazu gehören die zentralen Merkmale Verständlichkeit, Konsistenz und Aktualität (zeitliche Nähe). Diese Merkmale stimmen mit den in der Literatur berichteten zentralen Kennzeichen einer gelungenen Rezeption von Leistungsrückmeldungen überein (vgl. Helmke 2004; Helmke 2009, S. 309; Hosenfeld 2010, S. 20). Beispiel-Item: *Der Bericht ist prägnant und verständlich.*
- *Reflexion der Ergebnisse*: Die Reflexion der Schulinspektionsergebnisse umfasst kognitive und emotionale Prozesse, die im Anschluss an die Bereitstellung der Ergebnisse und ihre Rezeption entstehen. Günstigenfalls führen die Rückmeldungen zum Nachdenken über ihre Interpretation, zu potentiellen Erklärungsversuchen und zur Erwägung von Veränderungen. Im Idealfall lösen die Qualitätsanalyse und die unterschiedlichen Elemente der Ergebnisrückmeldung Impulse für die daran anschließende Schulentwicklungsarbeit aus. Da die Skala zum einen die Elemente von Rückmeldungen (mündliche Rückmeldung; Daten- und Textteile der schriftlichen Rückmeldung), zum anderen die Qualität der Rückmeldungen (Detaillierungsgrad und Begründung der Bewertungen) beinhaltet, erfolgt von schulischer Seite damit in erster Linie eine Einschätzung der Nützlichkeit der Ergebnisrückmeldungen und weniger der Intensität des tatsächlichen Reflexionsprozesses. Auch wenn im originären Rahmenmodell von Helmke (2004) die Phase *Reflexion* mit den beiden Merkma-

len *Suche nach Erklärungen* und *Erhebung zusätzlicher Informationen* operationalisiert ist, so wird im Rahmen dieser Arbeit die *Einschätzung der Nützlichkeit der Ergebnisse* auch als Reflexionsprozess begriffen, da sie die Auseinandersetzung mit den Ergebnissen voraussetzt und an die Phase *Rezeption* anknüpft. Beispiel-Item: *Der Bericht ist so detailliert, dass die Schule daraus gut Schwerpunkte und Maßnahmen für die Weiterentwicklung ableiten kann.*

### 8.2.4. Deskriptive Skalennennwerte

Im Anschluss an die explorative Identifizierung und konfirmatorische Prüfung der dimensional Struktur des Evaluationsbogens sowie der Skalenbeschreibung werden die deskriptiven Kennwerte der Skalen auf der Grundlage der Gesamtstichprobe (N = 1155) dokumentiert. Tabelle 13 enthält für die einzelnen Skalen die genauen Itemformulierungen und statistisch relevante Kennwerte der Items. Die einzelnen Items der Skalen haben eine akzeptable Trennschärfe. Die korrigierten Trennschärfen der Items liegen zwischen  $r_{it} = ,36$  und  $r_{it} = ,77$ . Nur sieben Trennschärfen der Items sind als mittelhoch einzustufen ( $r_{it} = ,36$  bis  $r_{it} = ,49$ ), alle übrigen Trennschärfekoeffizienten sind als hoch zu bewerten.

**Tab. 13: Itemformulierungen und statistisch relevante Kennwerte der Skalen der Sekundäranalyse (N = 1155)**

Antwortskala: 1 = stimme gar nicht zu, 2 = stimme eher nicht zu, 3 = stimme eher zu, 4 = stimme voll zu			
<b>Rezeption der Ergebnisse</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Bewertungen stimmen mit den Aussagen der mündlichen Bilanz nach dem Schulbesuch überein.	3,8	,51	,40
Der Bericht ist prägnant und verständlich.	3,6	,56	,49
Der Bericht wurde uns zeitnah (zum verabredeten Zeitpunkt) zugesandt.	3,6	,84	,36
<b>Reflexion der Ergebnisse</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Textteile des Berichts sind hilfreich für die weitere Arbeit unserer Schule.	3,4	,66	,77
Die Qualitätsanalyse hat unserer Schule wichtige Impulse für die weitere Entwicklung geliefert.	3,2	,72	,65
Die Datenteile des Berichts sind hilfreich für die weitere Arbeit unserer Schule.	3,3	,70	,67
Der Bericht ist so detailliert, dass die Schule daraus gut Schwerpunkte und Maßnahmen für die Weiterentwicklung ableiten kann.	3,5	,62	,72
Die mündliche Erstrückmeldung hat der Schule geholfen, wichtige Impulse für Handlungsfelder der weiteren Schulentwicklung zu identifizieren.	3,3	,72	,69
Die Bewertungen im Bericht sind hinreichend begründet.	2,8	,78	,48

<b>Vorbereitung der Qualitätsanalyse</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Abstimmung mit dem Qualitätsteam verlief reibungslos.	3,8	,49	,59
Eventuelle Unsicherheiten und Bedenken konnten geklärt werden.	3,6	,57	,53
Die Schule erhielt vor den Schulbesuchstagen alle Informationen, die notwendig waren.	3,8	,44	,45
Die Planung des Gesamtablaufs des Schulbesuchs verlief reibungslos.	3,7	,53	,51
<b>Konzeption des Schulbesuchs</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Anzahl der Unterrichtsbesuche war insgesamt angemessen.	3,1	,75	,68
Die Auswahl der Unterrichtsstunden bildet den Unterricht unserer Schule gut ab.	2,8	,78	,58
Der Zeitrahmen für den Schulbesuch war insgesamt angemessen.	3,3	,70	,62
<b>Durchführung des Schulbesuchs</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Der Schulbesuch hat die Abläufe in unserer Schule kaum beeinträchtigt.	3,1	,75	,59
Der Unterricht wurde durch die Anwesenheit der Qualitätsprüferinnen und -prüfer nicht beeinträchtigt.	3,3	,70	,57
Die Stimmung im Kollegium bezüglich der Qualitätsanalyse war während der Schulbesuchstage insgesamt positiv.	3,0	,67	,52
Das Qualitätsteam hat sich bei den Unterrichtsbesuchen und bei den Interviews freundlich und respektvoll verhalten.	3,8	,54	,39
<b>Empfundene Belastung durch die Qualitätsanalyse</b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Stimmung im Kollegium bezüglich der Qualitätsanalyse war vor dem Schulbesuch insgesamt positiv (rec).	2,4	,71	,37
Der Aufwand für die Zusammenstellung der Daten und Dokumente für das Schulportfolio war vertretbar (rec).	2,7	,79	,56
Aufwand und Ertrag der Qualitätsanalyse stehen in einem angemessenen Verhältnis (rec).	2,4	,83	,54

Im Rahmen der Reliabilitätsanalyse erreichen vier Skalen ein Cronbachs Alpha ( $\alpha$ ) über  $\alpha = ,70$  (s. Tab. 14). Die Klassifikation zur Interpretation dieser Werte besagt, dass ein Wert ab  $\alpha = ,70$  zugelassen ist (vgl. Brosius 2008, S. 808; Schnell u.a. 2008, S. 153). Die Reliabilitäten der Skalen *Rezeption der Ergebnisse* ( $\alpha = ,58$ ) und *Empfundene Belastung durch die QA* ( $\alpha = ,67$ ) entsprechen diesen Kriterien nicht und sind konventionell als inakzeptabel zu bewerten. Zu berücksichtigen ist dabei jedoch, dass  $\alpha$  in positiver Weise von der Anzahl der Indikatoren abhängt und die beiden genannten Skalen sich aus lediglich drei Variablen zusammensetzen. Zudem verweist Fromm (2010, S. 58) darauf, dass die geforderte Reliabilität in der Praxis häufig nicht erreicht wird. Da beide Skalen wichtig für die inhaltliche Interpretation sind, werden sie beibehalten.

**Tab. 14: Übersicht über die Skalen der Sekundäranalyse (N = 1155)**

Skala	Anzahl Items	Cronbachs Alpha ( $\alpha$ )	Min	Max	MW	SD
Rezeption der Ergebnisse	3	,58	1	4	3,6	,48
Reflexion der Ergebnisse	6	,86	1	4	3,3	,55
Vorbereitung der QA	4	,73	2	4	3,7	,38
Konzeption des Schulbesuchs	3	,79	1	4	3,1	,62
Durchführung des Schulbesuchs	4	,72	1	4	3,3	,49
Empfundene Belastung durch die QA	3	,67	1	4	2,5	,61

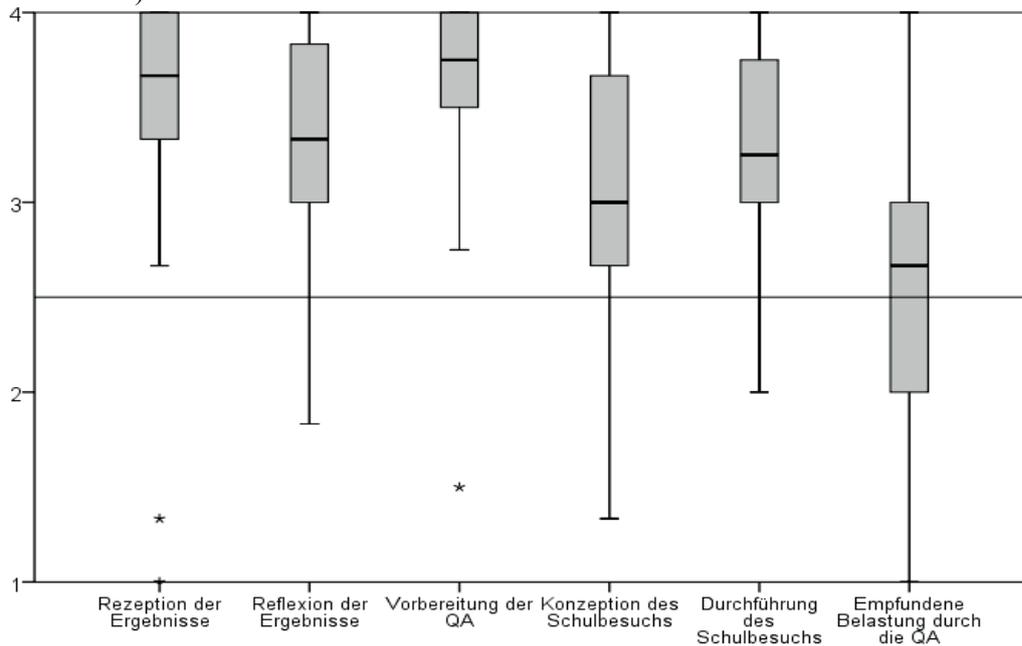
Antwortformat: 1 = stimme gar nicht zu, 2 = stimme eher nicht zu, 3 = stimme eher zu, 4 = stimme voll zu. Min = Minimum; Max = Maximum; MW = Mittelwert; SD = Standardabweichung; QA = Qualitätsanalyse NRW

Generell ist bei allen positiv konnotierten Skalen sowie den einzelnen Items zu beobachten, dass die Mittelwerte über 3,0 liegen und damit basierend auf dem vierstufigen Antwortformat sehr hoch ausfallen: Die Mittelwerte variieren bei den positiv konnotierten Skalen zwischen 3,1 und 3,7. Bei der negativ konnotierten Skala *Empfundene Belastung* wird ein durchschnittlicher Wert von MW = 2,5 erreicht und liegt damit genau auf der theoretischen Skalenmitte. Die Standardabweichungen (SD) der Skalen variieren zwischen SD = ,38 und SD = ,62. Bei der Skala *Vorbereitung der QA* lässt sich die geringste Streuung beobachten, während bei der Skala *Konzeption des Schulbesuchs* und *Empfundene Belastung durch die QA* die Angaben stärker streuen (s. Tab. 14). Insgesamt zeigen die Variablen und Skalen eine rechtssteile Verteilung mit Ausprägungen, welche an den oberen Rand der Antwortskala reichen. Die eingeschränkte Varianz und die homogen positiv bewerteten Variablen deuten damit auf einen Deckeneffekt hin. Die zusätzlich grafisch in Form von Boxplots<sup>36</sup> dargestellten Skalen liefern durch robuste Lage- und Streuungsmaße mehr Informationen als die bloße Betrachtung der Mittelwerte (s. Abb. 15).

<sup>36</sup> Lesehilfe für Boxplots: Ein Boxplot besteht aus zwei Geraden mit Endpunkten und aus einer Box (Quartilsabstand = mittlere 50 % der Werte). Der horizontale Strich innerhalb der Box gibt die Lage des Medians an, der die Verteilung halbiert. Die beiden Enden der Geraden außerhalb der Box zeigen die untersten und obersten 25 % der Werte an. Je länger die Box ist, desto stärker streuen die Beobachtungswerte im mittleren Bereich. Je länger die äußeren Geraden sind, desto stärker streuen die Ränder der Verteilungen. Die Extremwerte sind durch Sterne (\*) gekennzeichnet.

(1 = stimme gar nicht zu; 2 = stimme eher nicht zu; 3 = stimme eher zu; 4 = stimme voll zu;

N = 1155)



**Abb. 15: Boxplots der Skalen der Sekundäranalyse**

Insgesamt zeigt sich in Abb. 15, dass sich hinsichtlich der fünf positiv konnotierten Skalen die mittleren 50 % der Einschätzungen deutlich im Zustimmungsbereich oberhalb des theoretischen Mittelwerts von 2,5 bewegen. Vor allem die Skala *Vorbereitung der QA* wird von den Schulen in großer Übereinstimmung bzw. Dichte sehr positiv bewertet (Median = 3,8). Erwähnenswert ist zudem die differenzierte Betrachtung der Skalen *Rezeption* und *Reflexion der Ergebnisse*. Bei der Skala *Rezeption der Ergebnisse* liegen die mittleren 50 % der Einschätzungen im Vergleich zur Skala *Reflexion der Ergebnisse* im höheren Zustimmungsbereich und weisen zudem eine höhere Dichte auf. Der Median liegt bei der Skala *Reflexion der Ergebnisse* bei 3,3 und bei der Skala *Rezeption der Ergebnisse* bei 3,7. Im Hinblick auf die *Konzeption des Schulbesuchs* streuen die Einschätzungen der Schulen vergleichsweise breit und liegen mit einem Median von 3,0 zwar im positiven Bereich der Antwortskala, aber deutlich unter den Werten der anderen positiv konnotierten Skalen. Die größte Streuung weist die Skala *Belastung durch die QA* auf: Hier geht der Grad der empfundenen Belastung zwischen den Schulen sehr weit auseinander. Der Median liegt bei 2,7 und damit knapp oberhalb des theoretischen Mittelwerts.

### 8.2.5. Diskussion der Skalen

Sowohl die faktorenanalytische Identifikation als auch die inhaltliche Beschreibung der Skalen zeigt eine gute Passung zum zugrunde gelegten Modell auf (s. Kapitel 5). Während sich die Skalen *Rezeption* und *Reflexion der Ergebnisse* eindeutig im Modell verorten lassen, können die Skalen *Vorbereitung der Qualitätsanalyse*, *Konzeption und Durchführung des Schulbesuchs* sowie *Belastung durch die Qualitätsanalyse* den zusätzlichen Phasen des Modells *vor der Schulinspektion* und *während der Schulinspektion* zugeordnet werden. Insgesamt zeigt sich ein differenziertes Bild bzgl. des schulischen Umgangs mit der Qualitätsanalyse und ihren Ergebnissen: Während die Phasen vor und nach der Schulinspektion tendenziell positiv von den Schulen wahrgenommen werden, zeigt sich bei der Phase während der Schulinspektion eine kritischere Einschätzung.

Vor der Schulinspektion liegt auf der Datengrundlage der Evaluationsbögen die Belastung der Schulen durch die Qualitätsanalyse auf einem mittleren Niveau, während sich bei der Betrachtung der Einzelitems zeigt, dass insbesondere der Aufwand für die Zusammenstellung der Daten und Dokumente für das Schulportfolio vergleichsweise wenig Zustimmung findet (MW = 2,7). Die Skala *Vorbereitung der Qualitätsanalyse* wird von den Schulen hingegen eindeutig positiv hervorgehoben (MW = 3,7): Den Schulen werden von den Qualitätsteams alle Informationen zur Verfügung gestellt und eventuelle Bedenken werden geklärt. Sowohl die Abstimmung mit dem Qualitätsteam als auch die Planung des Schulbesuchs verlaufen aus der Perspektive der Schulen einwandfrei. Dieser Befund lässt sich vermutlich zu einem großen Anteil auf die gründliche Ausbildungsphase der Qualitätsprüfer und die guten organisatorischen Bedingungen der Qualitätsanalyse zurückführen.

Dieses positive Bild bezüglich der Phase vor dem Schulbesuch setzt sich während des Schulbesuchs nicht gleichermaßen fort: Hierbei wird insbesondere die Konzeption des Schulbesuchs kritischer bewertet: Die Einschätzung der Schulen, dass die Auswahl der Unterrichtsstunden dem tatsächlichen Unterrichtsbild entspricht, ist geteilt (MW = 2,8) und zeigt, dass einige Schulen hierbei eine Divergenz wahrnehmen. Hingegen würdigten die Schulen während des Schulbesuchs das freundliche und respektvolle Verhalten der Qualitätsteams (MW = 3,8), auch wenn ihr Stimmungsbild bezüglich der Qualitätsanalyse nur

moderat ausgeprägt war (MW = 3,0). Dies deutet auf eine positive Wahrnehmung der personellen Ressourcen und eine geringer ausgeprägte Akzeptanz der Konzeption und Operationalisierung der Qualitätsanalyse hin. Insgesamt könnte die vergleichsweise kritischere Einschätzung der Phase während der Schulinspektion mit Ängsten und Widerstand der schulischen Akteure zusammenhängen, da einzig in dieser Phase ein unmittelbarer Kontakt zu den Qualitätsprüfern besteht. Hingegen sind die Schulen in den übrigen zwei Phasen zwar mit der Ankündigung und den Ergebnissen der Qualitätsanalyse konfrontiert, stehen aber hierbei nicht in direktem Kontakt zum Verfahren und den Akteuren.

In der Phase nach der Qualitätsanalyse sind wieder positivere Einschätzungen aus der Perspektive der Schulen zu konstatieren. Im Mittel werden die Qualitätsberichte als prägnant, verständlich, detailliert und aktuell wahrgenommen. Zudem werden die Ergebnisse als übereinstimmend mit den mündlichen Aussagen der Qualitätsprüfer und als hilfreich für die weitere Arbeit der Schulen empfunden. Dabei wird der Rezeptionsprozess der Ergebnisse mit einem Mittelwert von 3,6 im Vergleich zur Reflexion der Ergebnisse (MW = 3,3) positiver eingeschätzt. Dieselbe Antworttendenz wurde auch im Rahmen der niedersächsischen Inspektionsbefragung beobachtet (vgl. Sommer 2011). Dies ist vermutlich damit zu erklären, dass die Rezeption auch über eine eher passive Auseinandersetzung mit den Ergebnissen vollzogen werden kann, während die Reflexion der Ergebnisse eine aktivere Beschäftigung mit den Rückmeldungen verlangt und somit vermutlich seltener erfolgt. Bei der Skala *Reflexion* hebt sich ein Item deutlich von den anderen ab: Die Begründungen der Bewertungen in den Qualitätsberichten erweisen sich aus Sicht der Schulen nicht immer vollständig nachvollziehbar (MW = 2,8).

Bei der Beschreibung und Charakterisierung der Skalen sind einige Einschränkungen zu berücksichtigen, die zum einen aus dem Verfahren der Qualitätsanalyse und der Evaluationsbefragung und zum anderen aus dem zugrunde gelegten Rohdatensatz resultieren. Da die Evaluationsbefragung zeitnah im Anschluss an die Zusendung der Qualitätsberichte erfolgt, können u. U. intensive Reflexionsprozesse in den Schulen noch nicht erfolgt sein. Zudem wurde der Evaluationsbogen mehrheitlich (in 87 % der Fälle) von der Schulleitung ausgefüllt, so dass zum Ausmaß der Auseinandersetzung einzelner Lehrkräfte mit den Ergebnissen der Qualitätsanalyse keine Angaben gemacht werden können.

Wie im Forschungsstand zur Schulinspektion deutlich geworden ist, schätzen Schulleitungen die Schulinspektion generell positiver ein als Lehrkräfte. Im Rahmen der niedersächsischen Inspektionsbefragung (N = 534 Schulen; vgl. Sommer 2011) konnte bei den Schulleitungen im Vergleich zur Gruppe der Lehrkräfte eine höhere Zustimmung bei der Beurteilung der Rezeption und Reflexion der Inspektionsergebnisse beobachtet werden. Eine weitere Einschränkung der Evaluationsbefragung kann darin gesehen werden, dass keine genauen Informationen darüber vorliegen, welcher Anteil des Kollegiums und welche Personengruppen (Steuergruppen etc.) den Qualitätsbericht gelesen und damit die Ergebnisse der Qualitätsanalyse rezipiert haben.

### **8.2.6. Zusammenhänge zwischen den Phasen und Rahmenbedingungen**

Ausgehend von den oben dargestellten Befunden, gemäß derer sich die Phasen des erweiterten Helmkeschen Modells identifizieren und ergänzen lassen und eine akzeptable Passung zu den empirischen Daten vorweisen, liegt die Vermutung nahe, dass die im Rahmenmodell postulierten Zusammenhänge zwischen den Phasen sich auch empirisch belegen lassen. Zur Überprüfung der Zusammenhänge werden in einem ersten Schritt Korrelationskoeffizienten (Produkt-Moment-Koeffizient nach Pearson) zwischen den einzelnen Skalen berechnet.

Tabelle 15 ist zu entnehmen, dass zwischen allen Skalen signifikante Zusammenhänge im mittleren bis hohen Bereich zu verzeichnen sind. Der Zusammenhang zwischen den Skalen *Rezeption* und *Reflexion* ist positiv und zeigt mit einem Korrelationskoeffizienten von  $r = ,56$  einen hohen Zusammenhang auf. Je eher sich also die Ergebnisrückmeldungen als konsistent und verständlich erweisen und durch eine zeitnahe Bereitstellung gekennzeichnet sind, desto positiver wird in den anschließenden Reflexionsprozessen die Nützlichkeit der Ergebnisrückmeldungen für die Schulentwicklungsarbeit eingeschätzt. Auffällig ist zudem der hohe Korrelationskoeffizient von  $r = ,60$  zwischen den Skalen *Durchführung des Schulbesuchs* und der *Reflexion der Ergebnisse* sowie von  $r = ,59$  zwischen den Skalen *Konzeption des Schulbesuchs* und der *Reflexion der Ergebnisse*: Je störungsfreier also der Schulbesuch verläuft und je respektvoller und freundlicher die Qualitätsprüfer auftreten, desto stärker findet eine Auseinandersetzung mit den Ergebnissen statt und desto nützlicher werden die

Ergebnisse der Qualitätsanalyse bewertet. Offenbar trägt hierbei der Eindruck, wie der Schulbesuch von der Qualitätsanalyse logistisch konzeptioniert und die Art und Weise, mit der der Schulbesuch tatsächlich vollzogen wird, dazu bei, als wie nützlich die Ergebnisse angesehen und reflektiert werden.

**Tab. 15: Skalen-Interkorrelationen der Sekundäranalyse**

Skalen	Reflexion der Ergebnisse	Vorbereitung der QA	Konzeption des Schulbesuchs	Durchführung des Schulbesuchs	Belastung durch die QA
Rezeption der Ergebnisse	,56**	,42**	,34**	,36**	-,30**
Reflexion der Ergebnisse		,53**	,59**	,60**	-,55**
Vorbereitung der Qualitätsanalyse (QA)			,45**	,52**	-,39**
Konzeption des Schulbesuchs				,53**	-,45**
Durchführung des Schulbesuchs					-,51**

\*\* Die Korrelation ist auf dem Niveau von ,001 (2-seitig) signifikant; N = 1155

Die negativ konnotierte Skala *Empfundene Belastung durch die QA* hingegen zeichnet sich durch negative Korrelationen zu allen anderen Skalen aus: D. h. je höher die empfundene Belastung durch die Qualitätsanalyse, desto problematischer und als unangemessener empfunden verläuft die Vorbereitung auf die Qualitätsanalyse ( $r = -,39$ ) und desto weniger werden Ergebnisse der Qualitätsanalyse rezipiert ( $r = -,30$ ), reflektiert und als nützlich angesehen ( $r = -,55$ ).

Als weitere statistische Methode, Zusammenhänge zwischen den einzelnen Skalen zu messen, wurde die Regressionsanalyse als strukturenprüfendes Verfahren eingesetzt, um die Zusammenhänge zwischen den Skalen zu beschreiben, erklären und prognostizieren (vgl. Backhaus u.a. 2010, S. 56f.). Als quantitative Größenordnung wird das globale Gütemaß zur Prüfung der Regressionsfunktion, das Bestimmtheitsmaß  $R^2$ , verwendet. Das Bestimmtheitsmaß misst die Güte der Anpassung der Regressionsfunktion an die empirischen Daten und ergibt sich aus dem Verhältnis von erklärter Streuung zur Gesamtstreuung (vgl. Backhaus u.a. 2010, S. 72-74). Für das Regressionsmodell mit der unabhängigen Variable (UV) *Rezeption* und der abhängigen Variable (AV) *Reflexion* liegt  $R^2$  bei ,32. Dieser Wert entspricht dem quadrierten, bereits o. g.

Korrelationskoeffizienten und besagt, dass 32 % der Varianz in der *Reflexion* durch den Regressor *Rezeption* erklärt wird. 68 % der Varianz werden durch andere, unbekannte Faktoren verursacht. Ein Varianzaufklärungspotenzial von 32 % gilt bei sozialwissenschaftlichen Fragestellungen als ein relativ guter Wert (vgl. Fromm 2010, S. 94). Die Varianzanalyse (ANOVA) weist den Einfluss der UV als höchst signifikant aus ( $p = ,000$ ).

Zur Identifizierung der Einflussgrößen auf die Ausprägung der wahrgenommenen Rezeption und Reflexion werden schrittweise Regressionsanalysen durchgeführt (vgl. Bühl 2008, S. 367ff.). Dabei werden Variablen schrittweise auf ihre Eignung zur Schätzung einer abhängigen Variable überprüft.<sup>37</sup> In einem ersten Schritt wird die Skala *Rezeption* als AV festgelegt und der durch die Einflussvariablen erklärbare Varianzanteil statistisch bestimmt (s. Tab. 16).

**Tab. 16: Modellzusammenfassungen der schrittweisen Regressionsanalyse für die Skala *Rezeption* als AV**

Modell	R	R <sup>2</sup>	Korrigiertes R <sup>2</sup>	Standardfehler des Schätzers	Durbin-Watson-Statistik
1	,42	,18	,18	,44	
2	,45	,21	,20	,43	
3	,47	,22	,22	,43	
4	,47	,22	,22	,42	1,70

1. Modell: Einflussvariablen: Vorbereitung der QA

2. Modell: Einflussvariablen: Vorbereitung der QA, Konzeption des Schulbesuchs

3. Modell: Einflussvariablen: Vorbereitung der QA, Konzeption des Schulbesuchs, Durchführung des Schulbesuchs

4. Modell: Einflussvariablen: Vorbereitung der QA, Konzeption des Schulbesuchs, Durchführung des Schulbesuchs, Empfundene Belastung durch die QA; N = 1155

Es zeigt sich in der vierstufigen Aufnahmereihenfolge der Variablen, dass die gewählten Einflussvariablen *Vorbereitung der QA*, *Konzeption des Schulbesuchs*, *Durchführung des Schulbesuchs* sowie *Empfundene Belastung durch die QA* insgesamt einen nur marginalen Einfluss auf die *Rezeption* der Ergebnisse

<sup>37</sup> Die Voraussetzungen für die Berechnung der Regressionsmodelle gelten als erfüllt: Die Kollinearitätsüberprüfung, d. h. die Unabhängigkeit der Einflussvariablen (UV) wurde anhand der Durbin-Watson-Statistik durchgeführt. Die erzielten Werte von unter 2 (s. Tab. 16 und 17) liegen in einem Bereich, der auf keinerlei Autokorrelation hindeutet. Im Kontext des Varianzhomogenität-Tests wurde die Nullhypothese, die besagt, dass kein systematischer Zusammenhang vorliegt, abgelehnt ( $p < ,001$ ). Zudem liegen die Verteilungen der Residuen sehr nah an der Normalverteilung.

ohne nennenswerten Zuwachs der Varianzaufklärung ( $,22 \geq R^2 \geq ,18$ ) aufweisen. Daraus kann geschlossen werden, dass die vorhandenen Variablen nicht wesentlich zur Erklärung der Gesamtvarianz der Skala *Rezeption* beitragen und dass insbesondere andere, bspw. schulspezifische, organisationale oder professionelle Merkmale einen größeren Einfluss auf die *Rezeption* haben könnten.

Hingegen erweisen sich diese Variablen für die Erklärung der Varianzen im Bereich der *Reflexion* durchaus als einflussreich (s. Tab. 17).

**Tab. 17: Modellzusammenfassungen der schrittweisen Regressionsanalyse für die Skala *Reflexion* als AV**

Modell	R	R <sup>2</sup>	Korrigiertes R <sup>2</sup>	Standardfehler des Schätzers	Durbin-Watson-Statistik
1	,60	,36	,36	,44	
2	,71	,50	,50	,39	
3	,75	,56	,56	,37	
4	,77	,59	,59	,35	
5	,78	,60	,60	,35	1,98

1. Modell: Einflussvariablen: Durchführung des Schulbesuchs  
 2. Modell: Einflussvariablen: Durchführung des Schulbesuchs, Rezeption der Ergebnisse  
 3. Modell: Einflussvariablen: Durchführung des Schulbesuchs, Rezeption der Ergebnisse, Konzeption des Schulbesuchs  
 4. Modell: Einflussvariablen: Durchführung des Schulbesuchs, Rezeption der Ergebnisse, Konzeption des Schulbesuchs, Empfundene Belastung durch die QA  
 5. Modell: Einflussvariablen: Durchführung des Schulbesuchs, Rezeption der Ergebnisse, Konzeption des Schulbesuchs, Empfundene Belastung durch die QA, Vorbereitung der QA; N = 1155

Über ein Drittel (36 %) der Varianz der Variable *Reflexion* kann durch den Prädiktor *Durchführung des Schulbesuchs* erklärt werden. Wird die Variable *Rezeption der Ergebnisse* in das Modell mitaufgenommen, so können damit 50 % der Gesamtvarianz der *Reflexion* erklärt werden, was einem sehr hohen Varianzaufklärungspotential entspricht. Weitere 10 % des Varianzaufklärungsanteils sind durch die Hinzunahme der übrigen drei Variablen *Konzeption des Schulbesuchs*, *Empfundene Belastung durch die QA* und *Vorbereitung der QA* zu erzielen. Allerdings erhöht sich mit Aufnahme der beiden letztgenannten Variablen das Varianzaufklärungspotential des Modells um 4 %, was als bescheiden zu interpretieren ist und deshalb zu überdenken wäre. Insbesondere die *Vorbereitung der QA* scheint so gut wie keine Erklärungskraft bezüglich der Varianz innerhalb der abhängigen Variable *Reflexion* zu haben.

### 8.3. Gewinn und Grenzen der Sekundäranalyse: eine kritische Reflexion

Die im Rahmen der Sekundäranalyse identifizierten sechs Skalen verdeutlichen den prozeduralen Umgang mit den Ergebnissen aus Schulinspektionen. Neben den zentralen Phasen im Anschluss an die Ergebnismeldungen stellen sowohl die Phase *vor* als auch *während* der Schulinspektion wichtige Bedingungen und Informationsquellen bezüglich des schulischen Umgangs mit den Ergebnissen dar. Die Phasen vor und nach der Schulinspektion erfahren im Vergleich zur Phase während der Schulinspektion deutlich positivere Einschätzungen. Dies könnte als Indiz bezüglich der Ängste und des Widerstands der schulischen Akteure interpretiert werden, welche beim Schulbesuch am unmittelbarsten wahrgenommen werden.

Im Folgenden wird eine kritische Einordnung der Datenqualität und Analyseverfahren der Studie I vorgenommen. Die von Flick (2009, S. 219) aufgezeigten Grenzen von Sekundäranalysen, insbesondere die geringe Kontrolle über die Datenqualität und die Datenfehler, gelten grundsätzlich auch für die vorliegende Auswertung. Dadurch, dass die Daten allerdings im Kontext einer Landesbehörde zu verorten sind, kann von einer hohen Datenqualität ausgegangen und von einer hohen Rücklaufquote profitiert werden. Eine wesentliche Kritik besteht in der möglichen sozial erwünschten Beantwortung der Evaluationsbögen. Dadurch, dass die Beantwortung der Evaluationsbögen für alle Schulen verpflichtend vom MSW NRW verordnet wurde, könnten sich die Schulen unbewusst oder bewusst gemäß einer Idealvorstellung der Qualitätsanalyse dargestellt haben. Die Gefahr der sozial erwünschten Antworten wäre sicherlich bei einer externen Befragung geringer gewesen. Eine weitere diskussionswürdige Fragestellung bezieht sich auf die fragebogenausfüllende Person bzw. Personengruppe. Bei den Ergebnissen der Sekundäranalyse wird von *schulischer* Rezeption und Reflexion gesprochen. Da jedoch nahezu 90 % der Evaluationsbögen von Schulleitungen ausgefüllt wurden, müsste bei der Auswertung der Evaluationsbögen stärker die Perspektive der Schulleitungen betont werden, die nicht zwingend die Wahrnehmung des Kollegiums widerspiegelt. Die Schulleitung ist die einzige schulische Instanz, die sich einer eingehenden Auseinandersetzung mit den Ergebnissen der Qualitätsanalyse nicht verwehren kann und

den Abschluss von Zielvereinbarungen als unmittelbare Reaktion auf die Inspektionsergebnisse mitverantwortet. Die starke Gewichtung der Schulleitungen bei der Evaluationsbefragung geht evtl. mit einer positiveren Wahrnehmung einher. Da der Forschungsstand (vgl. z. B. Gärtner/Wurster 2009) aufgezeigt hat, dass die Bewertung von externen Evaluationsformen oder auch speziell der Schulinspektion von Schulleitungen tendenziell positiver erfolgt als von Lehrkräften, stellt sich für die vorliegenden Befunde die Frage, wie sie sich darstellen würden, wenn vornehmlich Lehrkräfte befragt worden wären.

Zu berücksichtigen ist außerdem die Qualität des Evaluationsbogens als Messinstrument. Wie bereits im Vorfeld an einigen Stellen angemerkt wurde, sind Modifikationen und Optimierungen am Evaluationsbogen vorgenommen worden. Z. T. werden in diesem nicht alle Regeln zur Formulierung von Fragen und Antwortkategorien in sozialwissenschaftlichen Fragebögen erfüllt (vgl. Porst 2011). Wie die Itemkennwerte zeigen, differenzieren die Items zudem nicht genügend: Die Mittelwerte der einzelnen Items und Skalen liegen vornehmlich in einem überdurchschnittlichen Bereich der Antwortskalen, was auf einen Deckeneffekt hindeutet. Eine weitere Unsicherheit besteht in der Validität der identifizierten Phasen des Rezeptions-Nutzungs-Modells. Es gilt danach zu fragen, ob die durch die Faktorenanalyse erzeugten Komponenten wie beabsichtigt die einzelnen Phasen repräsentieren. Dieser Aspekt sollte bei der Zuordnung einiger Items zu übergeordneten Skalen mitberücksichtigt werden. Diskussionswürdig erscheinen zudem die starken Zusammenhänge zwischen den einzelnen Skalen, die auf den ersten Blick als Bestätigung der modellgeleiteten Annahmen zu interpretieren sind. Andererseits können die starken Zusammenhänge auch dahingehend interpretiert werden, dass der theoretische Prozessverlauf nicht differenziert abläuft, sondern einen einzelnen komplexen Vorgang darstellt. Somit wäre die Einteilung des Prozessverlaufs in die einzelnen Phasen nicht sinnvoll. Dagegen spricht, dass die Faktorenanalyse inhaltlich gut interpretierbare und insbesondere differenzierte Konstrukte liefert, die logisch aufeinander aufbauen und die Strukturierung des Modells bestätigen.

Zusammenfassend kann festgehalten werden, dass die zur Verfügung stehenden Daten und die modellgeleitete sekundärstatistische Auswertung einige Schwächen bergen. Dennoch stellen sie einen Gewinn und eine gute Zugangsmöglichkeit zu der Thematik *Umgang mit Schulinspektionsergebnissen* dar.

## 9. Studie II: Inhaltsanalyse der Zielvereinbarungen

Der Fokus in Studie II liegt auf der Untersuchung der Verarbeitung und Nutzung der Schulinspektionsergebnisse. Dazu stellen die zwischen Schulen und Schulaufsicht im Anschluss an die Qualitätsanalyse NRW geschlossenen Zielvereinbarungen eine relevante Informationsquelle dar. Im Rahmen der Zielvereinbarungsanalyse werden *formale* sowie *inhaltliche* Aspekte der Dokumente untersucht und die *Zielqualität* bewertet. Es erfolgt jedoch weder eine pädagogische Bewertung der Ziel- und Maßnahmeninhalte noch eine Evaluation der geplanten Aktivitäten. In einem ersten Schritt wird im Rahmen einer Dokumentenanalyse der Erkenntniswert der vorliegenden Zielvereinbarungen bestimmt. Daran schließt der Einsatz der Inhaltsanalyse als empirische Methode an, um die Zielvereinbarungen systematisch zu untersuchen.

An dieser Stelle sei ausdrücklich darauf hingewiesen, dass die Zielvereinbarungen lediglich die *geplante* Nutzung der Ergebnisse beinhalten und dass allein die Bestimmung von Zielen nicht automatisch zu einer Verbesserung der Schul- und Unterrichtsqualität führt. Allein aus dem Zielvereinbarungsdokument können keine Rückschlüsse auf die tatsächlich umgesetzten Maßnahmen, die geleistete schulische Arbeit und ihre Wirkungen gezogen werden. Die faktische Umsetzung der geplanten Maßnahmen wird in dieser Studie nicht untersucht. Wenn im Folgenden die Rede von der *Nutzung* der Ergebnisse ist, ist immer die *geplante Nutzung* bzw. die *Nutzungsabsicht* gemeint.

### 9.1. Dokumentenanalyse der Zielvereinbarungen

In einem ersten Schritt werden die Zielvereinbarungen einer Dokumentenanalyse unterzogen. Die Dokumentenanalyse ist nicht als Forschungsmethode zu verstehen, „sondern als Umschreibung einer spezifischen Zugangsweise zu schriftlichen Aufzeichnungen“ (Wolff 2009, S. 504; vgl. auch Glaser 2010, S. 366). Als Vorteile der Verwendung von Dokumenten in der empirischen Sozialforschung kann ihr nichtreaktiver Charakter, die direkte Verfügbarkeit der Daten und eine Vermeidung von Messfehlern durch das Entfallen der Datenerhebung gesehen werden (vgl. Flick 2007, S. 325-331; Wolff 2009, S. 502;

Mayring 2002, S. 49). Im nächsten Abschnitt werden die Zielvereinbarungen entlang folgender Kriterien analysiert: Art des Dokuments, äußere sowie innere Merkmale des Dokuments, Intendiertheit des Dokuments, Nähe des Dokuments zum Gegenstand und Herkunft des Dokuments (s. Mayring 2002, S. 48f.).

### 9.1.1. Beschreibung des Ausgangsmaterials

Für die Dokumentenanalyse konnten sämtliche bis Januar 2011 bereits existierenden *Zielvereinbarungen zu den Ergebnissen der Qualitätsanalyse NRW* einer<sup>38</sup> nordrhein-westfälischen Bezirksregierung herangezogen werden. Auch wenn die Zielvereinbarungen generell sowohl in den Schulen als auch in der Schulaufsicht vorliegen, wurde der Dokumentenzugang über die Schulaufsicht vorgezogen, da in dieser Institution alle verfügbaren Zielvereinbarungen zentral gebündelt vorlagen. Eine Sichtung der Original-Dokumente konnte nur in der Behörde erfolgen, während die kopierten Dokumente in anonymisierter Form auch extern ausgewertet werden durften. Für die Analyse wurden alle Namens- und Ortsangaben in den Zielvereinbarungen anonymisiert bzw. geschwärzt. Dieses Anonymisierungsvorgehen war als Bedingung der Dokumentenzustellung mit der Bezirksregierung vereinbart worden und unterstreicht die Vertraulichkeit und Bedeutsamkeit der Zielvereinbarungen.

Insgesamt waren bis zum Ende des ersten Schulhalbjahres 2010/11 114 Zielvereinbarungen verfügbar. Die Entstehungszeit der Dokumente liegt zwischen den Schuljahren 2007/08 und 2010/11. In Tabelle 18 wird ein Überblick über die Verteilung der analysierten Zielvereinbarungen auf die einzelnen Schulformen und die Schuljahre gegeben. Die meisten Zielvereinbarungen wurden in den Schulformen Hauptschule und Realschule getroffen (n = 25; 22 % und n = 24; 21 %). An dritter Stelle rangiert die Schulform Berufskolleg (n = 21; 18 %). Während aus den Schulformen Förderschule und Gymnasium jeweils ca. ein Dutzend Zielvereinbarungen zur Auswertung herangezogen werden, liegen in der Schulform Gesamtschule Zielvereinbarungen nur in geringer Anzahl vor (n = 2; 2 %).

---

<sup>38</sup> Aus Gründen der zugesicherten Anonymität wird der Bezirksregierungsstandort nicht genannt.

**Tab. 18: Anzahl der analysierten Zielvereinbarungen nach Schulform und -jahr**

Anzahl der Zielvereinbarungen nach Schulform (absolut und prozentual)			Anzahl der Zielvereinbarungen nach Schuljahr (absolut und prozentual)		
Grundschule	19	16,7 %	2006/07*	-	-
Förderschule	11	9,6 %	2007/08	21	18,4 %
Hauptschule	25	21,9 %	2008/09	15	13,2 %
Gesamtschule	2	1,8 %	2009/10	57	50,0 %
Realschule	24	21,1 %	2010/11**	21	18,4 %
Gymnasium	12	10,5 %			
Berufskolleg	21	18,4 %			
<b>Gesamt</b>	<b>114</b>	<b>100,0 %</b>	<b>Gesamt</b>	<b>114</b>	<b>100,0 %</b>

\* Im Schuljahr 2006/07 wurde die Qualitätsanalyse NRW implementiert. Daher erfolgte kein Abschluss von Zielvereinbarungen.

\*\* Es wurden Zielvereinbarungen aus dem ersten Schulhalbjahr 2010/11 berücksichtigt.

Die Hälfte der vorliegenden Zielvereinbarungen wurde im Schuljahr 2009/10 (n = 57; 50 %) und die geringste Anzahl an Zielvereinbarungen wurde im Schuljahr 2008/09 (n = 15; 13 %) getroffen. Die übrigen Zielvereinbarungen verteilen sich mit jeweils 18 % (n = 21) auf die Schuljahre 2007/08 sowie 2010/11. Aus dem Schuljahr 2006/07 liegen keine Zielvereinbarungen vor, da in diesem Jahr die Qualitätsanalyse NRW implementiert und mit den ersten Schulbesuchen begonnen wurde.

### 9.1.2. Quellenkritik

Im Anschluss an die Qualitätsanalyse NRW sollen auf der Grundlage der Qualitätsberichte Maßnahmen der Qualitätssicherung und -entwicklung geplant und in Form von Zielvereinbarungen schriftlich dokumentiert werden. Die Zielvereinbarungen sind von der obersten Schulaufsichtsbehörde als Element zur Steuerung der Qualitätsentwicklung vorgesehen und sollen nach einem landeseinheitlichen Verfahren erstellt werden. Die inhaltlich strukturierte Vorlage für die Zielvereinbarungen enthält Eingabefelder zur Ziel- und ggf. Teilzielbeschreibung, zu den Qualitätsbereichen, den Indikatoren und beabsichtigten Maßnahmen der Zielerreichung, zur geplanten und ggf. vereinbarten, externen Unterstützung sowie zur Terminierung der Bilanzgespräche (s. Zielvereinbarungsformular in der Handreichung zum institutionellen Zielvereinbarungsprozess zwischen Schulen und Schulaufsicht in NRW in: MSW NRW 2015b).

Die 114 analysierten Zielvereinbarungsdokumente sind durch diesen formalen Rahmen gekennzeichnet und liegen überwiegend in gedruckter, und in einer geringen Anzahl in elektronischer Form vor. Sie stellen repräsentative Primärdokumente dar, deren Zugang begrenzt ist, d. h. nur bestimmten Berufsgruppen innerhalb des MSW NRW, den Bezirksregierungen und Schulen gewährt wird. Da sie amtliche Aufzeichnungen darstellen, ist eine Reflexion hinsichtlich ihrer institutionellen Intention gefordert (vgl. Wolff 2009, S. 502-507). Die Implementierung der Zielvereinbarungen erfolgte mit der Intention, zusätzliche Impulse und ein standardisiertes Verfahren im Anschluss an die Qualitätsanalyse NRW zu initiieren, um den Umgang mit den Ergebnissen und die Nutzung dieser zu begünstigen. Damit sind die vorliegenden Zielvereinbarungen nicht primär aus einer inneren Bedürfnislage der Schulen entstanden, sondern als Ergebnis eines von der obersten Schulaufsichtsbehörde verordneten Steuerungsimpulses zur Qualitätssicherung und -entwicklung von Schule und Unterricht zu interpretieren. Darüber hinaus stellen Zielvereinbarungen strategische Dokumente dar, indem aus den dargelegten Zielen unter Berücksichtigung der Stärken-Schwächen-Analysen und möglicher Unterstützungsangebote operative Handlungen umgesetzt werden sollen. Die Angabe von konkreten Maßnahmen und den Indikatoren der Zielerreichung sollen gewährleisten, dass die Planungen möglichst konkret und überprüfbar erarbeitet werden und nicht nur vage Zielsetzungen bleiben. Allerdings führt die Verschriftlichung der geplanten Maßnahmen nicht zwangsläufig auch zur Umsetzung dieser.

Da es für den Abschluss von Zielvereinbarungen einer Einigung beider Parteien bedarf, zeichnet sich die Zielsetzung durch diskursive Prozesse aus. Hindernisse in diesem kooperativen Prozess könnten sich ergeben, wenn Schulleitung und Lehrkräfte ihre eigenen und von der Schulaufsicht losgelösten Zielpräferenzen einnehmen. In den Dokumenten fehlen jedoch Hinweise zu Entscheidungsprozessen bei der Vereinbarung von Zielen und Maßnahmen und den Interessen der Beteiligten: Erfolgt die Zielsetzung tatsächlich in einem diskursiven Prozess oder werden auf oberflächlicher Ebene mustergültige und konventionelle Ziele und Maßnahmen vereinbart, um die Forderung der oberen Verwaltungsebenen nach solch einem Dokument zu erfüllen? Diese potentiellen Szenarien gehen allein aus den Zielvereinbarungsdokumenten nicht hervor.

### 9.1.3. Fragestellungen

Folgende Fragen werden an die zu analysierenden Zielvereinbarungen gestellt:

- Durch welche formalen Aspekte (Umfang) sind die Zielvereinbarungen gekennzeichnet?
- Welche Inhalte (Qualitätsbereiche und -aspekte des Qualitätstableaus NRW) sind Gegenstand der Zielvereinbarungen?
- Gibt es einen Zusammenhang zwischen den Inhalten der Zielvereinbarungen und den von der Qualitätsanalyse diagnostizierten Defiziten der Schulen?
- Welche Qualität weisen die vereinbarten Ziele auf?

## 9.2. Methodisches Vorgehen: Inhaltsanalyse

Zur Beantwortung der o. g. Fragestellungen wird methodisch auf das inhaltsanalytische Verfahren zurückgegriffen, welches eine systematische Untersuchung der Zielvereinbarungen ermöglicht. Die Inhaltsanalyse zählt aufgrund ihres qualitativ-quantitativen Charakters zu den Standardverfahren sozialwissenschaftlicher Textanalyse (vgl. Mayring 2010, S. 602; Früh 2007, S. 42; Flick 2009, S. 144; Diekmann 2009, S. 586). Früh (2007, S. 27) definiert sie als

empirische Methode zur systematischen, intersubjektiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mitteilungen, meist mit dem Ziel einer darauf gestützten interpretativen Inferenz auf mitteilungsexterne Sachverhalte.

Abhängig von der Fragestellung der Untersuchung hat der Einsatz entweder der quantitativen oder der qualitativen Inhaltsanalyse seine Berechtigung. Während mittels quantitativer Inhaltsanalyse das Textmaterial hinsichtlich vorab festgelegter Merkmale quantifiziert werden kann, steht bei der qualitativen Inhaltsanalyse die Interpretation und die Aufdeckung latenter Strukturen im Vordergrund des Forschungsinteresses (vgl. Böhm-Kasper/Weishaupt 2008, S. 108; Bortz/Döring 2006, S. 151). Die Dichotomie beider Ansätze kritisierend kennzeichnet Früh die Inhaltsanalyse als quantitativ-qualitative Methode (vgl. Früh 2007, S. 68-74; s. auch Huber 1989). Seiner Ansicht nach sollten beide Elemente durch die Verwendung sowohl theoriegeleiteter (deduktiver) als auch

empiriegeleiteter (induktiver) Kategorienbildung zusammengeführt werden. Auch Mayring spricht sich für die Überwindung des Gegensatzes qualitativ vs. quantitativ aus, indem er argumentiert, dass der inhaltsanalytische Forschungsprozess von der Qualität zur Quantität und wieder zur Qualität verläuft (vgl. Mayring 2008, S. 16-20). Diesen Positionen der Verknüpfung von qualitativen und quantitativen Prozessen folgend, wird in dieser Studie sowohl ein qualitativer Zugang bei der Bildung eines Kategoriensystems als auch ein quantitativer Ansatz gewählt, bei dem ausgewählte Merkmale der Zielvereinbarungsdokumente in ihrer Häufigkeit bestimmt werden.

Da sich die Inhaltsanalyse durch ein regelgeleitetes Vorgehen auszeichnet, folgt die Auswertung der Zielvereinbarungen den Ablaufmodellen von Frequenz- sowie Valenz- und Intensitätsanalysen nach Mayring (vgl. Mayring 2008, S. 14f.). Während im Rahmen der Häufigkeits-/Frequenzanalysen bestimmte Textteile formaler oder inhaltlicher Art ausgezählt werden, findet bei Valenz- und Intensitätsanalysen eine Analyse des Textmaterials mithilfe einer zweistufigen (kategorialen) oder mehrstufigen (ordinalen) Einschätzskala statt (vgl. Mayring 2010, S. 601). Im Zentrum des inhaltsanalytischen Ablaufs steht das *Kategoriensystem*, dessen Bestandteile – die Kategorien – in einem Wechselverhältnis zwischen Theorie und Material entwickelt werden und mittels welcher sich das Analysematerial reduzieren bzw. verdichten lässt. Kategorien gelten als untereinander ähnlich wahrgenommene und einer bestimmten Merkmalsgruppe zugeordnete Kriterien (vgl. Früh 2007, S. 42). Es folgen die Bestimmung der Textbestandteile, die unter eine Kategorie fallen, die Nennung von Ankerbeispielen und Codierregeln, die durch eindeutige Zuordnungen Abgrenzungsprobleme zwischen den Kategorien klären. Die Codierung erfolgt durch möglichst mehrere Codierer. Schließlich wird das Textmaterial ausgewertet und interpretiert. Damit stellt der Gesamtprozess der Entwicklung, Durchführung und Interpretation einer Inhaltsanalyse eine einzige, lückenlose Argumentationskette dar (vgl. Früh 2007, S. 24). In den überwiegenden Analyseschritten kongruieren die beiden inhaltsanalytischen Techniken Frequenz- sowie Valenz- und Intensitätsanalysen; lediglich bei der Kategorien- bzw. Variablendefinition ergibt sich bei der Valenz- und Intensitätsanalyse als zusätzlicher Analyseschritt die Festlegung und Definition der Skalenpunkte.

### 9.2.1. Entwicklung des Kategoriensystems

Bei der Entwicklung eines Kategoriensystems lassen sich verschiedene Vorgehensweisen unterscheiden. Eine Variante bildet das deduktive Vorgehen: Dabei werden vorab festgelegte, aus anderen empirischen Untersuchungen oder aus der Theorie stammende Kategorien an das Material herangetragen. Umgekehrt werden beim induktiven, empiriegeleiteten Vorgehen die Kategorien aus dem Analysematerial gewonnen (vgl. Früh 2007, S. 153ff.). Diese Verfahrensweisen stehen nicht im gegensätzlichen Verhältnis zueinander, sondern existieren in der Forschungspraxis häufig als sich ergänzende Mischtypen (vgl. Bortz/Döring 2006, S. 151; Früh 2007, S. 73; Mayring 2008, S. 53).

Im Rahmen der *induktiven Kategorienbildung* wurden die wesentlichen Kriterien zur Beurteilung der formalen Struktur und des Inhalts der Zielvereinbarungen aus der Strukturierungsvorlage für Zielvereinbarungen abgeleitet (s. MSW NRW 2015b). Der formale Rahmen der Zielvereinbarungsdokumente erfordert Angaben zu den Zielen, den entsprechenden Qualitätsbereichen und den Maßnahmen der Zielerreichung. Als Grundsatz für die Formulierung der Ziele gibt das MSW NRW das SMART-Prinzip vor. Demnach sollten Ziele spezifisch, messbar, ausführbar, realistisch und terminierbar sein (s. Tab. 8). Zur Bestimmung der Zielqualität und für die Aufnahme in das Kategoriensystem erscheint lediglich die erste Bedingung des SMART-Prinzips – Spezifität – geeignet. Für die übrigen vier Postulate wäre eine fachliche Bewertung der Ziele im Kontext der jeweils individuellen schulischen Rahmenbedingungen und des Unterstützungsangebots gefordert.

Bei der *deduktiven Kategorienbildung* wurde zunächst auf den empirischen Forschungsstand zu Zielvereinbarungen im Kontext der Schulinspektion zurückgegriffen, um Anregungen und Anknüpfungspunkte für die eigene Analyse zu erhalten. Eine relevante Untersuchung zur Planung und Einleitung von Maßnahmen als Reaktion auf die Berliner Schulinspektion liefern Gärtner und Wurster (2009, S. 20-29). Allerdings basieren ihre Ergebnisse auf Befragungen von Schulleitungen und Lehrkräften, was eine Vergleichbarkeit und Übertragung des Analyserasters erschwert. Während bei Gärtner und Wurster die formalen Merkmale von Zielvereinbarungen nicht thematisiert werden, bildet die inhaltliche Kategorisierung der Ziele in Anlehnung an den Handlungsrahmen

Schulqualität, welcher das Berliner Pendant zum Qualitätstableau NRW darstellt, die Zugangsweise bei der inhaltlichen Analyse der Ziele. Weitergehende inhaltliche Anregungen für die Kategorienbildung sind zudem in einem vom Hessischen Kultusministerium herausgegebenen Leitfaden für die Entwicklung von Zielvereinbarungen zwischen Schule und Schulaufsicht nach der Schulinspektion enthalten (vgl. Hessisches Kultusministerium 2007). Darin werden beispielhaft vier Zielvereinbarungen dargestellt und kritisch kommentiert. Sie ermöglichen eine erste Orientierung bzgl. der Merkmale, Inhalte und Bewertungsmöglichkeiten für die Vereinbarung von Zielen.

Da keine weiteren Untersuchungen zu Zielvereinbarungen im Kontext der Schulinspektion bekannt sind, wurden für die deduktive Kategorienbildung außerdem Erkenntnisse aus der Auswertung von *Schulprogrammen* herangezogen und bestehende Kategoriensysteme an die vorliegenden Zielvereinbarungen herangetragen. Das Schulprogramm gilt analog zur Schulinspektion als ein Instrument zur Schulentwicklung und soll den Schulen dazu dienen, Schwerpunkte in der strategischen Schularbeit zu setzen. Insbesondere in den schriftlich fixierten Planungs- und Handlungskonzepten für Schulentwicklungsmaßnahmen wird eine Parallele der Schulprogramme zu Zielvereinbarungen im Kontext der Schulinspektion gesehen (vgl. Holtappels 2004, S. 15f.). Während das Schulprogramm als öffentliche Bekanntgabe mit dem Ziel der schulischen Außendarstellung verfasst wird, stellen die Zielvereinbarungen interne und vertrauliche Dokumente der Schulen und der Schulaufsicht dar. In Tabelle 19 sind die notwendigen Bestandteile der nordrhein-westfälischen Zielvereinbarungen und der grundsätzliche Aufbau von Schulprogrammen, der in dieser Form für Berlin, Nordrhein-Westfalen und Hamburg gilt, dar- und gegenübergestellt (vgl. Bestandteile der Schulprogramme in Köller 2009, S. 148-150). Die einzelnen Bestandteile sind in beiden Dokumentenarten weitestgehend identisch (s. Tab. 19). Insgesamt ist die Einbeziehung von Auswertungsstrategien der Schulprogramme zur Analyse von Zielvereinbarungen im Anschluss an die Schulinspektion grundsätzlich geeignet und legitimiert die deduktive Kategorienbildung. Durch die im Vergleich zu anderen Steuerungsinstrumenten frühe Verordnung zur Schulprogrammarbeit etwa um die Jahrtausendwende wurden bereits zahlreiche empirische Untersuchungen durchgeführt, welche einen Überblick über die Inhalte, Qualität und Wirkung der Schulprogramme geben.

**Tab. 19: Vorgaben zu den Bestandteilen von Schulprogrammen und Zielvereinbarungen im Rahmen der Qualitätsanalyse NRW**

<b>Bestandteile Schulprogramm</b> Geforderte Angaben zu folgenden Aspekten:	<b>Bestandteile Zielvereinbarung (QA NRW)</b> Geforderte Angaben zu folgenden Aspekten:
<b>Bestandsanalyse</b>	-
<b>Ziele der Entwicklungsvorhaben/ Schwerpunktsetzung</b> <ul style="list-style-type: none"> <li>• Unterrichtsentwicklung</li> <li>• Organisationsentwicklung</li> <li>• Personalentwicklung</li> <li>• Erziehung und Schulleben</li> </ul> <p>Formulierung der Ziele nach den SMART-Kriterien</p>	<b>Zielbeschreibung</b> <b>Qualitätsbereiche (QB) der Ziele</b> <ul style="list-style-type: none"> <li>• QB 1: Ergebnisse der Schule</li> <li>• QB 2: Lernen und Lehren – Unterricht</li> <li>• QB 3: Schulkultur</li> <li>• QB 4: Führung und Management</li> <li>• QB 5: Professionalität der Lehrkräfte</li> <li>• QB 6: Ziele und Strategien der Qualitätsentwicklung</li> </ul>
<b>Maßnahmenkatalog/Zeit- und Maßnahmenplanung für die Realisierung der Entwicklungsvorhaben</b> <ul style="list-style-type: none"> <li>• Vereinbarung von Maßnahmen zur Zielerreichung</li> <li>• Erstellen einer Zeitplanung</li> <li>• Festlegen von Verantwortlichkeiten</li> <li>• Formulierung von Indikatoren und Standards für die Zielerreichung</li> </ul>	<b>Maßnahmen in der Schule</b>
	<b>Indikatoren der Zielerreichung</b>
	<b>Datum der Zielerreichung</b>
	<b>Externe Unterstützung</b>
	<b>Terminierung der Bilanzgespräche</b>
<b>Evaluation</b>	<b>Bilanzgespräche</b>

Relevante Befunde zur Schulprogrammarbeit liegen in Form einer inhaltsanalytischen Auswertung Hamburger Schulprogramme (vgl. Holtappels/Müller 2002; 2004)<sup>39</sup>, einer inhaltlichen Analyse von bundesweiten Schulprogrammen<sup>40</sup> (vgl. Mohr 2006), einer quantitativen Inhaltsanalyse Berliner Schulprogramme<sup>41</sup> (vgl. Köller 2009, S. 143-194), sowie auf der Grundlage von Fallrekonstruktionen mittels Objektiver Hermeneutik<sup>42</sup> (vgl. Gruschka u.a. 2003) vor.

39 Holtappels und Müller untersuchten insgesamt 423 Hamburger Schulprogramme; diese entsprechen nahezu einer Vollerhebung (vgl. Holtappels/Müller 2004, S. 81).

40 Einbezogen wurden insgesamt 58 Schulprogramme im Kontext der IGLU-Studie aus den Bundesländern Brandenburg, Baden-Württemberg, Bremen, Hessen, Hamburg, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Schleswig-Holstein, Sachsen und Thüringen (vgl. Mohr 2006, S. 76; 94).

41 In die quantitative Inhaltsanalyse von Köller gingen insgesamt 26 Schulprogramme Berliner Schulen ein (vgl. Köller 2009, S. 113f.).

42 In dem Forschungsprojekt zu den Wirkungen der Schulprogrammarbeit wurden mittels Objektiver Hermeneutik 12 Einzelfallrekonstruktionen auf der Basis von Dilemma-interviews mit Schulleitungen, Lehrkräften und Schulaufsichtsbeamten sowie die Analyse von 70 Schulprogrammen in den Bundesländern Hamburg, Rheinland-Pfalz und Hessen durchgeführt (vgl. Gruschka/Heinrich/Köck/Martin/Pollmanns/Tiedtke 2003).

Zusammengefasst zeigen die Ergebnisse der analysierten Schulprogramme, dass sie z. T. durch deutliche Differenzen bezüglich ihrer Qualität, Struktur, Inhalte und ihres Umfangs gekennzeichnet sind. Inhaltlich enthalten die Schulprogramme den Ist-Zustand der Schulen in meist akzeptablen Ausführungen, welcher jedoch nur selten als Ausgangspunkt für eine kritische Auseinandersetzung genutzt wird und aus dem nur in unzureichendem Maße Kriterien zur Umsetzung von Maßnahmen und ihrer Evaluation abgeleitet werden.

Auf der Grundlage der o. g. Studien folgt die Darstellung ausgewählter Untersuchungsschritte und Elemente der Kategoriensysteme, die sich als nützlich und geeignet für die Kategorienbestimmung zur Analyse von Zielvereinbarungen erweisen. In der Hamburger Schulprogrammanalyse wurden die Entwicklungsschwerpunkte ausgewertet und die konzeptionelle Qualität der Schulprogramme bestimmt (vgl. Holtappels/Müller 2004, S. 99-101; ebd. 2002, S. 222-226): Zur Messung der Schulprogrammqualität wurden als Indikatoren die *konzeptionelle Fundierung*, der *Konkretisierungs- und Differenzierungsgrad* sowie die *innere Konsistenz* der Programmelemente festgelegt und auf einer vierstufigen Skala (++ = sehr gut; + = eher gut; - = Mängel, weniger gut; -- = insgesamt schlecht) bewertet. Das von Mohr (2006, S. 81-83) entwickelte Analyseraster von Schulprogrammen beinhaltet u.a. die formale Gestaltung der Schulprogramme, die Anzahl und die kategorielle Erfassung der Entwicklungsschwerpunkte sowie den Konkretisierungsgrad der geplanten Maßnahmen. Des Weiteren verfolgte Köller (2009) in ihrer Studie eine frequenz- und valenzanalytische Auswertungsstrategie, indem sie die formalen Merkmale der Dokumente erfasste, den Konkretisierungsgrad der Stärken und Schwächen auf einer dreistufigen Skala (sehr konkret, konkret und global/unspezifisch) sowie die inhaltlichen Bereiche der Entwicklungsvorhaben bestimmte. Eine weitere Vorgehensweise bei der Auswertung von Schulprogrammen enthält ebenfalls Anregungen für die eigene Studie: Zu schulaufsichtlichen Zwecken wurde von Herrgesell (2004) zur Sichtung und Auswertung von Schulprogrammen, in Rheinland-Pfalz als *Qualitätsprogramme* bezeichnet, ein Kategoriensystem entwickelt. Die darin enthaltenen Kategorien wurden zur Verdichtung der Daten gewichtet, so dass daraus der erforderliche Handlungsbedarf zur Beratung der Schulen für die Schulaufsicht erschlossen werden sollte (vgl. Herrgesell 2004, S. 150).

Die o. g. Auswertungskonzepte von Schulprogrammen dienten allesamt als Anregungen und Anknüpfungspunkte für die eigene Analyse von Zielvereinbarungen. Darüber hinaus konnte eine weitere Arbeit zur deduktiven Kategorienentwicklung herangezogen werden: Jäger (2004) gelang es in einer Untersuchung, ein (Wellen-)Modell für den Transfer von Erkenntnissen aus schulischen Entwicklungsprogrammen in Schulentwicklung zu konstruieren und dies anhand von empirischen Daten eines Modellversuchsprogramms<sup>43</sup> zu überprüfen. Dazu wurden Bewertungsmethoden und -kriterien zur Bestimmung der Zielqualität entwickelt. Im Vorfeld seiner Untersuchung hat Jäger angenommen, dass die Qualität der Ziele den Transfer beeinflusst. Zur Überprüfung dieser Hypothese wurden zum einen die aus den Zielen abgeleiteten Produkte und Ergebnisse der Entwicklungsarbeit erfasst und zum anderen die Entwicklungsziele nach der Qualität ihrer Beschreibung bewertet. In einem weiteren Schritt wurden Qualität und Quantität der Ziele durch Gewichtung in einem gemeinsamen Wert zusammengefasst. Hierfür wurden die drei wichtigsten Ziele und Unterziele sowie Arbeitsschritte zur Zielerreichung der schulischen Steuergruppen erfasst und bewertet (N = 227 Ziele aus 90 Schulen; vgl. Jäger 2004, S. 179-181). Als Bewertungskriterien wurden die Verständlichkeit sowie der Detaillierungsgrad der Ziele bestimmt und anhand von vier Kategorien eingeschätzt. Nachfolgend werden die von Jäger entwickelten Kategorien dargestellt, anhand derer die Einteilung der Ziele bezüglich ihrer Qualität vorgenommen wurde (Jäger 2004, S. 182-185):

**sehr spezifisch:** Ziele, Teilziele und Arbeitsschritte sind spezifisch und realistisch: Aus der Zielbeschreibung gehen die pädagogische Zielsetzung und die konkreten Schritte zur Erreichung des Ziels hervor. Es muss hervorgehen, wie die Teilziele erreicht werden.

**eher spezifisch:** Teilziele oder Arbeitsschritte sind unspezifisch: Hierbei sind die Ziele auf der obersten Ebene vergleichbar mit der ersten Kategorie (sehr spezifisch); die Beschreibung der Teilziele oder Arbeitsschritte erfolgt hingegen unpräzise, schlagwortartig und wenig detailliert.

---

<sup>43</sup> BLK-Modellversuchsprogramm QuiSS – Qualitätsverbesserung in Schulen und Schulsystemen: Laufzeit 1999 bis 2004 unter der Beteiligung von 14 Bundesländern.

**eher unspezifisch:** Teilziele und Arbeitsschritte sind unspezifisch oder unrealistisch: Die Ziele auf der oberen Ebene sind vergleichbar zur ersten Kategorie, jedoch sind sowohl Teilziele als auch Arbeitsschritte eher schlagwortartig und oberflächlich beschrieben, so dass aus der Beschreibung keine Ableitung des konkreten Vorgehens möglich ist.

**sehr unspezifisch:** Ziele sind unklar, nicht weiter spezifiziert bzw. nicht nachvollziehbar: Die Beschreibung bleibt auf der Ebene des Hauptziels ohne weitere Unterteilung, so dass Zusammenhänge nicht deutlich werden und die Ableitung konkreter Maßnahmen nicht sichtbar wird.

Resümierend kann festgehalten werden, dass sich die Unterteilung und Beschreibung dieser Kriterien zur Bestimmung der Zielqualität auch für die Analyse von Zielvereinbarungen eignet.

### 9.2.2. Darstellung des eingesetzten Kategoriensystems

Zur Beantwortung der Forschungsfragen wurde ein Kategoriensystem entwickelt, welches das Gerüst zur Bestimmung der Form, Inhalte und Qualität der zwischen Schulen und Schulaufsicht vereinbarten Ziele enthält (s. Tab. 20). Zusätzlich beinhaltet es in Form von Schulkennungen, Schulformen und Schuljahren sowie den Qualitätsaspekt-Bewertungen Identifikationskennzahlen, die z. T. außerhalb der Zielvereinbarungsdokumente erhoben wurden. Dazu wurden in einem zusätzlichen Analyseschritt parallel zur Sichtung der Zielvereinbarungen die entsprechenden Qualitätsberichte der Schulen in der Bezirksregierung herangezogen und durch die Zuordnung der Bewertungsstufen der einzelnen Qualitätsaspekte zu jedem Zielvereinbarungsdokument Zusammenhangsanalysen ermöglicht.

Die rein formalen Texteingenschaften der Zielvereinbarungen sind durch das frequenzanalytische Auszählen von Seiten- und Worthäufigkeiten<sup>44</sup> ungeachtet ihrer Bedeutung mühelos zu bestimmen. Zu diesem Zweck werden die Messeinheiten Seiten- und Wörteranzahl als Indikatoren herangezogen.

---

<sup>44</sup> Bei der Bestimmung der formalen Merkmale werden die Angaben der Dokumentenvorlage nicht mitgezählt.

**Tab. 20: Eingesetztes Kategoriensystem zur inhaltsanalytischen Auswertung von Zielvereinbarungen**

<b>Dimension</b>	<b>Kategorie/Variable</b>	<b>Ausprägung</b>
<b>Formale Identifikationskennzahlen</b> <i>Ebene Zielvereinbarung und Einzelziele</i>	Schulkennung	101 – 721 (114 Schulkennungen)
	Schulform	1. Grundschule 2. Förderschule 3. Hauptschule 4. Realschule 5. Gymnasium 6. Gesamtschule 7. Berufskolleg
	Schuljahr der Zielvereinbarung	1. 2006/07 2. 2007/08 3. 2008/09 4. 2009/10 5. 2010/11
	Bewertung <sup>45</sup> der Qualitätsaspekte auf Einzelschulebene	1. erheblich entwicklungsbedürftig 2. eher schwach als stark 3. eher stark als schwach 4. vorbildlich
<b>Formale Merkmale</b> <i>Ebene Zielvereinbarung</i>	Umfang des Zielvereinbarungsdokuments	1. Anzahl der Seiten 2. Anzahl der Wörter
	Zielhäufigkeit	Anzahl der genannten Ziele
	Teilzielhäufigkeit	Anzahl der genannten Teilziele
<b>Zielinhalt</b> <i>Ebene Einzelziele</i>	Themenbereiche der Ziele in Anlehnung an die <ul style="list-style-type: none"> <li>• Qualitätsbereiche</li> <li>• Qualitätsaspekte</li> <li>• Qualitätskriterien des Qualitätstableaus NRW</li> </ul>	Häufigkeit der genannten Zielinhalte gemäß der Struktur des Qualitätstableaus <ul style="list-style-type: none"> <li>• Qualitätsbereiche 1 – 6</li> <li>• Qualitätsaspekte 1.1 – 6.3</li> <li>• Qualitätskriterien 1.1.1 – 6.3.5</li> </ul>
<b>Zielqualität</b> <i>Ebene Zielvereinbarung und Einzelziele</i>	Grad der Zielspezifität	1. unspezifisch (1) 2. spezifisch (2) 3. sehr spezifisch (3)

Als weiteres Kennzeichen des Kategoriensystems ist eine Bestimmung der *Zielinhalte* entlang des Qualitätstableaus der Qualitätsanalyse NRW vorgesehen. Hierbei ist keine umfassende interpretatorische Leistung erforderlich, da die Zuordnung der Ziele zum Tableau sehr eindeutig vorgenommen werden kann und die Angabe des Qualitätsbereichs in der Vorlage der Zielvereinbarungen verlangt wird und somit eine inhaltliche Orientierung vorgibt.

<sup>45</sup> Die Bewertung der Qualitätsaspekte geht nicht aus den Zielvereinbarungsdokumenten hervor. Im Vorfeld der Inhaltsanalyse wurden für alle Einzelschulen, zu denen Zielvereinbarungen vorlagen, die Bewertungsstufen aller Aspekte erfasst, um sie bei der Auswertung zu den Zielinhalten in Beziehung setzen zu können. Bei der Dateneingabe konnte dies durch die Vergabe von Schulkennungen anonymisiert erfolgen.

Zur Bestimmung der Bewertungskriterien der *Zielqualität* kam eine intensitätsanalytische Auswertungsstrategie zum Einsatz, die eine Einschätzung der Zielqualität mithilfe graduell abgestufter Ausprägungen der ordinalen Kategorien erlaubt und danach fragt, inwiefern die Ziele *unspezifisch*, *spezifisch* oder *sehr spezifisch* beschrieben sind. Tabelle 21 beinhaltet die Beschreibung der einzelnen Ausprägungen der Zielspezifität zur Bestimmung der Zielqualität.

**Tab. 21: Ausprägungen der Zielspezifität zur Bestimmung der Zielqualität**

Grad der Zielspezifität	Beschreibung des Grads der Zielspezifität
Stufe 1 <i>unspezifisch</i>	<ul style="list-style-type: none"> <li>• Das Ziel wird zwar schlagwortartig benannt, jedoch wird es nicht durch die Verwendung von Teilzielen weiter spezifiziert bzw. ist nicht nachvollziehbar.</li> <li>• Die Beschreibung erfolgt ohne genaue Angabe von Maßnahmen <i>und</i> Indikatoren und ohne weitere Unterteilung, so dass Zusammenhänge nicht deutlich werden und die Ableitung konkreter Maßnahmen nicht sichtbar wird.</li> <li>• Die Formulierungen der Qualitätsbereiche oder Qualitätsaspekte gemäß des Qualitätstableaus werden für die Zielbeschreibung übernommen, ohne spezifische Schritte zur Erreichung des Ziels zu benennen.</li> </ul>
Stufe 2 <i>spezifisch</i>	<ul style="list-style-type: none"> <li>• Das Ziel und ggf. auch die Teilziele sind spezifisch, jedoch erfolgt die Beschreibung der Indikatoren und Maßnahmen unpräzise, schlagwortartig und wenig detailliert.</li> <li>• Eine Benennung von Maßnahmen <i>oder</i> Indikatoren erfolgt nicht.</li> <li>• Konkrete Maßnahmen werden nicht näher erläutert. Es werden nur Beispiele für Maßnahmen (durch ein „u.a.“ oder „z. B.“ gekennzeichnet) angeführt.</li> </ul>
Stufe 3 <i>sehr spezifisch</i>	<ul style="list-style-type: none"> <li>• Das Ziel, die Teilziele und Maßnahmen sind spezifisch beschrieben.</li> <li>• Aus der kleinschrittigen Zielbeschreibung gehen die pädagogischen Zielsetzungen und die konkreten Schritte zur Erreichung des Ziels klar hervor.</li> <li>• Die Beschreibung zur Zielerreichung umfasst eine spezifische Darstellung der Indikatoren <i>und</i> Maßnahmen und wird ggf. auch für Teilziele vorgenommen.</li> </ul>

Die graduell abgestuften Ausprägungen der Zielspezifität basieren auf einer Synopse der Bewertungskriterien von Zielen in Anlehnung an Jäger (2004) und Schulprogrammen in Anlehnung an Köller (2009). Die von Jäger aufgestellten vier Kategorien zur Bestimmung der Zielqualität wurden nach einer praktischen Erprobung an einem Teil des Zielvereinbarungsmaterials durch fünf Codierer im Rahmen eines universitären Kolloquiums als nicht praktikabel erachtet, da sie zu stark ausdifferenziert waren. Hingegen sind die Erläuterungen zu den Bewertungsstufen sehr präzise und anwendungsorientiert, so dass sie in abgewandelter Form übernommen werden konnten. Die von Köller vorgenommene dreistufige Ausdifferenzierung des Konkretisierungsgrads von Schulprogrammen war in einem ersten Pretestdurchgang besser anzuwenden.

So wurde schließlich im Kategoriensystem eine dreistufige Ausprägungsskala analog zu Köller mit den Skalenbeschreibungen analog zu Jäger gebildet.

Die Darstellung des entwickelten Kategoriensystems abschließend, sei auf die unterschiedlichen Ebenen bei der Auswertung der genannten Merkmale hingewiesen: Das Merkmal *Zielqualität* kann sowohl auf der Ebene der Einzelziele als auch auf der Ebene eines Zielvereinbarungsdokuments bestimmt werden, da ausgehend von der Spezifität der Einzelziele eine durchschnittliche Spezifität eines gesamten Dokuments berechnet werden kann. Bei dem Merkmal *Zielinhalt* ist dieses Vorgehen nicht möglich, so dass die Bestimmung der Zielinhalte somit lediglich auf der Ebene der Einzelziele erfolgt. Bei den formalen Merkmalen wird hingegen eine Bestimmung lediglich auf der Ebene eines Zielvereinbarungsdokuments als sinnvoll erachtet.

### 9.2.3. Codierung

Durch die Struktur der Dokumentenvorlage der Zielvereinbarungen wurde als Codiereinheit die Textpassage in den Zeilen der Zielvereinbarungsformulare bestimmt. Alle Dokumentenbestandteile wurden zunächst vollständig von der Autorin codiert. Die Hälfte der Dokumente wurde durch Zufallsauswahl zusätzlich von einer geschulten studentischen Hilfskraft der Westfälischen Wilhelms-Universität Münster codiert. Dabei wurde nicht das gesamte Kategoriensystem zugrunde gelegt, sondern lediglich der Ausschnitt zur Bestimmung der *Zielspezifität*. Dieses Vorgehen begründet sich dadurch, dass zum einen die übrigen Kategorien keine maßgebliche interpretatorische Leistung erfordern (z. B. die Bestimmung der Seiten- oder Zielanzahl) und auch die Bestimmung der Qualitätsbereiche und -aspekte überwiegend eindeutig erfolgt. Einzig die Codierung der *Zielspezifität* erfordert eine anspruchsvolle interpretatorische Leistung und trägt durch mehrfache Codierung zur Bestimmung der Objektivität bei. Trotzdem ist einzuräumen, dass sowohl der Einsatz mehrerer Codierer als auch die Codierung des gesamten Materials zweifelsohne sinnvoll gewesen wäre, sich allerdings aufgrund fehlender personeller Ressourcen nicht realisieren ließ.

Zur Überprüfung der Objektivität und Intercoderreliabilität wurden die Codierungen einer summativen Reliabilitätsprüfung unterzogen. Die Übereinstimmung der Codierungen der Autorin und der geschulten Codiererin wurde mithilfe des Jaccard-Koeffizienten ermittelt, der den relativen Anteil der Über-

einstimmungen beider Codiererinnen an denselben Zielvereinbarungen errechnet (vgl. Backhaus u.a. 2011, S. 403f.). Bei 156 unabhängig vorgenommenen Codierungen der Zielspezifität wurden 114 Übereinstimmungen erzielt, was einen Jaccard-Koeffizienten von 0,73 ergibt. Dieser Wert ist analog zum Korrelationsmaß zu interpretieren und kann als gute Übereinstimmung der Codierungen gewertet werden.

Die Auswertung der Zielvereinbarungen erfolgte computergestützt durch den Einsatz von MAXQDA, eines speziellen Programms für die qualitative Datenanalyse (vgl. Kuckartz/Rädiker 2010, S. 735-747; vgl. Kuckartz/Grunenberg 2010, S. 502; Kelle 2009, S. 488 f.). Die deskriptiven Berechnungen, Zusammenhangsbestimmungen sowie statistischen Tests für Mittelwerte wurden mithilfe der Statistiksoftware SPSS durchgeführt.

### **9.3. Ergebnisse**

Im Folgenden werden die formalen und inhaltlichen Merkmale der Zielvereinbarungen sowie die Qualität der Ziele auf der Ebene der Einzelziele und Zielvereinbarungen resp. Einzelschulen dargestellt. Anschließend wird der Zusammenhang zwischen den in den Zielvereinbarungen genannten Zielen und den durch die Qualitätsanalyse benannten relativen Schwächen einer Schule beleuchtet.

#### **9.3.1. Formale Merkmale der Zielvereinbarungen**

Die Ausprägungen der formalen Merkmale werden zunächst deskriptiv dargestellt und anschließend inhaltlich interpretiert. Die 114 untersuchten Zielvereinbarungen weisen insgesamt einen Umfang von 429 Seiten und 56.673 Wörtern auf und sind mit durchschnittlichen 497 Wörtern im Mittel 3,8 Seiten lang. Die Maßeinheit *Wörter* wird bei der Analyse formaler Zielvereinbarungsmerkmale zusätzlich herangezogen, da sie den Dokumentenumfang im Vergleich zur Maßeinheit *Seite(n)* exakter und unverfälscht angibt. Beide Umfangsmaßeinheiten zeichnen für die vorliegenden Zielvereinbarungen ein äußerst heterogenes Bild: Einige Zielvereinbarungen sind kurz und komprimiert verfasst; als Extrembeispiel kann eine einseitige Zielvereinbarung mit 29 Wörtern angeführt werden. Im Gegensatz dazu gibt es auch einige ausführliche und detailliert aus-

gearbeitete Dokumente: Auch hier sei als Ausreißer ein 22-seitiges Dokument mit über 2.500 Wörtern angeführt. Insgesamt sind die Zielvereinbarungen zum überwiegenden Anteil (75 %) durch einen überschaubaren Umfang von bis zu vier Seiten gekennzeichnet. Die Hälfte der Zielvereinbarungen ist auf zwei und drei Seiten verfasst. Der Modus beträgt drei Seiten. Jeweils 10 % der Zielvereinbarungen sind einseitig und fünfseitig konzipiert. Mehr als sechs Seiten lange Zielvereinbarungen liegen nur vereinzelt vor. In einem weiteren Schritt wird zu prüfen sein, ob der Dokumentenumfang tatsächlich mit der gemessenen Zielqualität zusammenhängt. Zuvor werden die Unterschiede des Umfangs der Zielvereinbarungen hinsichtlich der Schulformen und der Schuljahre betrachtet.

**Tab. 22: Umfang der Zielvereinbarungen in Seiten und Wörtern im Schulformvergleich**

Schulform	GS	FS	HS	RS	GY	BK	MW <sub>alle*</sub>
<b>n<sub>Zielvereinbarungen</sub></b>	<b>19</b>	<b>11</b>	<b>25</b>	<b>24</b>	<b>12</b>	<b>21</b>	<b>112</b>
<b>Seiten MW</b>	3,7	2,9	3,2	4,5	4,0	3,9	3,7
<b>SD</b>	1,2	2,2	2,5	3,6	2,4	4,4	2,7
<b>Min</b>	2	1	1	2	1	1	1
<b>Max</b>	6	8	13	19	9	22	22
<b>Summe</b>	<b>70</b>	<b>32</b>	<b>81</b>	<b>108</b>	<b>48</b>	<b>81</b>	<b>420</b>
<b>Wörter MW</b>	412	340	451	624	566	493	481
<b>SD</b>	139,5	320	482,7	457,3	399	549,2	391
<b>Min</b>	192	29	70	157	74	126	29
<b>Max</b>	690	1153	2508	1989	1323	2639	2639
<b>Summe</b>	<b>7.832</b>	<b>3748</b>	<b>11.279</b>	<b>14.984</b>	<b>6794</b>	<b>10.355</b>	<b>54.922</b>

N = 112; MW = Mittelwert, SD = Standardabweichung, Min = Minimum, Max = Maximum, GS = Grundschule, FS = Förderschule, HS = Hauptschule, RS = Realschule, GY = Gymnasium, BK = Berufskolleg.

\* ohne Gesamtschulen. Die Schulform Gesamtschule wird aus der deskriptiven Beschreibung der formalen Merkmale ausgenommen, da sie nur zwei Dokumente umfasst, die insgesamt 9 Seiten und 1.681 Wörter aufweisen.

Den größten durchschnittlichen Umfang weisen Zielvereinbarungen in der Schulform Realschule mit 4,5 Seiten auf, gefolgt von Gymnasien und Berufskollegs mit durchschnittlich je 4 und 3,9 Seiten (s. Tab. 22). Die komprimiertesten Zielvereinbarungen können in der Schulform Förderschule mit einer mittleren Seitenanzahl von 2,9 beobachtet werden. In den Grundschulen weisen die Zielvereinbarungen einen mittleren Umfang von 3,7 Seiten auf. Hierbei

zeigt sich ein innerhalb der Schulform homogenes Bild bezüglich des Umfangs in Seiten und Wörtern mit einer vergleichsweise geringen Standardabweichung von 1,2 Seiten und 139,5 Wörtern. Die größten Streuungen zeigen sich aufgrund der Ausreißer der 19- bzw. 22-seitigen Dokumente in den Schulformen Realschule und Berufskolleg. Allerdings ergab die Berechnung einer Varianzanalyse, dass sich der Umfang der Dokumente zwischen den Schulformen und Schuljahren statistisch nicht signifikant voneinander unterscheidet.

Neben der Bestimmung des Zielvereinbarungsumfangs anhand der Seiten- und Wörteranzahl wird ergänzend die Anzahl der Ziele und Teilziele eines Zielvereinbarungsdokuments als weiteres formales Merkmal herangezogen. Teilziele stellen die den Einzelzielen untergeordnete Hierarchieebene dar. Sie strukturieren die Einzelziele, sind als Zielbestandteile zu interpretieren und tragen zu einer hohen Zielspezifität bei. Hinsichtlich der Anzahl der formulierten Ziele und Teilziele ergibt sich über alle 114 Zielvereinbarungsdokumente eine Gesamtzahl von 339 Zielen und 541 Teilzielen. Insgesamt beinhalten 40 % der Zielvereinbarungsdokumente drei Ziele und ein Viertel der Dokumente weist eine Zielhäufigkeit von zwei auf, während insgesamt ein Zehntel der Dokumente mehr als vier Ziele beinhaltet. Über alle Schulformen hinweg weisen die Zielvereinbarungsdokumente eine durchschnittliche Häufigkeit von drei Zielen und 4,7 Teilzielen auf (s. Tab. 23).

**Tab. 23: Anzahl der Ziele und Teilziele in den Zielvereinbarungen im Schulformvergleich**

Schulform n (Zielvereinbarungen)	GS	FS	HS	RS	GY	BK	GE	Σ
<b>Anzahl Ziele absolut</b>	55	34	68	66	36	75	5	<b>339</b>
Durchschnittliche Anzahl innerhalb der Schulform	2,9	3,1	2,7	2,8	3,0	3,6	2,5	<b>3,0</b>
Prozentuale Anzahl über alle Schulformen	16 %	10 %	20 %	20 %	11 %	22 %	1 %	<b>100 %</b>
<b>Anzahl Teilziele absolut</b>	101	31	129	113	73	84	10	<b>541</b>
Durchschnittliche Anzahl innerhalb der Schulform	5,3	2,8	5,2	4,7	6,1	4,0	5,0	<b>4,7</b>
Prozentuale Anzahl über alle Schulformen	19 %	6 %	24 %	21 %	13 %	15 %	2 %	<b>100 %</b>
<b>Relation Ziel : Teilziel</b>	1 : 1,8	1 : 0,9	1 : 1,9	1 : 1,7	1 : 2	1 : 1,1	1 : 2	<b>1 : 1,6</b>

GS = Grundschule; FS = Förderschule; HS = Hauptschule; RS = Realschule; GY = Gymnasium; BK = Berufskolleg; GE = Gesamtschule; Σ = alle Schulformen; N = 114 Zielvereinbarungen

Wird die Zielhäufigkeit in Relation zur Anzahl der Zielvereinbarungsdokumente einer Schulform betrachtet, so zeigt sich neben dem Berufskolleg mit durchschnittlich 3,6 Zielen pro Zielvereinbarung die Schulform Förderschule mit einer mittleren Zielhäufigkeit von 3,1 als überdurchschnittlich. Hingegen weisen die Schulformen Gesamt-, Haupt-, Real- und Grundschule weniger als drei Ziele pro Zielvereinbarung auf. Wird die Zielhäufigkeit einer Schulform wiederum prozentual über alle Schulformen und somit stichprobengrößenabhängig betrachtet, so zeigt sich bei der Schulform Berufskolleg mit 75 Zielen (22 %), gefolgt von den Hauptschulen mit 68 Zielen (20 %) sowie den Realschulen mit 66 Zielen (20 %) die größte Zielhäufigkeit.

In einem zweiten Schritt wurde die Anzahl der Teilziele bestimmt. Dabei reicht die Spannweite von 2,8 Teilzielen in Förderschulen bis zu 6,1 Teilzielen in Gymnasien. Auffällig erscheint, dass die meisten derjenigen Schulformen, die Ziele in einer unterdurchschnittlichen Häufigkeit (< 3 Ziele) aufweisen, überdurchschnittlich viele Teilziele (> 4,7 Teilziele) formuliert haben: Dies trifft auf die Schulformen Grundschule, Hauptschule und Gesamtschule zu. Schließlich wurde die Relation von Zielen und Teilzielen auf der Ebene der Schulform berechnet: An Hauptschulen werden mit einem Verhältnis von 1 : 1,9 am häufigsten Teilziele formuliert, gefolgt von Realschulen (Verhältnis von 1 : 1,7). Die wenigsten Teilziele zeigen sich mit einem Verhältnis zu den Zielen von 1 : 0,9 bei den Zielvereinbarungen der Förderschulen.

### 9.3.2. Inhaltliche Merkmale der Zielvereinbarungen

In diesem Abschnitt wird eine Bestimmung der Zielinhalte vorgenommen, indem die Ziele inhaltlich den Qualitätsbereichen (QB), -aspekten und -kriterien des Qualitätstableaus NRW zugeordnet werden (zum Tableau s. MSW NRW 2008a). Wie Tabelle 24 illustriert, sind die meisten Ziele (62 %) dem *QB 2 Lernen und Lehren – Unterricht* und mit deutlichem Abstand (zu 16 %) dem *QB 6 Ziele und Strategien der Qualitätsentwicklung* zugeordnet. Als inhaltlich eher irrelevant sind in den Zielvereinbarungen mit jeweils weniger als 10 % die Qualitätsbereiche 1, 3,4 und 5 vertreten. Folglich spielen die Dimensionen *Ergebnisse der Schule, Schulkultur, Führung und Schulmanagement* und *Professionalität der Lehrkräfte* eine untergeordnete Rolle bei der Zieldefinition.

**Tab. 24: Prozentuale Häufigkeiten der Zielinhalte in den Qualitätsbereichen 1 bis 6 im Schulformvergleich**

Schulform n Ziele	GS n = 55	FS n = 34	HS n = 68	RS n = 66	GY n = 36	BK n = 75	GE n = 5	$\Sigma$ N = 339
QB 1	6 %	-	3 %	6 %	6 %	7 %	-	4 %
QB 2	56 %	68 %	66 %	68 %	56 %	59 %	60 %	62 %
QB 3	4 %	9 %	5 %	6 %	8 %	2 %	-	5 %
QB 4	9 %	3 %	9 %	6 %	14 %	5 %	20 %	9 %
QB 5	7 %	9 %	1 %	-	5 %	7 %	-	4 %
QB 6	18 %	11 %	16 %	14 %	11 %	20 %	20 %	16 %

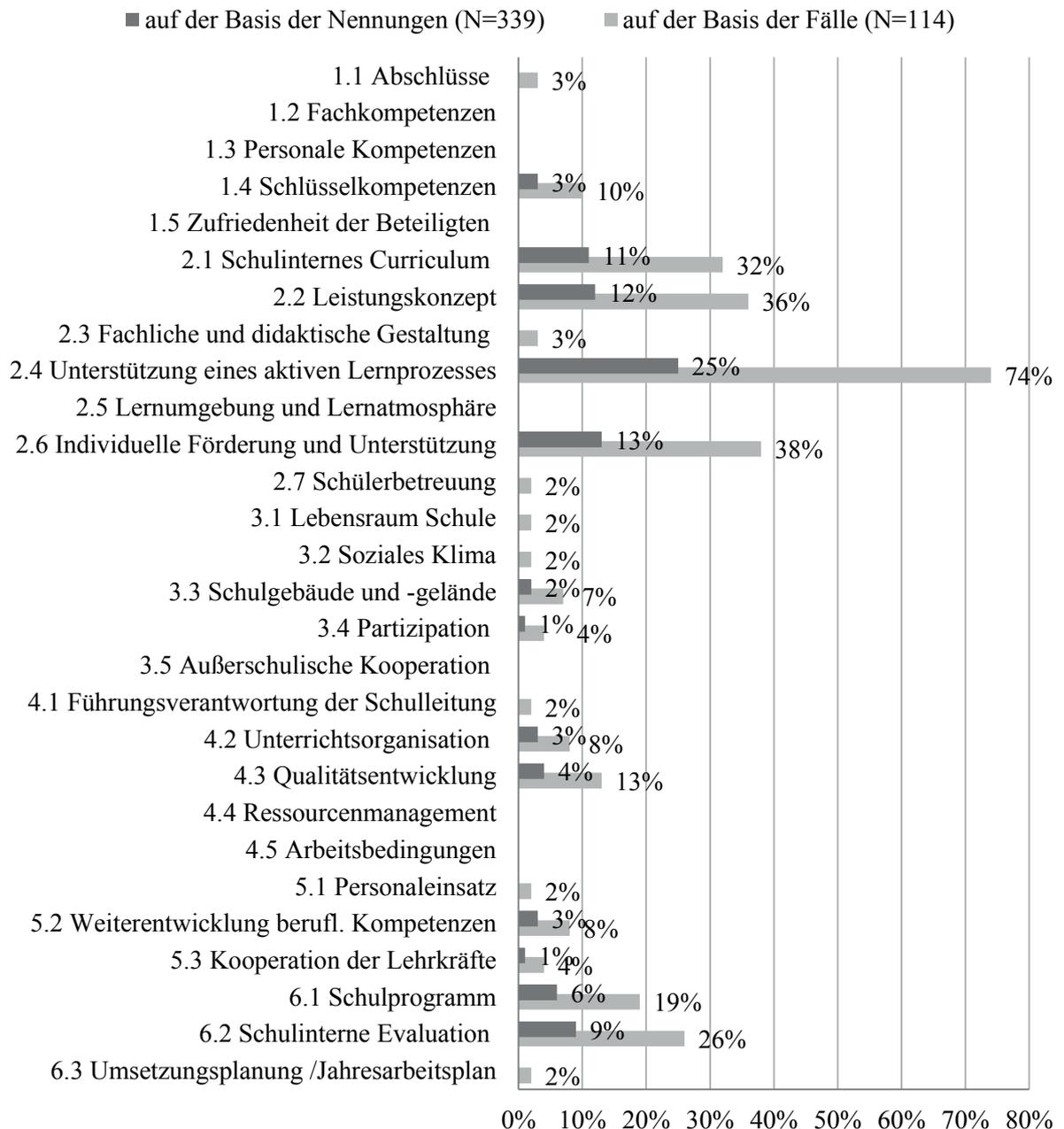
QB 1 = Ergebnisse der Schule; QB 2 = Lernen und Lehren – Unterricht; QB 3 = Schulkultur;  
 QB 4 = Führung und Schulmanagement; QB 5 = Professionalität der Lehrkräfte; QB 6 = Ziele und Strategien der Qualitätsentwicklung;  
 GS = Grundschule; FS = Förderschule; HS = Hauptschule; RS = Realschule; GY = Gymnasium;  
 BK = Berufskolleg; GE = Gesamtschule;  $\Sigma$  = alle Schulformen; N = 339 Ziele

Auf der Basis der Zielvereinbarungen (N = 114) sind 185% der Ziele dem *QB 2 Lernen und Lehren – Unterricht* zugeordnet. Das bedeutet, dass jedes Zielvereinbarungsdokument durchschnittlich nahezu zwei Ziele des QB 2 enthält. Mit deutlichem Abstand wurden in 47 % der Zielvereinbarungen die Ziele dem *QB 6 Ziele und Strategien der Qualitätsentwicklung* zugeordnet.

Bei der Zuordnung der Zielinhalte auf die einzelnen Qualitätsbereiche zeigt sich für die verschiedenen Schulformen eine im Wesentlichen homogene Verteilung ohne signifikante Unterschiede. Die Tendenz, dass der QB 2 innerhalb der Zielinhalte dominiert und der QB 6 mit deutlichem Abstand an zweiter Stelle rangiert, lässt sich für alle Schulformen übereinstimmend konstatieren (s. Tab. 24).

Die Zielinhalte lassen sich in einem weiteren Durchgang auch auf der Ebene der 28 Qualitätsaspekte bestimmen. Dies ist insofern sinnvoll, als dass die Qualitätsbereiche eine grobe Gliederung vorgeben und inhaltlich äußerst ausdifferenzierte Aspekte enthalten, so dass sich bei der Kategorisierung durch die Berücksichtigung der Qualitätsaspekte eine detailliertere und ausführlichere Einordnung der Zielinhalte durchführen lässt (s. Abb. 16). Innerhalb des hinsichtlich der Zielhäufigkeit dominanten *QB 2 Lernen und Lehren – Unterricht* zeigt sich wiederum ein vorherrschender Qualitätsaspekt (2.4): Der Aspekt 2.4 umfasst inhaltlich die *Unterstützung eines aktiven Lernprozesses* und beinhaltet dabei die Förderung der aktiven Teilnahme und des selbständigen Lernens der

Schüler im Unterricht durch die Berücksichtigung ihrer individuellen Lernwege, eigener Lösungsmöglichkeiten und ihrer Zusammenarbeit untereinander. Der Aspekt 2.4 ist in 25 % aller Ziele und in 74 % aller Zielvereinbarungen enthalten (s. Abb. 16).



**Abb. 16: Prozentuale Häufigkeit der Zielinhalte (> 1 %) auf der Ebene der Qualitätsaspekte 1.1 bis 6.3**

Als weitere häufig genannte Zielinhalte innerhalb des QB 2 gelten die Aspekte 2.6 *Unterricht – Individuelle Förderung und Unterstützung* (13 % der Ziele; 38 % der Zielvereinbarungen), 2.2 *Leistungskonzept – Leistungsanforderung und -bewertung* (12 % der Ziele; 36 % der Zielvereinbarungen) und 2.1 *Schulinternes Curriculum* (11 % der Ziele; 32 % der Zielvereinbarungen).

Innerhalb des am zweithäufigsten in den Zielen enthaltenen *Qualitätsbereichs 6 Ziele und Strategien der Qualitätsentwicklung* wird der Qualitätsaspekt 6.2 *Schulinterne Evaluation* am häufigsten genannt (9 % der Ziele; 26 % der Zielvereinbarungen).

Zusätzlich zu der Auswertung der Zielinhalte auf Ebene der Qualitätsbereiche und -aspekte wird im Folgenden eine Ausdifferenzierung der Zielinhalte auf Grundlage der nächsten nachgelagerten Gliederungsebene – den Qualitätskriterien – vorgenommen (s. Tab. 25). Die Auswertung stützt sich dabei auf die Kriterien der drei am häufigsten genannten Qualitätsaspekte: Den Qualitätsaspekten 2.4 – *Unterstützung eines aktiven Lernprozesses* (25 % aller Ziele), 2.6 – *Individuelle Förderung und Unterstützung* (13 %) sowie 2.2 – *Leistungskonzept – Leistungsanforderung und Leistungsbewertung* (12 %). Diesen Qualitätsaspekten lassen sich insgesamt 167 Ziele zuordnen. Bei der differenzierten Betrachtung des Qualitätsaspekts 2.4 wird deutlich, dass das Kriterium 2.4.1 „Der Unterricht fördert eine aktive Teilnahme der Schülerinnen und Schüler“ mit 45 % von den Schulen am häufigsten als Zielinhalt genannt wird. Bei der genaueren Analyse des Qualitätsaspekts 2.6 zeigen sich neben überwiegend wenig berücksichtigten zwei häufig genannte Kriterien (2.6.1 und 2.6.3). Dabei kann konstatiert werden, dass im Bereich *Individuelle Förderung* eher globale Ziele wie die Verfügbarkeit von Instrumenten und Kompetenzen (59 %) sowie eine systematische Förderung eines individualisierenden Unterrichts (23 %) im Vordergrund stehen. Deutlich weniger Priorität erfahren spezifische Themen der individuellen Förderung, welche leistungsschwächere Schüler und solche mit besonderen Begabungen oder Konzepte zur Sprach- und Leseförderung fokussieren. Analoges gilt für den Qualitätsaspekt 2.2: Das eher allgemein formulierte Kriterium „Die Schule hat für alle Fächer Grundsätze zur Leistungsbeurteilung festgelegt“ wird innerhalb des Aspekts in 79 % der Fälle genannt, während die Transparenz und Praktizierung dieses Konzepts in der Prioritätensetzung weniger Relevanz erfährt (16 % und 5 %; s. auch Tab. 25).

**Tab. 25: Prozentuale Häufigkeiten der Ziele auf Ebene der Kriterien der Qualitätsaspekte 2.4, 2.6 und 2.2 (absteigend sortiert)**

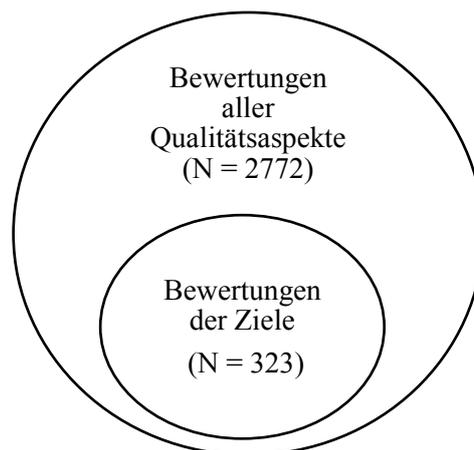
<b>2.4 Unterricht – Unterstützung eines aktiven Lernprozesses (N = 87 Ziele)</b>		Häufigkeit
2.4.1	Der Unterricht fördert eine aktive Teilnahme der Schülerinnen und Schüler.	45 %
2.4.2	Der Unterricht fördert die Zusammenarbeit zwischen den Schülerinnen und Schülern und bietet ihnen Möglichkeiten zu eigenen Lösungen.	6 %
2.4.3	Der Unterricht berücksichtigt die individuellen Lernwege der einzelnen Schülerinnen und Schüler.	14 %
2.4.4	Die Schülerinnen und Schüler erhalten Gelegenheit zu selbständiger Arbeit und werden dabei unterstützt.	23 %
2.4.5	Der Unterricht fördert strukturierte und funktionale Partner- bzw. Gruppenarbeit.	12 %
<b>2.6 Individuelle Förderung und Unterstützung (N = 41 Ziele)</b>		Häufigkeit
2.6.1	Die Schule verfügt über Instrumente und Kompetenzen zur individuellen Lernstandsdiagnose und Förderplanung.	59 %
2.6.2	Die Schule verwirklicht ein Konzept zur systematischen Förderung leistungsschwächerer Schülerinnen und Schüler.	7 %
2.6.3	Die Schule unterstützt die systematische Förderung eines individualisierenden, differenzierenden Unterrichts.	23 %
2.6.4	Die Schule fördert Schülerinnen und Schüler mit besonderen Begabungen.	2 %
2.6.5	Die Schule verwirklicht ein Konzept zur systematischen Sprachförderung.	7 %
2.6.6	Die Schule verwirklicht ein Konzept zur Leseförderung.	2 %
<b>2.2 Leistungskonzept – Leistungsanforderung und Leistungsbewertung (N = 39 Ziele)</b>		Häufigkeit
2.2.1	Die Schule hat für alle Fächer Grundsätze zur Leistungsbewertung festgelegt.	79 %
2.2.2	Alle Beteiligten (u.a. Schülerinnen und Schüler) kennen die vereinbarten Grundsätze zur Leistungsbewertung.	16 %
2.2.3	Die Lehrkräfte der Schule halten sich an die Grundsätze zur Leistungsbewertung.	5 %
2.2.4	Die Schule honoriert besondere Leistungen der Schülerinnen und Schüler.	-

Die unterschiedliche Gewichtung einzelner Qualitätsaspekte und Kriterien innerhalb eines Qualitätsbereichs zeigt, dass eine differenziertere Betrachtung der Zielinhalte auf Aspekt- und Kriterienebene gefordert ist. Die Angabe der Qualitätsbereiche 1 bis 6 bietet zwar eine anfängliche Orientierungsfunktion für die Kategorisierung der Zielinhalte, bleibt jedoch inhaltlich zu vage und unspezifisch.

### 9.3.3. Zusammenhang zwischen den Zielinhalten und den festgestellten Stärken und Schwächen der Schulen

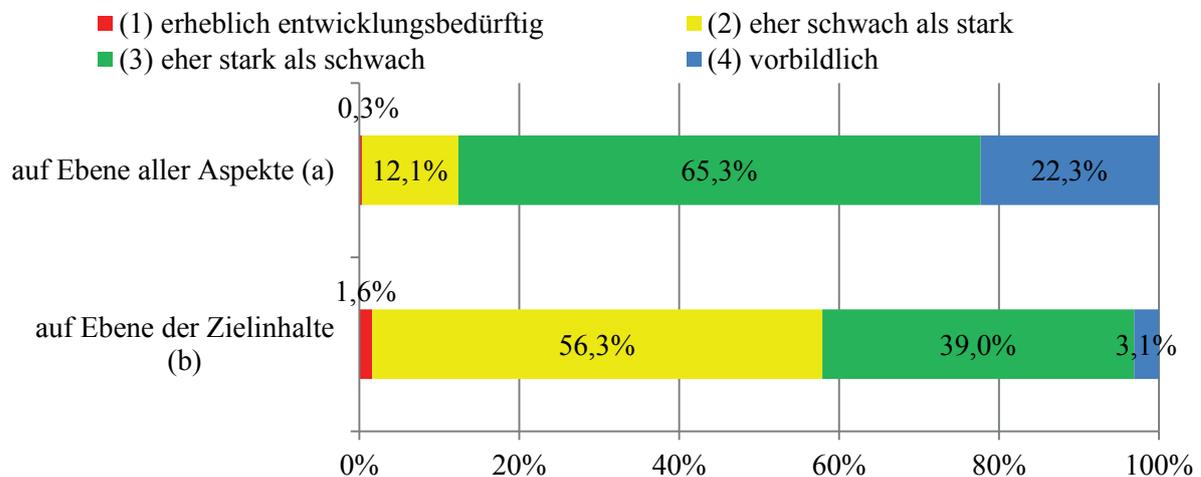
Zur Prüfung des Zusammenhangs zwischen den Zielinhalten und den diagnostizierten Schwächen und Stärken der Schulen wird in einem ersten Schritt zunächst die Verteilung der Bewertungsstufen von 1 (erheblich entwicklungsbedürftig) bis 4 (vorbildlich) für *alle* bewerteten Qualitätsaspekte der 114 Schulen ermittelt. Da drei der insgesamt 28 Qualitätsaspekte (1.1; 1.2; 4.5) bislang von der Qualitätsanalyse nicht bewertet wurden und für drei der 114 Schulen keine Bewertungen vorliegen, ergibt sich als Gesamtheit aller bewerteten Qualitätsaspekte eine Anzahl von 2772 Bewertungen. Die durchschnittliche Bewertung der Qualitätsaspekte entspricht der Bewertungsstufe 3 (eher stark als schwach;  $MW = 3,1$ ;  $SD = ,59$ ).

In einem zweiten Schritt erfolgt eine Gegenüberstellung aller bewerteten Qualitätsaspekte ( $N = 2772$ ) zu den Bewertungen der Qualitätsaspekte auf der Ebene der Ziele ( $N = 323$ ; für drei Schulen liegen keine Bewertungen vor). Die 323 bewerteten Qualitätsaspekte auf der Ebene der Ziele stellen dabei eine Teilmenge bzw. Subgruppe aller bewerteten Qualitätsaspekte dar (s. Abb. 17) und wurden im Mittel mit der Bewertungsstufe 2 (eher schwach als stark) beurteilt ( $MW = 2,4$ ;  $SD = ,58$ ).



**Abb. 17: Bewertung der Ziele ( $N = 323$ ) als Teilmenge aller bewerteten Qualitätsaspekte ( $N = 2772$ )**

Der Befund des t-Tests (one sample t-test) zeigt, dass die Zielbewertungen mit  $T = -20,452$ ,  $df = 322$  und  $p \leq ,001$  vom zugrunde gelegten Mittelwert der Bewertungen aller Qualitätsaspekte signifikant abweichen. D. h. die Bewertungen der Ziele sind nicht mit den Bewertungen der Referenzstichprobe identisch. Die Unterschiede zwischen den Bewertungen aller Qualitätsaspekte und der Zielinhalte werden in einem weiteren Schritt spezifiziert und die einzelnen Bewertungsstufen in ihrer prozentualen Häufigkeit dargestellt. Abbildung 18 zeigt die Gegenüberstellung des Verhältnisses der Bewertungsstufen einerseits über alle Qualitätsaspekte und andererseits über diejenigen auf der Ebene der Zielinhalte.



a: auf der Ebene von 25 bewerteten Aspekten der Gesamtstichprobe (N=2772 Bewertungen), da drei Aspekte (1.1, 1.2, 4.5) im Rahmen der Qualitätsanalyse nicht bewertet werden.  
 b: auf der Ebene aller Zielbewertungen (n=323), da für drei Schulen (16 Ziele) keine Bewertungen vorliegen.

**Abb. 18: Vergleichende Darstellung der Bewertungsstufen 1 bis 4 auf der Ebene der Ziele (N = 323) und aller bewerteten Qualitätsaspekte (N = 2772)**

Betrachtet man zunächst die Verteilungen der Bewertungsstufen über alle Aspekte hinweg (oberer Balken), so zeigt sich, dass die Bewertungsstufe 1 (erheblich entwicklungsbedürftig) äußerst selten (0,3 %) und die Bewertungsstufe 2 (eher schwach als stark) zu einem prozentualen Anteil von 12,1 % vergeben wird. Hingegen erhalten die meisten Schulen als Bewertung die Stufe 3 (eher stark als schwach; 65,3 %), gefolgt von der Stufe 4 (vorbildlich; 22,3 %). Wird dann die Verteilung der Bewertungsstufen aller Zielinhalte fokussiert (unterer Balken in der Abb. 18), so wird eine deutliche Verschiebung offensichtlich: Sowohl der Anteil der Bewertungsstufe 2 als auch 1 verfünffacht sich im Ver-

gleich zur Ebene aller bewerteten Qualitätsaspekte (1: Erhöhung von 0,3 % auf 1,6 %; 2: Erhöhung von 12,1 % auf 56,3 %). Damit werden in den Zielvereinbarungen zum überwiegenden Anteil (56,3 %) Qualitätsaspekte aufgegriffen, welche im Rahmen der Qualitätsanalyse als *eher schwach als stark* (Bewertungsstufe 2) eingestuft wurden. Hingegen werden weniger Ziele auf der Ebene von Aspekten vereinbart, die eine höhere Bewertungsstufe aufweisen. Im Vergleich der beiden Ebenen (alle Aspekte vs. Bewertungen der Ziele) lässt sich bei der Bewertungsstufe 3 ein Rückgang um 26,3 % auf 39 % und bei der Bewertungsstufe 4 ein Rückgang um 19,2 % auf 3,1 % beobachten. Der vergleichsweise hoch ausgeprägte Anteil der Stufe 3 auf der Ebene der Ziele ist mit der Definition der Bewertungsstufe 3 zu erklären: „Die Schule weist bei diesem Qualitätsaspekt mehr Stärken als Schwächen auf. Die Schule kann die Qualität einiger Kriterien noch weiter verbessern; die wesentlichen Kriterien werden erfüllt“.

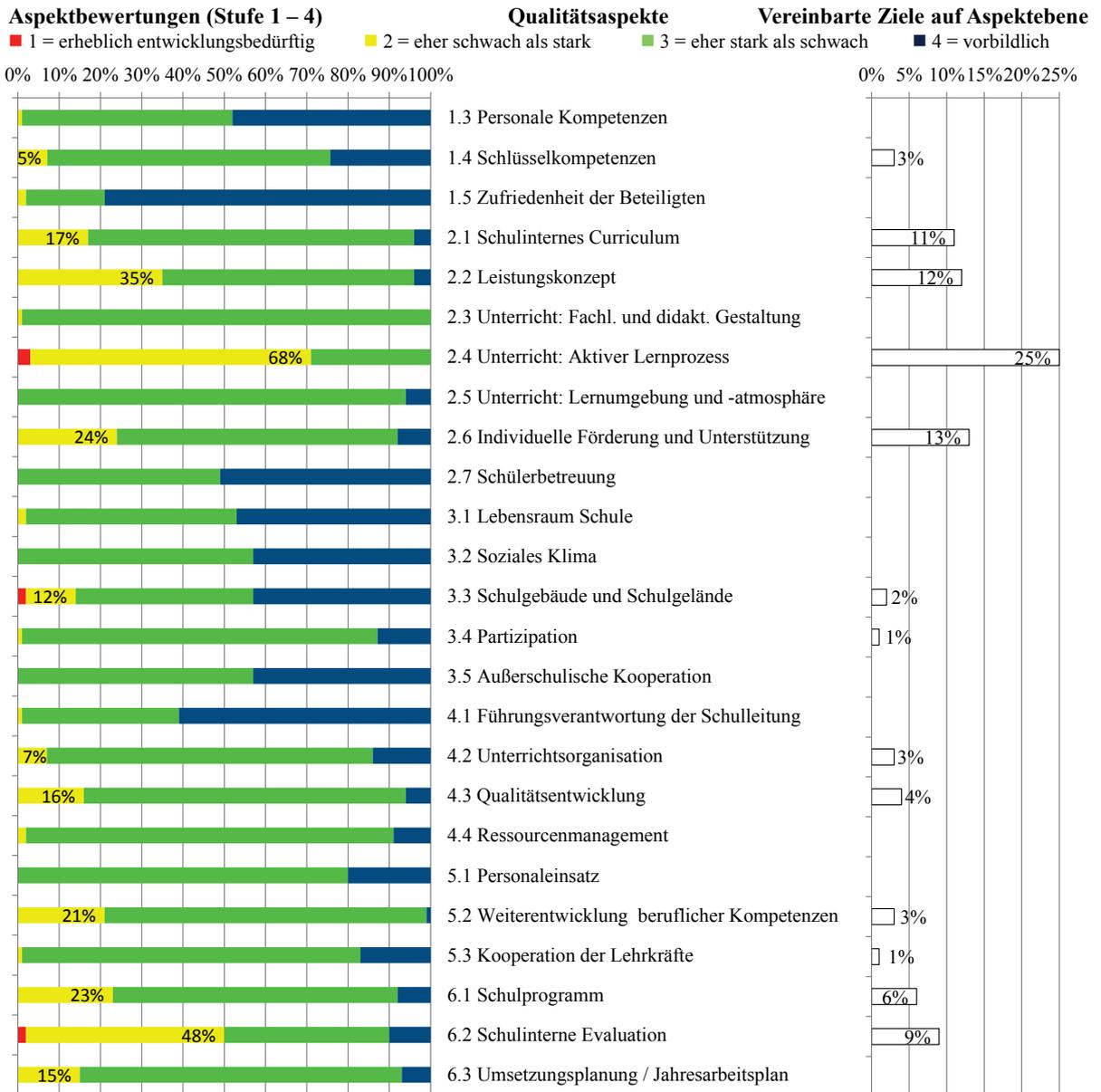
Im Rahmen der statistischen Überprüfung der Signifikanz dieser Unterschiede wurde die Frage geklärt, ob es eine signifikante Über- oder Unterrepräsentanz einzelner Bewertungsstufen bei den Qualitätsaspekten gibt, auf die sich die vereinbarten Ziele stützen. Hierzu wurden Abweichungen der beobachteten von den erwarteten Häufigkeiten beleuchtet. Die Ergebnisse der standardisierten Residuen<sup>46</sup> zeigen für die Aspekte, welche in den Zielen aufgegriffen werden, dass bei den Bewertungsstufen im negativen Bereich (Stufe 1 und 2) eine signifikante Häufung ( $p < ,001$ ), hingegen bei den Bewertungsstufen im positiven Bereich (Stufe 3 und 4) eine signifikante Unterrepräsentanz vorliegt. Auch der Chi-Quadrat-Test nach Pearson, der die Unabhängigkeit der beiden Variablen der Kreuztabelle überprüft, ergibt ein signifikantes ( $p < ,001$ ) Ergebnis und bestätigt, dass es einen Zusammenhang zwischen beiden Variablen (Bewertungsstufe und Ebene aller Qualitätsaspekte oder Zielinhalte) gibt und dass die Hypothese der Gleichverteilung der einzelnen Bewertungsstufen zu verwerfen ist.

Abbildung 19 veranschaulicht, auf der Basis welcher Qualitätsaspekte und der entsprechenden Aspektbewertungen Ziele vereinbart werden. Dabei werden die Ausprägungen der Bewertungsstufen der prozentualen Häufigkeit der Ziele

---

<sup>46</sup> Als Faustregel wurde ein standardisiertes Residuum von 3,3 oder größer herangezogen, welches eine Abweichung der beobachteten von der erwarteten Häufigkeit auf einem Signifikanzniveau von  $p < ,001$  anzeigt (vgl. Bühl 2008, S. 216).

gegenübergestellt. Zielinhalte mit Angaben von unter 1% sind in der Grafik nicht berücksichtigt. Augenscheinlich ist eine starke Relation zwischen eher schwach bewerteten Aspekten und den vereinbarten Zielen erkennbar.

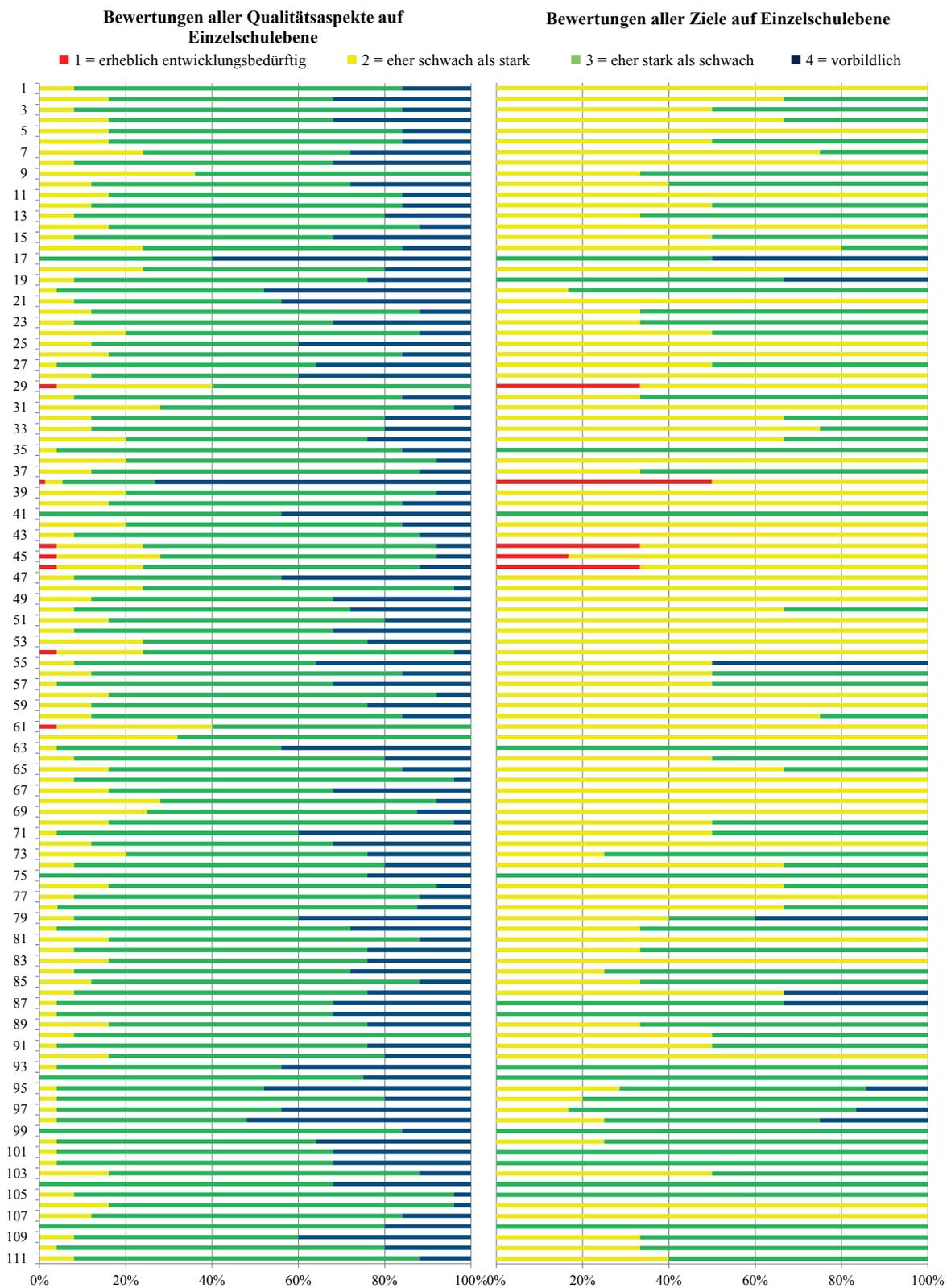


**Abb. 19: Kombinierte Darstellung der Aspektbewertungen und der Häufigkeit der Zielinhalte auf der Basis der Qualitätsaspekte (N = 114 ZV)**

Der Aspekt 2.4, dem der größte Anteil der vereinbarten Ziele zuzuordnen ist (25 % aller Nennungen), wird zu 68 % mit der Bewertungsstufe 2 bewertet. Weitere eher schwach bewertete Aspekte (6.2; 2.2; 2.6; 6.1; 5.2; 2.1) spiegeln sich sichtbar in den Zielen wider, während sich überwiegend positiv bewertete Aspekte (1.5; 4.1; 2.7; 3.1; 3.2) nicht in den Handlungsabsichten finden lassen. Insgesamt konzentrieren sich die Zielinhalte auf einen Ausschnitt des Qualitätstableaus, während viele Qualitätsaspekte bei der Zieldefinition wenig bis gar nicht berücksichtigt werden.

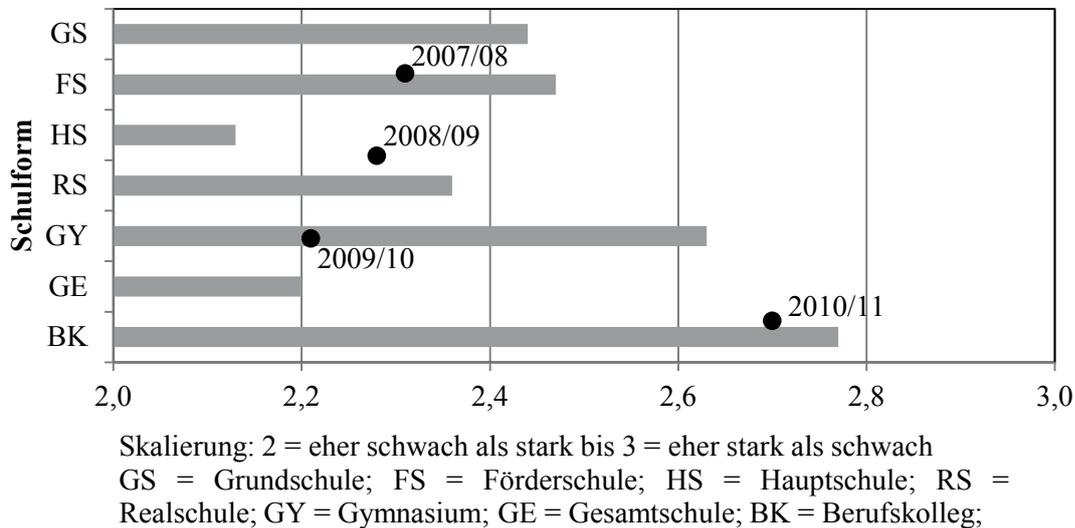
In Abbildung 20 ist die prozentuale Häufigkeit aller Aspektbewertungen und der Bewertungen der Zielinhalte auf Einzelschulebene gegenübergestellt. Insgesamt liegen für 111 Schulen entsprechende Daten vor. Während in der linken Spalte, d. h. den prozentualen Häufigkeiten der Bewertungen aller Qualitätsaspekte die Bewertungsstufen 3 und 4 (Farbgebung grün und blau) dominieren, werden bei der Formulierung von Zielen insbesondere diejenigen Qualitätsaspekte berücksichtigt, welche mit der Bewertungsstufe 2 (gelb) bewertet wurden.

Ein Vergleich der Einzelschulen offenbart einen durchaus sehr unterschiedlichen Umgang mit den Ergebnissen der Qualitätsanalyse: Während fast die Hälfte aller betrachteten Schulen (50 von insgesamt 111 Schulen) ausschließlich ihre als schwach bewerteten Qualitätsaspekte zum Fokus von Entwicklungsvorhaben macht und damit tatsächlich die gewünschte Reaktion auf die Ergebnismeldungen zeigt, verweigern sich offensichtlich einige Schulen einer solchen Reaktion. An neun Schulen zeigt sich, dass diese ihre eher schwach bewerteten Aspekte bei der Formulierung von Zielen nicht aufgreifen (Schulen 19, 35, 63, 87, 88, 93, 101, 102, 105) und Ziele in Bereichen definieren, in denen sie bereits *eher stark als schwach* aufgestellt sind. Die übrigen 52 Schulen wählen sowohl schwach als auch stark bewertete Qualitätsaspekte als Entwicklungsziele aus, wobei die schwach bewerteten Qualitätsaspekte häufiger Berücksichtigung finden, sich also auch hier eine Tendenz zur intendierten Reaktion nachweisen lässt. Insgesamt können die in den Abbildungen 19 und 20 dargestellten Zusammenhänge als unmittelbare, evidenzbasierte Reaktion auf gemessene schulische Schwachstellen gedeutet werden.



**Abb. 20: Gegenüberstellung aller Aspektbewertungen und der Bewertungen der Ziele auf Einzelschulebene (N = 111 Schulen)**

Im Folgenden werden die dargelegten Zusammenhänge um eine Betrachtung der durchschnittlichen Zielbewertungen differenziert nach Schulform und Schuljahr erweitert (s. Abb. 21).

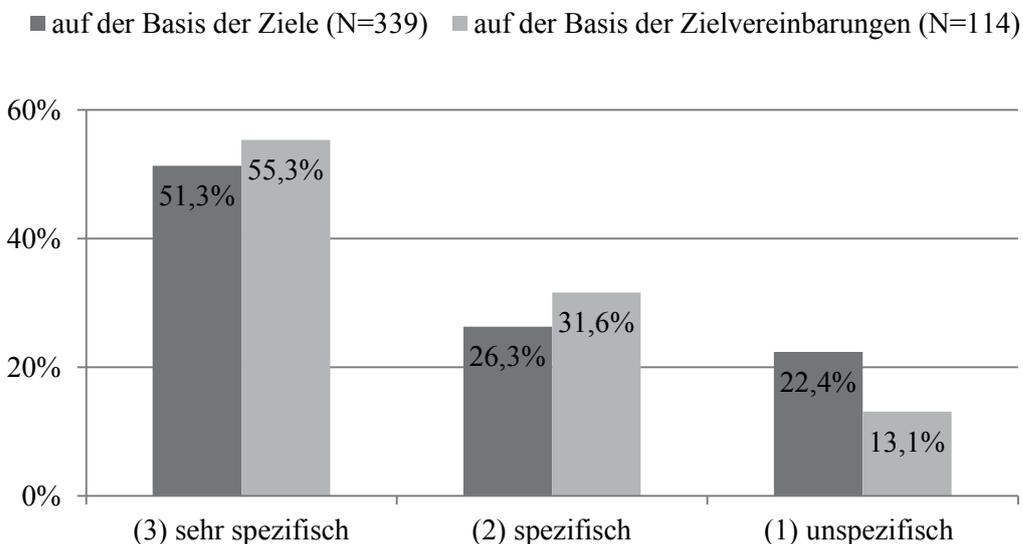


**Abb. 21: Durchschnittliche Zielbewertungen nach Schulform (Balken) und Schuljahr (Punkte); (N = 114 Zielvereinbarungen)**

Zwischen den Schulformen und Schuljahren ergeben sich statistisch signifikante Bewertungsunterschiede der Qualitätsaspekte, auf deren Grundlage die Ziele vereinbart wurden. Hinsichtlich der Schulformunterschiede liefert die Varianzanalyse (Einfaktorielle ANOVA) ein signifikantes Ergebnis ( $p < ,001$ ). Mittels Anwendung der Post-Hoc-Verfahren Duncan-Test und Scheffé-Prozedur ließ sich aufdecken, dass die Schulform Hauptschule durch geringere durchschnittliche Bewertungen der Zielinhalte signifikante Unterschiede zum Gymnasium sowie zum Berufskolleg aufweist. In einer weiteren Berechnung der Mittelwertunterschiede (einfaktorielle ANOVA) der Aspektbewertungen, auf deren Grundlage die Zielbestimmung erfolgt, zeigt sich im Schuljahresvergleich, dass sich das aktuellste Schuljahr 2010/11 signifikant positiv vom Schuljahr 2009/10 unterscheidet. Der Post-Hoc-Test identifiziert zwei homogene Untergruppen: In der ersten Gruppe versammeln sich die Schuljahre 2007/08 bis 2009/10 und die zweite Gruppe besteht aus dem Schuljahr 2010/11. Dies bedeutet, dass sich das Schuljahr 2010/11 von den übrigen drei Schuljahren statistisch signifikant durch höhere Bewertungen unterscheidet.

### 9.3.4. Zielqualität: Bestimmung der Zielspezifität

Die Zielqualität wird anhand des Kriteriums *Zielspezifität* mit den drei möglichen Ausprägungen *sehr spezifisch* (3), *spezifisch* (2) und *unspezifisch* (1) bestimmt. Dabei ist zu berücksichtigen, dass die Zielspezifität ein relevantes, jedoch nicht das alleinige Kriterium für Zielqualität abdeckt (s. SMART-Kriterien). Die Ergebnisdarstellung wird aus zwei Perspektiven vorgenommen: Zum einen auf der Ebene der Ziele (N = 339) und zum anderen auf der Ebene der Zielvereinbarungen (N = 114). Für die Bestimmung der Zielspezifität der einzelnen Dokumente wurde das arithmetische Mittel der einzelnen Ziele berechnet und durch Auf- oder Abrundung dieses Werts für das gesamte Dokument eine der drei Ausprägungen vergeben. Auf der Ebene der Ziele zeigt sich eine mittlere Zielspezifität von 2,3 (SD = ,81), d. h. die Ziele sind überwiegend *spezifisch* mit einer Tendenz zum Ausprägungsmerkmal *sehr spezifisch*. Die einzelnen Zielvereinbarungsdokumente weisen eine vergleichbare mittlere Zielspezifität von 2,4 (SD = ,72) auf. In Abbildung 22 zeigt sich folgende Verteilung der Zielspezifität:



**Abb. 22: Prozentuale Häufigkeit der Zielspezifität auf der Ebene der Ziele (N = 339) und der Zielvereinbarungen (N = 114)**

Auf der Basis der Ziele ist knapp über die Hälfte der Ziele *sehr spezifisch* formuliert, d. h. dass durch die zugehörigen Teilziele, Indikatoren und kleinschrittig beschriebenen Maßnahmen das konkrete Verfahren zur Erreichung des Ziels

klar hervorgeht. Etwa ein Viertel der Ziele ist *spezifisch* beschrieben, d. h. dass die Ziele und Teilziele zwar spezifisch beschrieben sind, jedoch die Bestimmung der Indikatoren und Maßnahmen unpräzise, schlagwortartig und wenig detailliert erfolgt. Ein weiteres knappes Viertel der Ziele ist schließlich *unspezifisch* verfasst, da ihr Inhalt nur schlagwortartig benannt wird und weder durch die Unterteilung in Teilziele weiter spezifiziert wird noch die Ableitung konkreter Maßnahmen erkennbar wird. Auf der Basis der Zielvereinbarungen ergeben sich geringfügig höhere Anteile der sehr spezifischen und spezifischen Ziele und ein zu fast 10 % geringerer Anteil unspezifischer Ziele.

In einem weiteren Schritt wurden die Zusammenhänge zwischen der Zielspezifität und den Qualitätsaspekten, Bewertungsstufen, Schulformen sowie Schuljahren untersucht. Auf Ebene der Qualitätsaspekte zeigt sich eine differenzierte Verteilung der Zielspezifität, welche sich allerdings statistisch als nicht signifikant erweist. Lediglich beim Aspekt *2.4 Unterricht – Unterstützung eines aktiven Lernprozesses* zeigt sich ein statistisch bedeutsamer unterrepräsentierter Anteil der unspezifischen Ziele (24 %) und ein überrepräsentierter Anteil von sehr spezifischen Zielen (65 %). Auch im Hinblick auf die Bewertungsstufen der Aspekte ergeben sich nicht signifikante Unterschiede in der Zielspezifität. Die Berechnung der standardisierten Residuen zeigt lediglich, dass in *eher stark als schwach* (3) bewerteten Aspekten überproportional häufig die Ziele unspezifisch formuliert sind. Während sich die Zielspezifität im Verlauf der Schuljahre nicht signifikant verändert, lassen sich diesbezüglich zwischen den Schulformen signifikante Unterschiede erkennen. Besonders in den Zielvereinbarungen der Realschulen ( $n = 66$ ) weisen die Ziele eine hohe Spezifität auf ( $MW = 2,7$ ), gefolgt von Gymnasien ( $n = 36$ ;  $MW = 2,5$ ). Dagegen kann für die Zielvereinbarungen der Förderschulen ( $n = 34$ ;  $MW = 1,8$ ) und Berufskollegs ( $n = 75$ ;  $MW = 2,0$ ) eine deutlich geringere Zielspezifität konstatiert werden. Der Post-Hoc-Test nach Duncan identifiziert die Förderschulen und Berufskollegs mit vergleichsweise weniger spezifischen Zielen als erste homogene Untergruppe, die Schulformen Grundschule, Hauptschule und Gymnasium konzentrieren sich mit einer mittleren Zielspezifität in der zweiten homogenen Untergruppe und schließlich wird die dritte Gruppe von der Schulform Realschule mit vergleichsweise überdurchschnittlich spezifischen Zielen gebildet.

### 9.3.5. Zusammenhang zwischen der Zielspezifität und formalen Merkmalen der Zielvereinbarungen

Die Kreuztabellen 26 und 27 enthalten die Ausprägungshäufigkeiten<sup>47</sup> der Zielspezifität unter der Bedingung einerseits der Seitenanzahl und andererseits der Zielanzahl innerhalb eines Zielvereinbarungsdokuments. Es zeigen sich z. T. beträchtliche Differenzen von bis zu 45 % der Zielspezifität in Abhängigkeit der formalen Merkmale der Zielvereinbarungen.

Hinsichtlich der Zielspezifität und der Seitenanzahl besteht ein schwacher, positiver und signifikanter ( $p < 0,005$ ) Zusammenhang von Kendall-Tau- $b = ,23$  (als anderes Zusammenhangsmaß für ordinal skalierte Daten wurde Gamma = ,33 berechnet). D. h. sehr spezifische Ziele gehen tendenziell mit einem höheren Zielvereinbarungsumfang einher, während unspezifische Ziele überwiegend auf einer Seite (45,5 % der unspezifischen Ziele) bzw. wenigen Seiten verfasst sind.

**Tab. 26: Kreuztabelle Zielspezifität und Seitenanzahl**

			Seitenanzahl					$\Sigma$
			1	2	3	4	> 5	
Zielspezifität	1 unspezifisch	Anzahl	5	2	4	2	2	15
		% innerhalb von Seitenanzahl	45,5 %	6,9 %	13,3 %	13,3 %	6,9 %	13,2 %
	2 spezifisch	Anzahl	3	15	6	6	6	36
		% innerhalb von Seitenanzahl	27,3 %	51,7 %	20,0 %	40,0 %	20,7 %	31,6 %
	3 sehr spezifisch	Anzahl	3	12	20	7	21	63
		% innerhalb von Seitenanzahl	27,3 %	41,4 %	66,7 %	46,7 %	72,4 %	55,3 %
Gesamt		Anzahl	11	29	30	15	29	114
N = 114 Zielvereinbarungen		% innerhalb von Seitenanzahl	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %	100 %

Im Gegensatz dazu zeigt sich hinsichtlich des Zusammenhangs der Zielspezifität und der Zielanzahl eine schwache, negative und signifikante ( $p < ,005$ ) Korrelation von Kendall-Tau- $b = -,25$  (Gamma = -,37). Somit geht ein höherer Grad der Zielspezifität mit einer geringeren Anzahl von Zielen innerhalb eines

<sup>47</sup> Die Verwendung der Chi-Quadrat-Verteilung und der Kreuztabellen mit den Abweichungen der beobachteten von den erwarteten Häufigkeiten war im vorliegenden Fall nicht zulässig (vgl. Bühl 2008; Voß/Khlayva/Schöneck 2008, S. 128).

Zielvereinbarungsdokuments einher. Der Anteil der sehr spezifischen Zielvereinbarungsdokumente ist besonders hoch, wenn das Dokument ein oder zwei Ziele umfasst (80 %; 71 %).

**Tab. 27: Kreuztabelle Zielspezifität und Zielanzahl**

			Anzahl der Ziele					Σ
			1	2	3	4	> 5	
Zielspezifität	1 unspezifisch	Anzahl	2	3	6	1	3	15
		% innerhalb von Zielanzahl	13,3 %	10,7 %	13,3 %	7,1 %	25,0 %	13,2 %
	2 spezifisch	Anzahl	1	5	18	5	7	36
		% innerhalb von Zielanzahl	6,7 %	17,9 %	40,0 %	35,7 %	58,3 %	31,6 %
	3 sehr spezifisch	Anzahl	12	20	21	8	2	63
		% innerhalb von Zielanzahl	80,0 %	71,4 %	46,7 %	57,1 %	16,7 %	55,3 %
Gesamt		Anzahl	15	28	45	14	12	114
N = 114 Zielvereinbarungen		% innerhalb von Zielanzahl	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %	100 %

## 9.4. Bewertende Zusammenfassung und Schlussfolgerungen aus der Zielvereinbarungsanalyse

Im Rahmen der inhaltsanalytischen Auswertung der Zielvereinbarungen wurden die formalen Merkmale dieser Dokumente bestimmt sowie die Inhalte und die Qualität der vereinbarten Ziele ermittelt. Bzgl. der Ausprägungen dieser Merkmale liegen gewisse aus der Theorie und Forschung abgeleitete Annahmen zu Grunde: So ist bspw. anzunehmen, dass Zielvereinbarungen am ehesten die Überführung von Zielen in tatsächlich umgesetzte Maßnahmen befördern können, wenn sie überschaubar, in der inhaltlichen Zielbestimmung eindeutig sowie in der Gesamtkonzeption durch die kleinschrittige Beschreibung konkreter Indikatoren zur Zielerreichung und den geplanten Maßnahmen möglichst spezifisch verfasst sind. Zudem sollten die Schulen bei der Zielauswahl evidenzbasiert auf ihre spezifischen, im Rahmen der Qualitätsanalyse identifizierten Handlungsfelder reagieren. Vor diesem Hintergrund kann eine Einschätzung abgegeben werden, ob die Erwartungen an die Qualitätsanalyse und das Instrument *Zielvereinbarung*, Schulentwicklungsprozesse zu initiieren, erfüllt werden könnten.

Bei der Analyse der 114 herangezogenen Zielvereinbarungsdokumente zeigt sich eine Streuung im Umfang der Dokumente und in der Ausführlichkeit der Darstellung. Die beobachteten Unterschiede erweisen sich allerdings zwischen den Schulformen und Schuljahren als nicht signifikant. Über die formalen Textmerkmale können Schlussfolgerungen zur Handhabbarkeit und Überschaubarkeit der Zielvereinbarungsdokumente gezogen werden. Der größte Teil der Zielvereinbarungen ist auf bis zu vier Seiten verfasst und bleibt somit im Umfang überschaubar und ist nicht überladen. Im Hinblick auf die Zielhäufigkeit können innerhalb der 114 Zielvereinbarungsdokumente 339 Ziele und 541 Teilziele angeführt werden, die mit durchschnittlich drei Zielen pro Zielvereinbarung noch gut überblickt werden können. Damit weisen die Dokumente einen vom Umfang abgeleiteten angemessenen Informationsgehalt auf.

Die Analyse der Zielinhalte auf der Grundlage des Qualitätstableaus NRW wurde mit der Intention durchgeführt, einerseits Aufschluss über aktuelle schulische Entwicklungs- und Handlungsbedarfe zu erhalten und andererseits unter Einbeziehung der Bewertungsstufen der Qualitätsaspekte eine evidenzbasierte Reaktion auf die Ergebnismeldungen zu überprüfen. Die Beschränkung bzw. Konzentration der vereinbarten Ziele auf vergleichsweise wenige Aspekte des gesamten Qualitätstableaus ist deutlich zu konstatieren. Mit Abstand am stärksten ist dabei der Qualitätsbereich *QB 2 Lernen und Lehren – Unterricht* in insgesamt 62 % der Ziele und in insgesamt 185 % der Zielvereinbarungen vertreten. Innerhalb des Qualitätsbereichs zeigte sich der Qualitätsaspekt *2.4 Unterstützung eines aktiven Lernprozesses* gefolgt vom Qualitätsaspekt *2.6 Individuelle Förderung und Unterstützung* in der Wahl der Zielbestimmung den anderen Themenfeldern deutlich überlegen. An zweiter Stelle folgt mit großem Abstand der *QB 6 Ziele und Strategien der Qualitätsentwicklung*, dem sich insgesamt 16 % der Ziele und 47 % der Zielvereinbarungen zuordnen lassen. Zum überwiegenden Anteil (58 %) werden die Ziele in Themenfeldern definiert, in denen die Schulen im Rahmen der Qualitätsanalyse als *eher schwach als stark* (Bewertungsstufe 2) bewertet worden sind. Die Bewertungen der Qualitätsaspekte, auf deren Grundlage die Ziele vereinbart wurden, weisen eine signifikante Häufung im negativen Bereich (Bewertungsstufen 1 und 2) und eine signifikante Unterrepräsentanz im positiven Bewertungsspektrum (Bewertungsstufen 3 und 4) auf. Dies lässt sich als Hinweis interpretieren, dass die Schulen

ihre von der Qualitätsanalyse identifizierten Handlungsfelder in den Zielvereinbarungen aufgreifen und damit – zunächst konzeptionell – direkt auf die Ergebnisse der Qualitätsanalyse reagieren. Dieser Befund kann als Indikator der direkten, evidenzbasierten Ergebnisnutzung der Qualitätsberichte für die Ausgestaltung der Zielvereinbarungen gedeutet werden. Insgesamt entsprechen die Ergebnisse der vorliegenden Inhaltsanalyse z. T. den bereits verfügbaren empirischen Befunden einer Berliner und einer bayerischen Schulleiterbefragung (vgl. Gärtner/Wurster 2009; Huber 2008). Analog zu den Ergebnissen der vorliegenden Zielvereinbarungsanalyse lässt sich in der Studie von Gärtner und Wurster (2009)<sup>48</sup> der Großteil der Zielinhalte dem Qualitätsbereich 2 *Lehr- und Lernprozesse* zuordnen. Allerdings ist der Anteil der Zielinhalte im Bereich des Unterrichts in NRW in deutlich stärkerem Ausmaß vertreten: Diesem Inhaltsbereich lassen sich in NRW etwa doppelt so viele Ziele zuordnen (62 % in NRW vs. 29 % in Berlin). Auf dem zweiten Rang sind auf der Grundlage des Berliner Orientierungsrahmens Zielinhalte des Qualitätsbereichs 5 *Lehrerprofessionalität und Personalentwicklung* vorzufinden (24 % aller Nennungen) – einem Bereich, der in NRW die geringsten Nennungen aufweist (4 % aller Nennungen). Werden die Erkenntnisse von Huber (2008) zum Vergleich herangezogen, so zeigt sich, dass die Mehrzahl der bayerischen Schulleiter (N = 311) den Eindruck bestätigt, die Zielvereinbarungen würden die wichtigsten Empfehlungen des Evaluationsberichts aufgreifen. Huber konkludiert auf der Grundlage seiner Befunde (Huber 2008, S. 275):

Insgesamt lässt sich feststellen, dass der Katalog an Maßnahmen, die zur Umsetzung der Zielvereinbarungen ergriffen wurden, genau jene Arbeitsfelder abdeckt, die sich bei der Auswertung der Evaluationsberichte der Schulen (...) als Handlungsfelder aufgedrängt haben.

Als Einschränkung des Vergleichs der Zielinhalte der drei Bundesländer ist das unterschiedliche Kategoriensystem – in Form des Qualitätstableaus NRW, des Orientierungsrahmens Berlin sowie des Qualitätstableaus zur externen Evaluation in Bayern – zu nennen. Des Weiteren sind die differierenden Methoden zu

---

<sup>48</sup> In der Studie von Gärtner und Wurster (2009) wurden Angaben von Schulleitungen zu den drei wichtigsten vereinbarten Zielen ausgewertet (N = 227 bzw. 180) und den Qualitätsbereichen und -merkmalen zugeordnet.

berücksichtigen, d. h. zum einen der Einsatz einer Dokumentenanalyse, zum anderen einer schriftlichen Befragung von Schulleitungen.

Schließlich wurde im Rahmen der Inhaltsanalyse auch die Qualität der einzelnen Ziele anhand der folgenden drei Ausprägungen der Zielspezifität bestimmt: sehr spezifisch (3), spezifisch (2) und unspezifisch (1). Die Ziele (N = 339) weisen eine mittlere Zielspezifität von 2,3 auf. Während sich keine signifikanten Unterschiede der Zielspezifität im Zusammenhang mit den Qualitätsaspekten beobachten lassen, werden Differenzen in der Zielspezifität zwischen einzelnen Schulformen deutlich. Wird die Zielqualität in Abhängigkeit der formalen Merkmale – gemessenen anhand der Seiten- und Zielanzahl – untersucht, so zeigen sich folgende signifikante und durch schwache bis mittlere Korrelationen nachweisbare Zusammenhänge: Die Berechnungen zeigen auf, dass sehr spezifische Ziele tendenziell mit einem höheren Dokumentenumfang einhergehen, während unspezifische Ziele eher auf einer bzw. wenigen Seiten verfasst sind. Auch geht ein höherer Grad der Zielspezifität mit einer geringeren Anzahl der vereinbarten Ziele eines Zielvereinbarungsdokuments einher.

Zusammengefasst zeigt sich neben einer eher breiten Spanne des Umfangs der Zielvereinbarungen ein für alle Schulen tendenziell homogenes Bild der inhaltlichen und qualitativen Merkmale. Insbesondere auf der Ebene der Zielinhalte ergeben sich über alle Schulen hinweg charakteristische Befunde, indem sowohl bei den gemessenen Schwächen als auch der Bestimmung der Zielinhalte das Kerngeschäft der Schulen – der Unterricht – mehrheitlich genannt wird. Insgesamt kann auf der Konzeptions- und Planungsebene eine evidenzbasierte Nutzung der Ergebnisrückmeldungen der Qualitätsanalyse und ein Anstoß von Schulentwicklungsprozessen als empirisch bestätigt gelten.

Ausgehend von den erzielten Ergebnissen lassen sich direkte Implikationen für die Ausgestaltung von Zielvereinbarungen und für die Steuerungsebene ableiten. Die untersuchten Zielvereinbarungen haben unter Berücksichtigung der Zielsetzungstheorie eine Reihe von Merkmalen erfüllt, welche das Handeln positiv beeinflussen können. Es empfiehlt sich je nach Zielanzahl ein Umfang der Zielvereinbarungsdokumente von zwei bis vier Seiten. Es sollten zudem möglichst nicht mehr als drei Ziele in einem Dokument vereinbart werden. Bei der Formulierung der Zielsetzungen sollte eine kleinschrittige, spezifische und möglichst in mehrere Teilziele differenzierte Zielbeschreibung vorgenommen

werden, aus der die konkreten Schritte des Ziels und die Darstellung der Indikatoren und Maßnahmen klar hervorgehen. Darüber hinaus könnte im Zielvereinbarungsprozess die Bereitstellung konkreter Beispiele, wie gute Zielvereinbarungen zu formulieren sind, hilfreich sein. Zum einen könnten – in einer vergleichbaren Form, wie dies im Bundesland Hessen erfolgt ist (vgl. Hessisches Kultusministerium 2007) – Zielvereinbarungsbeispiele aufgezeigt und kommentiert werden, zum anderen wären konkrete Kriterien nützlich, an denen beispielsweise die Zielspezifität festzumachen ist.

Die Analyse der Zielinhalte hat deutlich zeigen können, dass die Angabe des Qualitätsbereichs im Zielvereinbarungsformular nicht ausreichend erscheint, da die einzelnen Qualitätsbereiche stark ausdifferenziert sind. Hierbei wäre in den Zielvereinbarungsformularen die Änderung in die Angabe der entsprechenden Qualitätsaspekte ratsam. Für die Steuerungsebene ergeben sich aus der Kenntnis, in welchen Bereichen sich die schulischen Schwächen bündeln, gezielte Interventionsmöglichkeiten, z. B. in Form neuer Fortbildungskonzepte usw. Ob dieses Steuerungswissen tatsächlich in der Praxis mündet, wäre zu überprüfen und bietet einen Ansatzpunkt für weitere Forschung.

## **10. Studie III: Onlinebefragung der Schulaufsicht**

Auf der Grundlage des dargelegten Stands zur Schulinspektions- und Rezeptionsforschung ist deutlich geworden, dass die Erforschung des Rezeptions- und Nutzungsprozesses von Schulinspektionsergebnissen aus der Perspektive der Schulaufsicht ein in der Bildungsforschung in Deutschland noch nicht bearbeitetes Feld darstellt (vgl. Gärtner/Wurster 2009; Huber 2006; Döbrich/Gerecht/Lankert/Schnell 2007; van Ackeren 2003; Böttcher/Kotthoff 2007b). Um in diesem Zusammenhang die Forschungslücke zu schließen und die Perspektive der Schulaufsicht im Kontext der Schulinspektion zu untersuchen, fand im Sommer 2011 eine Onlinebefragung der Schulaufsicht in NRW statt. In diesem Kapitel erfolgt eine Vorstellung der methodischen Einbettung der Befragung, an welche die Darstellung des Erhebungsinstruments, der Stichprobe sowie der Ergebnisse anschließt.

### **10.1. Fragestellungen**

Auf der Grundlage des weiterentwickelten Rahmenmodells von Helmke und der multiperspektivischen Herangehensweise des Triangulationsdesigns dieser Arbeit lässt sich neben den Schulen die Schulaufsicht als weitere Adressatin der Ergebnisse von Schulinspektionen charakterisieren und nimmt auch im Nutzungsprozess eine aktive Rolle ein. Ergänzend zu den in den Studien I und II erhobenen Forschungsbefunden soll eine Datengrundlage generiert werden, die einen Überblick über einzelne Phasen und Rahmenbedingungen des Rezeptions-Nutzungs-Prozesses von Ergebnissen der Qualitätsanalyse aus der Innenperspektive der Schulaufsicht liefert. In Studie III werden folgende zentralen Fragestellungen fokussiert:

Wie gestaltet sich aus der Perspektive der Schulaufsicht ...

- der Umgang mit den Schulinspektionsergebnissen?
- der Zielvereinbarungsprozess?
- die Bereitstellung von Unterstützungsstrukturen im Anschluss an die Schulinspektion?

## 10.2. Methodisches Vorgehen: Onlinebefragung

Die Entscheidung, die Befragung der Schulaufsicht als Onlineerhebung durchzuführen, wurde weitgehend aus forschungsökonomischen Gesichtspunkten getroffen. Angesichts der wachsenden Bedeutung des Mediums Internet und der Nutzung neuer Medien im Arbeitsalltag war anhand des dienstlichen E-Mail-Verzeichnisses der Schulabteilungen ein direkter, umfassender Zugang zu den Mitgliedern der Schulaufsicht möglich. Über die dienstlichen E-Mail-Adressen konnte sichergestellt werden, dass die Zielgruppe uneingeschränkt über einen Internetzugang verfügt und somit eine Verzerrung der Stichprobe hinsichtlich dieses Merkmals auszuschließen war (vgl. Diekmann 2009, S. 530; Jacob u.a. 2011, S. 121; Couper/Coutts 2006, S. 228f.; Maurer/Jandura 2009, S. 65f.).

Im Allgemeinen bietet diese Erhebungsform eine Reihe von Vorteilen, die vor allem in den pragmatischen Aspekten wie der Zeiteffizienz bei der Erhebung, Auswertung und Präsentation der Daten sowie den geringen Kosten und dem Aufwand für Druck, Austeilung, Kodierung und Dateneingabe begründet sind (vgl. Thielsch/Weltzin 2009, S. 70; Kuckartz u.a. 2009, S. 9; 110f.; Diekmann 2009, S. 522f.; Jacob u.a. 2011, S. 116-118; Pötschke 2009, S. 77ff.). Weitere methodische Vorteile liegen im Ausbleiben von Interviewereinflüssen und Dateneingabefehlern sowie einem ehrlicheren Antwortverhalten, geringeren Effekten sozialer Erwünschtheit und der höher empfundenen Anonymität (vgl. Thielsch/Weltzin 2009, S. 71; Taddiken 2009, S. 94). Im Vergleich zu Papierfragebögen lassen sich zudem komplexe Filterführungen realisieren und das Befragtenverhalten aufzeichnen.

Neben den Vorteilen von Onlinebefragungen sind entsprechend auch methodenbedingte Nachteile im Forschungsprozess zu berücksichtigen. Zu diesen zählen die fehlende Kontrolle der Durchführungsbedingungen der Datenerhebung, welche zu Schwierigkeiten bei der Durchführungsobjektivität führen können. Weitere negative Methodeneffekte von Onlinebefragungen werden durch sog. Technologieeffekte, z. B. bei der Darstellung des Fragebogens durch verschiedene Software und dem kognitiv belastenderen und langsameren Lesen am Bildschirm verursacht (vgl. Taddicken 2009, S. 100).

Die Methodenwahl abschließend begründend wird herausgestellt, dass die Onlinebefragung – wie jede andere Methode auch – Vor- und Nachteile aufweist und für den vorliegenden Forschungszweck zielführend erscheint, da sie eine forschungsökonomische Befragungsmöglichkeit darstellt. Die vorliegende Befragung wurde internetgestützt als Web-Survey durchgeführt, d. h. der Fragebogen wurde auf einem Webserver abgelegt. Die Zielgruppe wurde via E-Mail eingeladen, über den Zugriff auf eine bestimmte URL (Uniform Resource Locator) an der Befragung teilzunehmen und konnte über die angegebene Webseite den Fragebogen online ausfüllen. Die Datenspeicherung erfolgte direkt auf dem Server. Als Softwareprogramm kam EFS Survey 8.0 von *Globalpark* zum Einsatz (vgl. Globalpark AG 2011). Dieses Programm gilt als Marktführer zur Durchführung von internetbasierten Befragungen im deutschsprachigen Raum und wird den Hochschulen als *Unipark* angeboten.

### **10.3. Erhebungsinstrument**

Generell gelten für die Erstellung von Onlinebefragungen die gleichen methodischen und konzeptionellen Anforderungen wie für traditionelle Befragungsformen auch (vgl. Porst 2011; Kirchhoff u.a. 2010; Jacob u.a. 2011, S. 124-133). Die Spezifika und Unterschiede von Onlinebefragungen im Vergleich zu herkömmlichen Befragungsformen sind vor allem in der Programmierung, der optischen Gestaltung und der Dokumentation begründet (vgl. Gräf 2010, S. 35f.; Maurer/Jandura 2009, S. 67).

#### **10.3.1. Fragebogenkonstruktion**

Grundlage für die Entwicklung des Fragebogens bildet das erweiterte Rahmenmodell von Helmke (s. Kap. 5). Der Fragebogen knüpft an die einzelnen Phasen (Rezeption, Reflexion und (geplante) Aktion) und Rahmenbedingungen aus der Perspektive der Schulaufsicht an und umfasst die in der Tabelle 28 dargestellten Themenkomplexe I bis V, welche thematisch ähnlich gelagerte Fragen beinhalten. Die Reihenfolge der Fragen orientiert sich an der Abfolge der Qualitätsanalyse (Bericht – Zielvereinbarungen – Unterstützungsstrukturen), so dass die Teilnehmer die Folgefragen als passend für den Fragebogenverlauf empfinden sollten. Der Onlinefragebogen umfasst insgesamt (ohne Filterfüh-

runge) 19 Seiten, auf denen die folgenden zentralen 18 Fragenbereiche erhoben werden (s. Tab. 28):

**Tab. 28: Inhaltliche Themen des Online-Fragebogens der Schulaufsicht**

<b>I. <i>Aufgaben der Schulaufsicht</i></b>
- Veränderungen der Tätigkeitsfelder
- Auseinandersetzung mit dem Instrument und den Ergebnissen der QA
- Hospitation bei einer Qualitätsanalyse
<b>II. <i>Qualitätsbericht</i></b>
- Anzahl gelesener Qualitätsberichte
- Qualität gelesener Qualitätsberichte
- Akzeptanz von Qualitätsberichten
<b>III. <i>Qualitätsanalyse</i></b>
- Erkenntnisgewinn durch Qualitätsanalysen
- Akzeptanz und Folgen der Qualitätsanalyse
- Professionalität der Qualitätsprüfer
<b>IV. <i>Zielvereinbarungen</i></b>
- Anzahl getroffener Zielvereinbarungen
- Positive Effekte der Zielvereinbarungen
- Zusammenhang der Ziele mit den Qualitätsberichten
- Häufigkeiten der Zielinhalte
- Inhaltliche Bestimmung der Zielvereinbarungen
- Überprüfung der Einhaltung der Zielvereinbarungen
- Zielvereinbarungsprozess
<b>V. <i>Unterstützungsstrukturen im Anschluss an die Qualitätsanalyse</i></b>
- Anzahl bestimmter Unterstützungsleistungen
- Angebot an Unterstützungsstrukturen

Da der Rezeptions-Nutzungs-Prozess der Ergebnisse aus Qualitätsanalysen durch die in der Schulaufsicht tätigen Akteure bislang ein unerforschtes, bzw. ein mit der Untersuchung von Böttcher, Keune und Neiwert (2010) forschungsaufkeimendes Feld darstellt, mussten überwiegend neue Items für die standardisierte Befragung entwickelt werden. Neben der Mehrzahl an eigenentwickelten Items konnten leicht modifizierte und an die Qualitätsanalyse in NRW angepasste Items aus den Untersuchungen von Gärtner und Wurster (2009), Böttcher, Keune und Neiwert (2010) sowie Ulber (2010) übernommen werden. Dabei berücksichtigen die beiden letztgenannten Untersuchungen die Schulaufsicht, während in der Befragung von Gärtner und Wurster (2009) Lehrkräfte und Schulleitungen adressiert sind. Im Folgenden werden die zentralen Merkmale des Fragebogens dargestellt:

*Einleitungsseite:* Die Befragung wird mit einem kurzen Einleitungstext eröffnet, der die Teilnehmer darauf hinweist, dass alle Angaben vertraulich be-

handelt und anonym ausgewertet werden. Zudem werden die Teilnehmer über den Zeitrahmen der Befragung (15 Minuten) und die Funktion des Fortschrittsbalkens im Fragebogen informiert. Der Hinweis zum Befragungsthema, zum Ziel der Befragung und der durchführenden Person und Institution wird vorab in der Einladungsmail zur Umfrage gegeben.

*Demografie:* Bezüglich der demografischen Angaben der Befragten werden Geschlecht, Alter, Tätigkeitsfeld und -ort erfragt. Entgegen der oft genannten Empfehlung, demografische Fragen vorzugsweise an das Ende eines Fragebogens zu stellen (vgl. Kuckartz u.a. 2009, S. 36; Gräf 2010, S. 79, S. 104; Diekmann 2009, S. 485; Jacob u.a. 2011, S. 183, S. 189), wurden die Einzelfragen zu den sozialstatistischen Merkmalen der Akteure im vorliegenden Fall am Beginn des Fragebogens platziert, um Abbrüche am Ende des Fragebogens zu vermeiden (vgl. Baur/Florian 2009, S. 125). Zudem wurden die demografischen Fragen nicht als Pflichtfragen, sondern mit der Möglichkeit zur Nichtbeantwortung programmiert, da ihre Antworterzwingung unter Umständen zu einer Ablehnung der Bearbeitung des Fragebogens geführt hätte.

*Filterführung:* Zur Identifikation derjenigen Schulaufsichtsbeamten innerhalb der Stichprobe, die noch keine Qualitätsberichte gelesen oder noch keine Zielvereinbarungen mit den Schulen geschlossen haben, wurde eine intelligente Filterführung programmiert. Somit wurden nur Befragte bei entsprechenden Bejahungen mit den Folgefragen konfrontiert.

*Schlussseite:* Auf der Schlussseite wird den Befragten kurz für ihre Teilnahme gedankt. Zudem ist eine Kontaktadresse hinterlegt, falls Fragen zur Erhebung oder zu den Ergebnissen bestehen. Die vorletzte Seite enthält freien Raum für Kommentare.

*Layout:* Um Seriosität, Relevanz und Nutzerfreundlichkeit des Fragebogens hervorzuheben, wurde für die Gestaltung der Fragebogenseiten ein schlichtes Design gewählt. Es wurden Logos gemäß der Corporate-Identity-Vorgaben der Westfälischen Wilhelms-Universität Münster und des Instituts für Erziehungswissenschaft in das Rahmendesign eingebunden. Zur Navigation wurden *Zurück-* und *Weiter-*Buttons definiert. Die Fragenblöcke wurden bildschirmweise aufgebaut, um Scrolling zu vermeiden und den wesentlichen Inhalt einer Fragebogenseite mit einem Blick erfassbar zu machen. Damit für die Befragten stets erkennbar sein sollte, an welcher Stelle im Fragebogen sie sich befinden,

wurde eine Fortschrittsanzeige als Balkendiagramm mit der relativen Fortschrittsangabe integriert.

*Skalierung:* Als Antwortformat der geschlossenen Einstellungsabfragen wird ein vierstufiges Antwortformat mit den Ausprägungen *1 = trifft gar nicht zu*, *2 = trifft eher nicht zu*, *3 = trifft eher zu* und *4 = trifft voll zu* verwendet, um die Tendenz zur Mitte bei einer ungeraden Anzahl an Skalenstufen zu vermeiden. Die Häufigkeit der gelesenen Berichte oder der getroffenen Zielvereinbarungen wird mittels einer fünfstufigen Skala (*1 = keine*, *2 = weniger als die Hälfte*, *3 = die Hälfte*, *4 = mehr als die Hälfte*, *5 = alle*) erfasst. Neben geschlossenen Fragen mit vorgegebenen Antwortkategorien werden offene und halboffene Fragen (Hybridfragen), bei welchen zusätzlich zu geschlossenen Antwortskalen offene Antwortmöglichkeiten vorgesehen sind, eingesetzt. Auf das Zurückgreifen der Befragten auf eine *weiß-nicht-* oder *keine Angabe-*Kategorie wird bewusst verzichtet, da sie tendenziell das Signal vermittelt, für die Befragten inhaltlich nicht interessant zu sein (vgl. Gräf 2010, S. 79).

*Pflichtfragen:* Der Fragebogen wurde so programmiert, dass die Befragten auf unvollständig ausgefüllte Fragen hingewiesen werden (Vollständigkeitscheck DAC bei Unipark). Dabei wird die Folgeseite nur dann angezeigt, wenn alle Fragen einer Seite beantwortet sind. Lediglich die Beantwortung der demografischen Fragen sowie die offenen Antwortmöglichkeiten konnten auf Wunsch unausgefüllt bleiben, um zur nächsten Fragebogenseite zu gelangen.

### **10.3.2. Pretest: Maßnahmen zur Qualitätssicherung des Fragebogens**

Aufgrund des Alleinstellungsmerkmals der Schulaufsicht in NRW und der geplanten Vollerhebung war die Durchführung eines Pretests mit der Zielgruppe nicht realisierbar. Eine Möglichkeit hätte in der Erprobung des Fragebogens bei Schulaufsichtsbeamten anderer Bundesländer bestanden. Von dieser Variante wurde allerdings Abstand genommen, da der Fragebogen sowohl sprachlich als auch prozesshaft auf das Verfahren der Qualitätsanalyse NRW zugeschnitten ist und für den Einsatz in anderen Bundesländern eine Anpassung an die dort vorherrschenden Verfahren bedingt hätte. Dadurch wäre im Pretest ein anderes Instrument im Umlauf gewesen als in der Feldphase. Dennoch muss eingeräumt werden, dass die Durchführung eines Pretests sicherlich Vorteile ge-

bracht hätte. Um den Fragebogen nicht ungeprüft einzusetzen, wurde dieser vor der Freischaltung der Umfrage sowohl inhaltlich als auch hinsichtlich seiner technischen Durchführbarkeit einer ausführlichen Pretestphase unterzogen. Es wurden einerseits Expertenurteile zum Fragebogen eingeholt, andererseits erfüllten zwei Fragebogenkonferenzen im Rahmen universitärer Kolloquien eine weitere Pretest-Funktion (vgl. Jacob u.a. 2011, S. 194). Die technische Überprüfung richtete sich auf die korrekte visuelle Darstellung, Filterführung und das Öffnen des Fragebogens mit verschiedenen gängigen Internetbrowsern. Hierfür war der Zugriff auf den Fragebogen von einer bei einer Bezirksregierung beschäftigten Testperson sinnvoll, um auch die Durchführbarkeit der Befragung an den dortigen Browsereinstellungen zu gewährleisten und um den Fragebogen unter möglichst realistischen Feldbedingungen zu prüfen.

Der inhaltliche Pretest wurde von insgesamt acht Personen mit sozialwissenschaftlichen Kenntnissen anhand des Pretest-Tools der Befragungssoftware *Unipark* vorgenommen. Er beinhaltete Aspekte der Fragenformulierung, des Fragebogaufbaus, der Skalierung von Antworten sowie der Verständlichkeit der Fragen. Zudem sollte sowohl auf redundante als auch auf fehlende Aspekte geachtet und die Bearbeitungszeit des Fragebogens ermittelt werden. Ein wichtiger Schritt, um auch eine gesteigerte Akzeptanz bei den Mitgliedern der Schulaufsicht zu erzielen, war der Einsatz zweier Experten aus der Schulaufsicht und einer ehemaligen Schulleitung, das Instrument auf seine Tauglichkeit hin zu bewerten. Durch diese qualitative Prüfung gelang es, die Fragen zu konkretisieren und somit das Instrument umfassend zu optimieren.

## **10.4. Durchführung der Befragung**

### **10.4.1. Rekrutierung**

Bei der Zielgruppe der Befragung handelt es sich um Angehörige einer spezifischen Berufsgruppe, den Arbeitnehmern im Beamtenverhältnis in der oberen und unteren Schulaufsicht<sup>49</sup> in NRW. In die Befragung wurden die schulfachli-

---

<sup>49</sup> Die Bezirksregierungen nehmen die Dienst- und Fachaufsicht über Realschulen, Gesamtschulen, Gymnasien und Berufskollegs wahr. Die Schulämter als untere Schulaufsichtsbehörden nehmen die Dienst- und Fachaufsicht über Grundschulen und die Fachaufsicht über Förder- und Hauptschulen wahr.

chen Mitarbeiter der Dezernate 41-45<sup>50</sup> einbezogen. Die Mitarbeiter der Dezernate 46-48<sup>51</sup> sowie juristische Mitarbeiter wurden nicht befragt, da der Abschluss von Zielvereinbarungen mit den Schulen im Anschluss an die Qualitätsanalyse nicht in ihren Aufgabenbereich fällt.

Um eine hohe Rücklaufquote zu gewährleisten, wurde anhand verschiedener Zugänge auf die geplante Befragung aufmerksam gemacht: Die Schulabteilungsleiter wurden über postalische Anschreiben mit der Beschreibung der geplanten Befragung und des Ablaufs um Zustimmung gebeten, die Befragung in ihrer Abteilung durchführen zu dürfen. Zudem war die geplante Befragung Inhalt einer landesweiten Schulabteilungsleiterkonferenz, so dass die Abteilungsleiter auch auf diesem Weg über das Befragungskonzept informiert wurden.

Das Einladungsschreiben für die Onlinebefragung wurde am 30.05.2011 elektronisch an die Leiter der fünf Schulabteilungen in den Bezirksregierungen verschickt. Darin wurden sie gebeten, die Einladungs-E-Mail an die schulfachlichen Dezenten in ihren Abteilungen weiterzuleiten. Das Einladungsschreiben beinhaltete neben dem Link, über den der Zugriff auf den Fragebogen erfolgte, eine kurze Vorstellung des Dissertationsthemas, den Zweck und die Dauer der Befragung. Zudem wurde eine zeitnahe Auswertung der Daten und eine Rückmeldung der Ergebnisse zugesichert. Vier Abteilungsleiter stimmten der Befragung zu, eine fünfte Bezirksregierung hat sich nicht an der Onlinebefragung beteiligt. Dadurch, dass sowohl der Versand des Einladungs- als auch des Erinnerungsschreibens an die Teilnehmer der Befragung direkt über die Abteilungsleiter erfolgte, konnte aus datenschutzrechtlichen Gründen kein direkter Kontakt zur Zielgruppe der Befragung aufgebaut und weder die Form noch der genaue Zeitpunkt der Weiterleitung der Einladungs- und Erinnerungsmails überprüft werden. Dennoch wurde dieses anonymisierte Vorgehen im Hinblick auf eine höhere Akzeptanz und Bereitschaft zur Teilnahme an der Befragung gewählt.

---

<sup>50</sup> Zuständigkeiten: Dezernat 41: Grund- und Förderschulen; Dezernat 42: Haupt-, Real- und Verbundschulen; Dezernat 43: Gymnasien und gymnasiale Oberstufe der Gesamtschulen; Dezernat 44: Gesamtschulen; Dezernat 45: Berufskollegs.

<sup>51</sup> Zuständigkeiten: Dezernat 46: Lehreraus- und -fortbildung; Dezernat 47: Personal- und Stellenplanangelegenheiten; Dezernat 48: Schulrecht und Schulverwaltung, Schulbau, Kirchensachen, Ersatzschulen, Sport, Sportstättenbau, Weiterbildung, Kunst und Kulturpflege, öffentliche Bibliotheken; Dezernat 49: Qualitätsanalyse an Schulen.

Die Feldphase der Onlinebefragung wurde zunächst für den Zeitraum vom 30.05.2011 bis zum 22.06.2011 angegeben. Da bis zum 20.06.2011 jedoch keine zufriedenstellende Rücklaufquote beobachtet werden konnte und sich in dieser Zeit mehrere Feier- sowie Brückentage bündelten, wurde am 20.06.2011 ein Erinnerungsschreiben verschickt und der Zeitraum zur Bearbeitung des Fragebogens bis zum 08.07.2011 verlängert. Das Erinnerungsschreiben hatte einen positiven Effekt auf die Rücklaufquote. Damit umfasst die Befragung insgesamt eine Feldzeit von 39 Tagen.

#### **10.4.2. Teilnahmeverhalten: Rücklauf und Ausschöpfungsquote**

Die durchschnittliche Gesamtbearbeitungszeit des Fragebogens durch die Teilnehmer lag bei 14 Minuten und 58 Sekunden (Median: 12 Minuten und 2 Sekunden). Somit wurde das für die Befragung angegebene Zeitfenster von 15 Minuten und eine generelle Empfehlung für die Dauer von Onlinebefragungen von 10 bis 15 Minuten eingehalten (vgl. Baur/Florian 2009, S. 124). Insgesamt wurde der Fragebogen 143 Mal aufgerufen und in 63 Fällen vollständig ausgefüllt. Dies entspricht einer Beendigungsquote von 44 %. Zusätzlich zu den 63 vollständig ausgefüllten Fragebögen lassen sich neun unvollständig bearbeitete Fragebögen mit geringen Item-Nonresponse-Ausfällen im letzten Sechstel des Fragebogens für die Datenauswertung nutzen. Im Folgenden bilden diese 72 Fragebögen die Grundlage der quantitativen Analysen. Werden die in die Auswertung eingegangenen Fragebögen zu den Zugriffen insgesamt in Beziehung gesetzt, so ergibt sich eine Abbruchquote von 50 % (72 von 143 Fällen). Dabei sind in 28 Fällen stichprobenrelevante Ausfälle beobachtet worden, in denen die erste Seite aufgerufen und abgebrochen wurde. Somit liegt die Nettobeteiligung an der Befragung bei 115 Fragebogenzugriffen. Innerhalb dieser Stichprobe wurden 52 (45,2 %) Fragebogenzugriffe abgebrochen und 63 (54,8 %) beendet. Legt man eine potentielle Stichprobengesamtheit von 177<sup>52</sup> der in Frage kommenden schulfachlichen Dezernenten in der oberen und unteren Schulaufsicht aus den vier teilnehmenden Bezirksregierungen zugrunde, so

---

<sup>52</sup> Diese Zahl beruht auf den der Autorin zur Verfügung gestellten Informationen durch die vier an der Umfrage beteiligten Bezirksregierungen. Es liegen keine Informationen zu Personen vor, die aus Krankheits- oder Urlaubsgründen nicht anwesend waren.

ergibt sich eine Ausschöpfungsquote (Rücklaufquote) von 41 %. Die erzielte Rücklaufquote entspricht durchaus den berichteten Referenzwerten im Rahmen von Onlinebefragungen (vgl. Couper/Coutts 2006, S. 223, 234; Maurer/Jandura 2009, S. 66; Gräf 2010, S. 60; Pötschke 2009, S. 83). Wird die Tatsache berücksichtigt, dass kein personalisierter Zugang mit entsprechender Möglichkeit einer persönlichen Kontaktierung der Teilnehmer gegeben war und die Feldphase zudem in eine bedingt durch mehrere Feiertage typische Urlaubszeit fiel, ist die Teilnahmequote als akzeptabel zu kennzeichnen.

Die Onlinebefragung der Schulaufsicht war als Vollerhebung konzipiert. Dieser Anspruch konnte zum einen schon durch die Teilnahmeverweigerung einer Bezirksregierung, zum anderen durch die erreichte Rücklaufquote von 41 % nicht eingelöst werden. Generell können auch unvollständig ausgeschöpfte Stichproben eine gute Qualität aufweisen. Hierbei ist die Analyse der Ausfallgründe und die Identifizierung ggf. systematischer Verzerrungen gefordert (vgl. Pötschke 2009, S. 84ff.). Zu fragen ist, ob das Teilnahmeverhalten mit den Merkmalen und Inhalten des Fragebogens zusammenhängt (vgl. Maurer/Jandura 2009, S. 64). Ein wichtiger Grund für die Ablehnung der Befragungsteilnahme könnte darin bestanden haben, dass sich diejenigen Schulaufsichtsbeamten, die weder einen Bericht gelesen haben, noch Zielvereinbarungen mit den Schulen getroffen haben – was angesichts der regionalgeleiteten Auswahl der Schulbesuche und der noch nicht flächendeckend erfolgten Qualitätsanalysen aller Schulen in NRW durchaus denkbar ist – gegen die Teilnahme an der Befragung entschieden haben könnten. Diese Vermutung wird durch Mitteilungen zweier Schulaufsichtsbeamten gestützt, in denen der Autorin via E-Mail mitgeteilt wurde, dass von der Beantwortung des Fragebogens aus o. g. Gründen Abstand genommen wird. Für die Verweigerung an der vorliegenden Befragung, vor allem derjenigen Personen, die den Link zur Befragung gar nicht oder lediglich die Einleitungsseite aufgerufen haben, geht eine weitere Erklärung in mehrere Richtungen: Zum einen könnte einem Teil der Zielgruppe das Thema nicht ausreichend interessant oder gar zu komplex erschienen sein, um dafür Zeit in die Beantwortung zu investieren. Zum anderen könnte auch fehlende Zeit oder Motivation die Beantwortung des Fragebogens verhindert haben. Die hohe Anzahl von Abbrüchen zu Beginn des Fragebogens, vor allem auf der Startseite ist vermutlich durch sogenannte Lurker zu erklären, d. h. Personen,

die sich den Fragebogen kurz anschauen wollten, um diesen ggf. zu einem späteren Zeitpunkt auszufüllen. Schließlich kommen Datenschutzbedenken als weiterer Grund für Teilnahmeverweigerungen in Betracht.

Betrachtet man die Abbruchquoten im Fragebogenverlauf, so lässt sich festhalten, dass die meisten Abbrüche auf den ersten vier von insgesamt neunzehn Seiten zu finden sind. Danach nimmt die Abbruchquote stark ab, steigt bei der Hälfte der Seiten leicht an und ab Seite 10 finden nur vereinzelte Abbrüche statt. Insgesamt sind bereits auf der Startseite (Seite 1) 25 % (absolut: 35) und zusammengenommen weitere 22 % (absolut: 32) der Abbrüche auf den Seiten 2 (Angaben zur Person) und 3 (Tätigkeitsfelder) zu beobachten. Jeweils 3 % der Abbrüche (absolut: 4) sind auf der Seite 4 (Auseinandersetzung mit dem Instrument und den Ergebnissen der Qualitätsanalyse) und auf der Seite 9 (Akzeptanz der Qualitätsanalyse) erfolgt. Die übrigen Abbrüche sind punktuell über einzelne Seiten verteilt. Im letzten Viertel des Fragebogens werden keine Abbrüche mehr beobachtet. Da auf der Einleitungsseite das Thema und die Ziele der Befragung kurz beschrieben sind, können Abbrüche dadurch erklärt werden, dass Personen die Studie für uninteressant hielten oder nicht bereit waren, die kalkulierten 15 Minuten zu investieren.

Neben dem Ausmaß an Item-Nonresponse gilt zusätzlich die Beantwortungsquote der offenen, nicht verpflichtend auszufüllenden Fragen als Merkmal der Teilnahmebereitschaft an der Befragung. Im Fragebogen kamen drei offene Fragen zum Einsatz, zwei Hybridfragen sowie ein Kommentarfeld am Ende des Fragebogens. Diese richteten sich an alle Befragten, sobald die Angabe zu den abgeschlossenen Zielvereinbarungen positiv beantwortet wurde (Filterfunktion). Die Beantwortungsbereitschaft lag bei den offenen Fragen mit 74 % bis 88 % in einem sehr hohen Bereich. Hinsichtlich der Hybridfragen wurde eine Beantwortungsbereitschaft zwischen 10 % und 18 % verzeichnet, was darauf hindeutet, dass durch die geschlossenen Fragen die zentralen Inhalte weitestgehend abgedeckt werden. Das offene Kommentarfeld wurde von einem Viertel der Befragten genutzt.

## 10.5. Stichprobe

Die Stichprobe der befragten Schulaufsichtsbeamten basiert auf Angaben von 72 Schulaufsichtsbeamten. Die Verteilung der Stichprobe auf die einzelnen Bezirksregierungen ist heterogen. Aus Gründen der zugesicherten Anonymität erfolgt keine namentliche Erwähnung der Bezirksregierungen. Aus zwei der vier Bezirksregierungen rekrutiert sich jeweils ein Drittel der Stichprobe, weitere 15 % und 20 % der Befragten sind an den anderen beiden Bezirksregierungen beschäftigt, während eine Bezirksregierung nicht an der Befragung teilgenommen hat. Darüber hinaus ist die Stichprobe dadurch gekennzeichnet, dass der überwiegende Teil (63 %) in der oberen Schulaufsicht in den Schulabteilungen der Bezirksregierungen tätig ist. 37 % der befragten Personen sind in der unteren Schulaufsicht, d. h. in den Schulämtern mit der Zuständigkeit für die Schulformen Grund-, Haupt- und Förderschule tätig. Die befragten Schulaufsichtsmitarbeiter sind mehrheitlich 56 Jahre und älter (74 %), nur ein geringer Teil ist zwischen 46 und 55 (19 %) und unter 46 Jahre (7 %) alt. Die Gruppe der unter 46-Jährigen fällt erwartungsgemäß klein aus und ist durch den berufsbiografisch vergleichsweise späten Eintritt in die Schulaufsicht zu erklären.

Das Geschlechterverhältnis ist in der Stichprobe paritätisch, allerdings lassen sich in Abhängigkeit der Funktionsstellung einige Unterschiede beobachten: Der Anteil der Beamtinnen innerhalb der oberen Schulaufsicht (42,2 %) liegt zu 13,4 % unter dem Anteil ihrer männlichen Kollegen (55,6 %). Entsprechend sind in der unteren Schulaufsicht 23 % mehr weibliche als männliche Beamte tätig. Dafür können schulformspezifische Begründungen angeführt werden, da insbesondere das Lehrerprofil an Grundschulen weiblich dominiert ist und die unteren Schulaufsichtsbehörden unter anderem die Dienst- und Fachaufsicht über Grundschulen wahrnehmen. Zudem liegt der Anteil jüngerer Mitarbeiter (45 Jahre und jünger) in der unteren Schulaufsicht zu 13,2 % über dem entsprechenden Anteil in der oberen Schulaufsicht. Hinsichtlich beider Merkmale der Stichprobe sind die Unterschiede allerdings nach Prüfung mittels Chi-Quadrat-Test nicht signifikant. Ob die Stichprobe mit den angegebenen Merkmalen als repräsentativ für die Grundgesamtheit der nordrhein-westfälischen Schulaufsicht gilt, kann nicht genau beantwortet werden, da keine amtlichen Daten über das soziodemografische Profil von Schulaufsichtsbeamten veröffentlicht sind.

## 10.6. Ergebnisse

### 10.6.1. Explorative Faktorenanalyse

Da der Fragebogen unter der Annahme konzipiert wurde, dass die verschiedenen Items Skalen komponieren, wurde eine explorative Faktorenanalyse durchgeführt (s. zum Verfahren Studie I). Da die einzelnen Items ausreichend hohe Kommunalitäten aufwiesen, wurde die EFA entgegen der Forderung nach einer umfangreichen Stichprobe mit dem vorliegenden Datensatz von  $N = 72$  Fällen berechnet (vgl. Bühner 2011, S. 344f.). Bezüglich der Konstruktvalidierung entfiel die CFA, da zur Berechnung der konfirmatorischen Faktorenanalyse eine Stichprobe von mindestens 100 bzw. 200 Fällen herangezogen werden sollte (vgl. Backhaus/Plinke/Weiber 2011). Die Verarbeitung fehlender Werte beruht auf paarweise verwendeten Fällen. In die EFA gingen 43 Variablen ein, von denen aber 17 aus weiteren Berechnungen ausgeschlossen wurden, da sie die geforderte Ladungshöhe von  $\geq ,50$  nicht erfüllten oder ihre eindeutige inhaltliche Zuordnung zu den Komponenten nicht möglich erschien. Insgesamt erfüllen 19 Variablen die statistischen Voraussetzungen einer eindeutigen Zuordnung zu den Komponenten. Sieben Variablen wurden trotz vorhandener Nebenladungen aus inhaltlichen Erwägungen beibehalten. Es wurde eine sechsfaktorielle Lösung ermittelt, die nach der varimax-rotierten Berechnung zusammen 72 % der Gesamtvarianz der Variablen erklärt.

### 10.6.2. Skalenbeschreibung und -ausprägung

Folgende sechs Skalen konnten bestimmt werden: *Rezeption\_quantitativ*, *Rezeption\_qualitativ*, *Reflexion*, *Akzeptanz der Qualitätsanalyse*, *wahrgenommene Professionalität der Qualitätsprüfer* sowie *negative Folgen der Qualitätsanalyse*. Insgesamt entsprechen die Skalen inhaltlich den Phasen und Rahmenbedingungen des postulierten Modells in Anlehnung an Helmke (2004). Es folgt eine inhaltliche Beschreibung der Skalen:

***Rezeption (quantitativ und qualitativ)***: Da jedem Schulaufsichtsbeamten mehrere Qualitätsberichte vorliegen, wird die Rezeption der Ergebnisse aus Qualitätsanalysen aus zwei Perspektiven bestimmt. Die Skala *Rezeption\_quantitativ* erfasst die Rezeption der Qualitätsberichte anhand von quantitativen

Einschätzungen. Erfragt wird auf der Grundlage von Merkmalen der Qualitätsberichte z. B. der Anteil prägnanter und verständlicher Qualitätsberichte. Beispiel-Item: *Wie viele der Ihnen vorliegenden Qualitätsberichte sind prägnant und verständlich?* Die Skala *Rezeption\_qualitativ* fokussiert die Intensität der Rezeption. Beispiel-Item: *Ich lese und bearbeite die Qualitätsberichte genau.*

**Reflexion:** Anhand dieser Skala wird die wahrgenommene Nützlichkeit der Ergebnisrückmeldungen und die Möglichkeiten der Ergebnisnutzung erfasst. Die Einschätzung erfolgt auf der Grundlage der Qualitätsberichte und Zielvereinbarungen hinsichtlich der strukturierten Steuerung schulischer Prozesse und der Impulsgebung zur Schulentwicklung. Beispiel-Item: *Die Qualitätsberichte sind geeignet, sachgerechte Empfehlungen für die Schulen abzuleiten.*

**Akzeptanz der Qualitätsanalyse:** Die Skala gibt Auskunft über die Akzeptanz der Qualitätsanalyse aus der Perspektive der Schulaufsicht. Thematisiert wird die Regelmäßigkeit, Notwendigkeit sowie der Nutzen von Qualitätsanalysen. Auf der Grundlage des Forschungsstands stellt eine positive Einstellung sowie Akzeptanz der Qualitätsanalyse eine wichtige Voraussetzung für eine gelungene Rezeption und Nutzung der Ergebnisse dar. Beispiel-Item: *Die Qualitätsanalyse ist notwendig, um die Qualität der Schulen insgesamt zu erhöhen.*

**Wahrgenommene Professionalität der Qualitätsprüfer:** Die Skala wurde leicht modifiziert und gekürzt aus der Untersuchung zur Wirkung von Schulinspektionen in Berlin von Gärtner und Wurster (2009, S. 74f.) adaptiert. Sie erhebt die Kompetenz und Diagnosegüte der agierenden Qualitätsprüfer, welche eine wichtige Voraussetzung für die Qualität der resultierenden Ergebnisse aus Qualitätsanalysen darstellen. Beispiel-Item: *Die Qualitätsprüfer/innen können die Unterrichtspraxis kompetent einschätzen.*

**Negative Folgen der Qualitätsanalyse:** Die für den vorliegenden Fragebogen adaptierten Items wurden zur Einschätzung negativer und erwünschter Folgen der Schulinspektion aus der Perspektive von Lehrkräften und Schulleitungen im Kontext der Untersuchung zur Wirkung von Schulinspektionen für das Bundesland Berlin entwickelt (vgl. Gärtner/Wurster 2009, S. 32f.;74). Die Skala erfasst die durch die Qualitätsanalyse ausgelösten negativen Begleiterscheinungen, bspw. Konflikte in den Schulen und zusätzliche Belastungen. Beispiel-Item: *Die Qualitätsanalyse hat zusätzliche Belastungen für die Schulen zur Folge.*

**Tab. 29: Itemformulierungen und statistisch relevante Skalen-Kennwerte der Befragung der Schulaufsicht (N = 72)**

Skalierung:			
a: Vierstufige Antwortskala: 1 = trifft gar nicht zu; 2 = trifft eher nicht zu; 3 = trifft eher zu; 4 = trifft voll zu			
b: Fünfstufige Antwortskala: 1 = kein Bericht; 2 = weniger als die Hälfte der Berichte; 3 = die Hälfte der Berichte; 4 = mehr als die Hälfte der Berichte; 5 = alle Berichte			
<b>Rezeption_quantitativ<sup>b</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Wie viele der Ihnen vorliegenden Qualitätsberichte sind so detailliert, dass die Schulen daraus gut Schwerpunkte und Maßnahmen für die Weiterentwicklung ableiten können?	4,3	,75	,68
Wie viele der Ihnen vorliegenden Qualitätsberichte enthalten hinreichend begründete Bewertungen?	4,0	,78	,64
Wie viele der Ihnen vorliegenden Qualitätsberichte helfen, Stärken der Schulen zu erkennen?	4,4	,75	,74
Wie viele der Ihnen vorliegenden Qualitätsberichte sind prägnant und verständlich?	4,3	,69	,66
Wie viele der Ihnen vorliegenden Qualitätsberichte helfen, Schwächen und Handlungsfelder der Schulen zu erkennen?	4,3	,90	,58
<b>Rezeption_qualitativ<sup>a</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Ich lese und bearbeite die Qualitätsberichte genau.	3,6	,71	,92
Ich setze mich mit den Befunden der Qualitätsanalysen intensiv auseinander.	3,7	,63	,92
<b>Reflexion<sup>a</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Qualitätsberichte haben für meine Arbeit in der Schulaufsicht einen großen Nutzen.	3,1	,73	,83
Die Qualitätsberichte stellen für die Schulen und mich eine gemeinsame Verständigungsgrundlage für Schulentwicklungsprozesse dar.	3,4	,60	,81
Die Qualitätsberichte haben für die Schulen einen großen Nutzen.	3,1	,69	,73
Die Qualitätsberichte sind geeignet, nachhaltige Impulse für die Schulentwicklung zu liefern.	3,2	,68	,71
Die Qualitätsberichte sind geeignet, sachgerechte Empfehlungen für die Schulen abzuleiten.	3,3	,63	,77
Durch Zielvereinbarungen im Anschluss an die Qualitätsanalyse gelingt es mir, die Ergebnisse aus Qualitätsanalysen für die Schulentwicklung zu nutzen.	3,2	,53	,69
Durch Zielvereinbarungen im Anschluss an die Qualitätsanalyse gelingt es mir, die Schulen darin zu unterstützen, ihre Ziele zu erreichen.	3,1	,52	,59
Durch Zielvereinbarungen im Anschluss an die Qualitätsanalyse gelingt es mir, schulische Prozesse strukturiert zu steuern.	2,9	,63	,63
<b>Akzeptanz der Qualitätsanalyse<sup>a</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Qualitätsanalyse ist notwendig, um die Qualität der Schulen insgesamt zu erhöhen.	2,9	,84	,77
Ich halte die Qualitätsanalyse für ein nutzbringendes Instrument der Schulentwicklung.	3,2	,77	,79
Qualitätsanalysen sollten regelmäßig durchgeführt werden.	3,3	,91	,79

<b>Wahrgenommene Professionalität der Qualitätsprüfer<sup>a</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Qualitätsprüfer/innen können die Qualität der schulischen Arbeit auf der Grundlage des Qualitätstableaus gut einschätzen.	3,2	,51	,56
Die Qualitätsprüfer/innen können die Leistungen der Schulen vor dem Hintergrund der Zusammensetzung der Schülerschaft zuverlässig einschätzen.	2,7	,64	,56
Die Qualitätsprüfer/innen können die Unterrichtspraxis kompetent einschätzen.	2,9	,75	,68
Die Qualitätsprüfer/innen können die Entwicklungsbedarfe der Schulen zuverlässig einschätzen.	3,1	,61	,67

<b>Negative Wahrnehmung der Qualitätsanalyse<sup>a</sup></b>	<b>MW</b>	<b>SD</b>	<b>rit</b>
Die Qualitätsanalyse hat zusätzliche Belastungen für die Schulen zur Folge.	3,1	,83	,55
Das Verhältnis von Aufwand und Ertrag der Qualitätsanalyse ist für die Schulen nicht angemessen.	2,4	,87	,69
Durch die Qualitätsanalyse treten zusätzliche Spannungen und Konflikte in den Schulen auf.	2,3	,71	,50
Die durch Qualitätsanalyse ausgelösten Impulse und Effekte verpuffen in den Schulen schnell.	2,2	,72	,44

In Tabelle 29 sind für die einzelnen Skalen die genauen Itemformulierungen und statistisch relevante Kennwerte der Items veranschaulicht. Alle sechs Skalen weisen mit hohen internen Konsistenzen von  $\alpha = ,75$  bis  $\alpha = ,95$  gute messtheoretische Eigenschaften auf (s. Tab. 30).

**Tab. 30: Übersicht über die Skalen der Befragung der Schulaufsicht**

<b>Skala</b>	<b>Anzahl Items</b>	<b>Cronbachs Alpha (<math>\alpha</math>)</b>	<b>Wertebereich der Skala</b>	<b>MW</b>	<b>SD</b>
Rezeption_quantitativ	5	,85	1 - 5	4,3	,61
Rezeption_qualitativ	2	,95	1 - 4	3,6	,65
Reflexion	8	,91	1 - 4	3,2	,50
Akzeptanz der QA	3	,89	1 - 4	3,1	,76
Wahrgenommene Professionalität der Qualitätsprüfer	4	,80	1 - 4	3,0	,50
Negative Folgen der QA	4	,75	1 - 4	2,5	,59

Vierstufige Antwortskala: 1 = trifft gar nicht zu; 2 = trifft eher nicht zu; 3 = trifft eher zu; 4 = trifft voll zu  
Fünfstufige Antwortskala: 1 = kein Bericht; 2 = weniger als die Hälfte der Berichte; 3 = die Hälfte der Berichte; 4 = mehr als die Hälfte der Berichte; 5 = alle Berichte;  
QA = Qualitätsanalyse; MW = Mittelwert; N = 72 Schulaufsichtsmitarbeiter

Alle Skalenmittelwerte liegen im positiven Bereich der Antwortskalen. Der Skalenmittelwert liegt bei der Skala *Rezeption\_quantitativ* (fünfstufige Skala) bei MW = 4,3 (SD = ,61), bei der Skala *Rezeption\_qualitativ* (vierstufige Skala)

la) bei  $MW = 3,6$  ( $SD = ,65$ ) und somit in beiden Fällen über der jeweiligen theoretischen Skalenmitte von 3 und 2,5. Damit erfüllen die Qualitätsberichte aus der Perspektive der Schulaufsicht durchaus die Voraussetzungen einer erfolgreichen Rezeption: Sie sind zum überwiegenden Anteil prägnant und verständlich verfasst sowie anschlussfähig im Hinblick auf das Erkennen von schulischen Schwächen und Stärken und die Überführung in spezifische Maßnahmen zur Weiterentwicklung der Schulen. Die Ergebnisse werden von den Befragten gründlich rezipiert und erfahren eine intensive Auseinandersetzung.

Für die Skala *Reflexion* ergibt sich ein durchschnittlicher Wert von 3,2, der inhaltlich so zu interpretieren ist, dass der Nutzen von Qualitätsberichten als durchaus hoch wahrgenommen wird, indem sie eine gemeinsame Verständigungsgrundlage für Schulen und die Schulaufsicht bilden sowie nachhaltige Impulse für die Schulentwicklung liefern. Auch die Zielvereinbarungen werden als durchaus nützlich für die Zielerreichung und Steuerung der schulischen Prozesse wahrgenommen. Die Skala *Akzeptanz* liegt ebenfalls im positiven Skalierungsbereich ( $MW = 3,1$ ), wird jedoch im Vergleich zur *Rezeption* und *Reflexion* geringfügig negativer bewertet. Die Meinungen der Schulaufsichtsbeamten streuen bzgl. der Akzeptanz vergleichsweise am stärksten ( $SD = ,76$ ). Ebenso wird auch die *Professionalität der Qualitätsprüfer* mit einem Mittelwert von 3,0 zwar durchaus positiv, jedoch kritischer als die Rezeption, Reflexion und Akzeptanz der Qualitätsanalyse bewertet. Die Skala *Negative Folgen der Qualitätsanalyse* liegt mit einem Mittelwert von 2,5 genau auf der theoretischen Skalenmitte und lässt sich somit nicht eindeutig dem positiven oder negativen Skalenbereich zuordnen. Somit ist zwar eine Grundakzeptanz der Qualitätsanalyse und der Qualitätsprüfer vorhanden, jedoch wird die Qualitätsanalyse von der Schulaufsicht differenziert wahrgenommen.

### **10.6.3. Zusammenhänge zwischen den Phasen und Rahmenbedingungen**

Nahezu alle Skalen korrelieren untereinander auf einem Niveau von  $p \leq ,01$  signifikant. Dabei weisen die signifikanten Korrelationen mittlere bis starke Ausprägungen auf (s. Tab. 31). Schulaufsichtsbeamte, welche den Qualitätsberichten hohe rezeptionsförderliche Eigenschaften bescheinigen, setzen sich auch intensiver mit den Ergebnissen auseinander und erleben eine höhere Nütz-

lichkeit der Ergebnismeldungen ( $r = ,59$ ;  $p = ,000$  mit der Skala *Rezeption\_quantitativ*;  $r = ,64$ ;  $p = ,000$  mit der Skala *Rezeption\_qualitativ*). Dabei wird deutlich, dass die beiden verschiedenen Facetten (quantitativ und qualitativ) der Skala *Rezeption* nur auf einem mittleren Niveau miteinander korrelieren ( $r = ,41$ ;  $p = ,001$ ). Die qualitative Facette der *Rezeption* hängt geringfügig stärker mit der *Reflexion* zusammen ( $r = ,64$ ). Je mehr die Ergebnisse rezipiert und reflektiert werden, desto weniger werden negative Folgen der Qualitätsanalyse wahrgenommen ( $r = -,32$  und  $r = -,35$ ). Offensichtlich tragen auch die Rahmenbedingungen deutlich zum Ausmaß der Rezeption und Reflexion bei: Je mehr die Qualitätsanalyse akzeptiert wird und je ausgeprägter die Professionalität der Qualitätsprüfer wahrgenommen wird, desto positiver werden auch die Ergebnisse der Qualitätsanalyse rezipiert und reflektiert. Zwischen den *negativen Folgen der Qualitätsanalyse* und der *Akzeptanz* sowie *Professionalität der Qualitätsprüfer* ist ein negativer, allerdings schwacher und nicht signifikanter Zusammenhang zu verzeichnen.

**Tab. 31: Skalen-Interkorrelationen der Befragung der Schulaufsicht**

	(2)	(3)	(4)	(5)	(6)
Rezeption_quantitativ (1)	,41**	,59**	,39**	,53**	-,32**
Rezeption_qualitativ (2)		,64**	,38**	,32**	-,30*
Reflexion (3)			,65**	,52**	-,35**
Akzeptanz (4)				,62**	-,02
Professionalität Qualitätsprüfer (5)					-,21
Negative Folgen QA (6)					

\*\* Die Korrelation ist auf dem Niveau von ,01 (2-seitig) signifikant.

\* Die Korrelation ist auf dem Niveau von ,05 (2-seitig) signifikant. N = 72 Schulaufsichtsmitarbeiter

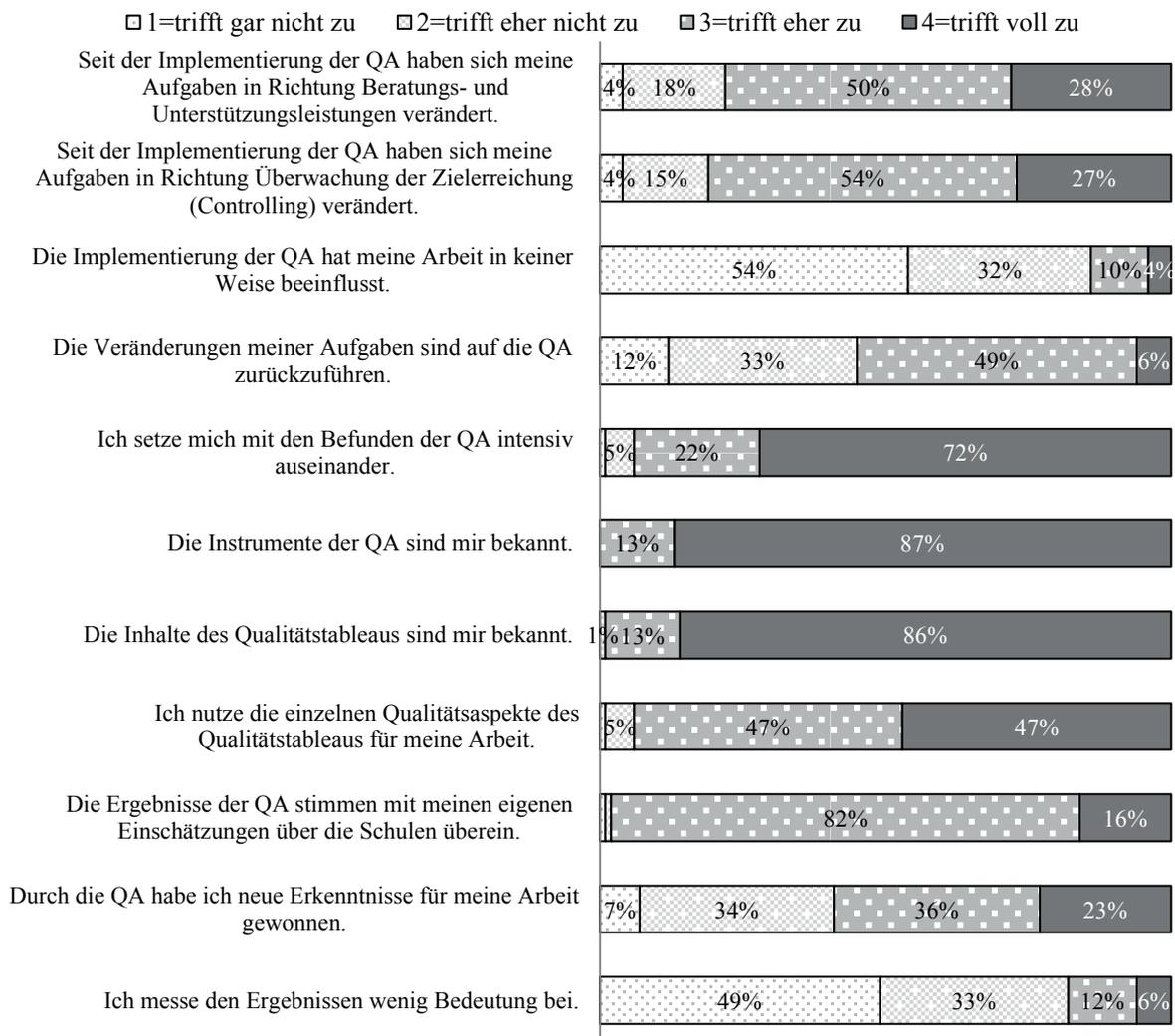
Als wesentliche Befunde der schrittweisen Regressionsanalyse lässt sich für die *Rezeption\_quantitativ* ein nur geringes Varianzaufklärungspotential ( $< R^2 = ,30$ ) der Rahmenbedingungen festhalten. Hingegen erweisen sich die Rahmenbedingungen *Akzeptanz der Qualitätsanalyse* und *Rezeption\_quantitativ* bzgl. ihres Erklärungspotentials der Phase *Reflexion* als einflussreich. Mit diesen Variablen können 52 % der Gesamtvarianz der *Reflexion* erklärt werden, was einem sehr hohen Varianzaufklärungspotential entspricht. Die beiden Phasen *Rezeption* und *Reflexion* haben allerdings ein nur geringes Varianzaufklärungspotential der Phase *Folgen der Qualitätsanalyse* von  $R^2 = ,17$ .

#### **10.6.4. Auseinandersetzung der Schulaufsicht mit der Qualitätsanalyse NRW und ihren Ergebnissen**

Die Unterstützung und Beratung der Schulen im Anschluss an die Qualitätsanalyse sowie die Überprüfung der umgesetzten Maßnahmen setzt seitens der Schulaufsichtsbeamten eine intensive Auseinandersetzung mit den einzelnen Instrumenten und den Ergebnissen der Qualitätsanalyse voraus. Inwiefern dies erfolgt, wurde im Rahmen der Onlinebefragung der Schulaufsicht erfasst und wird im Folgenden anhand mehrerer Indikatoren erörtert. Dabei dienen zum einen quantitative und „objektive“ Größen wie die Anzahl gelesener Qualitätsberichte, Anzahl der Zielvereinbarungen und Hospitationen als Indikatoren der schulaufsichtlichen Auseinandersetzung mit der Qualitätsanalyse und ihren Ergebnissen. Zum anderen wird dies über Fragen zu Veränderungen der Tätigkeitsfelder sowie dem Ausmaß der wahrgenommenen Auseinandersetzung und Nutzung der Ergebnisse direkt erhoben.

Den Einfluss der Qualitätsanalyse auf ihre Arbeit nehmen nahezu alle befragten Schulaufsichtsbeamten wahr (86 %). Die Frage, ob sich die Aufgaben der Schulaufsicht seit der Implementierung der Qualitätsanalyse NRW eher in Richtung *Beratungs- und Unterstützungsleistungen* oder *Controlling* verändert hätten, liefert keinen eindeutigen Befund: Etwa drei Viertel der befragten Schulaufsichtsmitarbeiter bejahen eine Veränderung beider Aufgabenfelder (Addition der Antwortoptionen *trifft voll zu* und *trifft eher zu*) und etwa die Hälfte (55 %) führt die Veränderungen der Aufgaben auf die Qualitätsanalyse zurück (s. Abb. 23). Die befragte Stichprobe verfügt über eine durchaus gute Kenntnis der Instrumente und hat mehrheitlich praktische Erfahrung mit dem Verfahren der Qualitätsanalyse sammeln können: Etwa zwei Drittel (69 %) der befragten Schulaufsichtsbeamten haben bereits an einer Qualitätsanalyse hospitiert. Darüber hinaus sind allen befragten Schulaufsichtsbeamten die Instrumente der Qualitätsanalyse sowie die Inhalte des Qualitätstableaus bekannt (bezüglich des Tableaus hat nur eine Person verneinend mit „trifft eher nicht zu“ geantwortet). Die genaue Kenntnis der Instrumente und Inhalte der Qualitätsanalyse zeigt sich auch daran, dass 94 % der befragten Schulaufsichtsmitarbeiter die einzelnen Qualitätsaspekte für ihre Arbeit nutzen.

Eine Übereinstimmung der Ergebnisse der Qualitätsanalyse mit den eigenen Einschätzungen wird von der überwiegenden Mehrheit der Befragten als zutreffend wahrgenommen (15 % trifft voll zu, 81 % trifft eher zu;  $MW = 3,1$ ;  $SD = ,47$ ). Dass die Qualitätsanalyse jedoch nicht nur die schulaufsichtlichen Erfahrungswerte über die Schulen repliziert, zeigt sich an der Beantwortung des Items „Durch die Qualitätsanalyse habe ich neue Erkenntnisse für meine Arbeit gewonnen.“: Diese Aussage halten rund 60 % der Befragten für zutreffend ( $MW = 2,8$ ;  $SD = ,90$ ). Die deutliche Mehrheit der Befragten (82 %) misst den Ergebnissen der Qualitätsanalyse eine hohe Bedeutung bei.



**Abb. 23: Häufigkeitsverteilung der Items zur schulaufsichtlichen Auseinandersetzung mit der Qualitätsanalyse und ihren Ergebnissen (N = 72)**

Die Befragungsergebnisse zeigen, dass eine faktische Auseinandersetzung mit den schulischen Ergebnissen im Anschluss an die Qualitätsanalyse durch die Schulaufsicht in starkem Ausmaß stattgefunden hat: Jeweils ein Drittel der Stichprobe hat bereits einen bis fünf (31 %) oder sechs bis zehn (34 %) Qualitätsberichte gelesen; jeweils sieben Prozent der Befragten geben an, 16 bis 20 und mehr als 20 Qualitätsberichte gelesen zu haben. Während insgesamt 20 % der Befragten angaben, 11 bis 15 Berichte gelesen zu haben, hat lediglich eine Person noch keinen Bericht gelesen. Ein weiterer Indikator für die Auseinandersetzung mit den Ergebnissen der Qualitätsanalyse liegt im Umfang abgeschlossener Zielvereinbarungen. Nahezu alle Befragten haben bereits mit mehreren Schulen in ihrer Zuständigkeit Zielvereinbarungen im Anschluss an die Qualitätsanalyse geschlossen. Der größte Teil (42 %) hat entsprechende Vereinbarungen mit weniger als der Hälfte der Schulen, knapp ein Fünftel (19 %) bereits mit allen Schulen und zwei Personen bislang mit keiner Schule geschlossen. 13 % der Schulaufsichtsmitarbeiter gaben an, mit der Hälfte der Schulen zum Befragungszeitpunkt Zielvereinbarungen geschlossen zu haben.

Ob sich die schulaufsichtliche Wahrnehmung der verschiedenen Phasen (*Rezeption, Reflexion, Negative Folgen der Qualitätsanalyse*) und der Rahmenbedingungen (*Akzeptanz, Professionalität der Qualitätsprüfer*) in Abhängigkeit der Anzahl gelesener Qualitätsberichte und abgeschlossener Zielvereinbarungen unterscheidet, wurde mittels Mittelwertvergleiche (T-Test und Einfaktorische ANOVA) geprüft (s. Tab. 32). Für diese Berechnung wurden aus Gründen einer besseren Vergleichbarkeit die gelesenen Berichte und abgeschlossenen Zielvereinbarungen gruppiert. Die Mittelwerte der Skalen unterscheiden sich in Abhängigkeit der Hospitation, der Anzahl gelesener Qualitätsberichte und der Anzahl abgeschlossener Zielvereinbarungen nur marginal. Die Mittelwertunterschiede sind allesamt statistisch nicht signifikant, was beweist, dass auf der Grundlage der vorliegenden Stichprobe die Intensität der Auseinandersetzung mit den Ergebnissen der Qualitätsanalyse und ihre Kenntnis keine Auswirkung auf die Wahrnehmung der Phasen und Rahmenbedingungen des theoretischen Modells hat.

**Tab. 32: Mittelwertunterschiede der Skalen in Abhängigkeit der Intensität der Auseinandersetzung mit der Qualitätsanalyse und ihren Ergebnissen**

Skala	Hospitation		Anzahl gelesener Qualitätsberichte			Anzahl abgeschlossener Zielvereinbarungen	
	MW (SD)		MW (SD)			MW (SD)	
	ja (N = 44)	nein (N = 20)	< 5 (N = 23)	6 - 10 (N = 24)	> 11 (N = 24)	< Hälfte (N = 29)	> Hälfte (N = 27)
<b>Rezeption_ quant</b>	4,3 (,63)	4,3 (,61)	4,4 (,53)	4,1 (,74)	4,3 (,53)	4,2 (,53)	4,3 (,58)
<b>Rezeption_qual</b>	3,7 (,55)	3,6 (,67)	3,4 (,85)	3,7 (,55)	3,8 (,46)	3,7 (,55)	3,7 (,59)
<b>Reflexion</b>	3,1 (,53)	3,2 (,42)	3,2 (,50)	3,0 (,49)	3,3 (,48)	3,1 (,46)	3,3 (,42)
<b>Akzeptanz</b>	3,2 (,65)	3,1 (,70)	3,2 (,65)	3,1 (,72)	3,1 (,90)	3,2 (,61)	3,2 (,61)
<b>Professionalität Qualitätsprüfer</b>	3,0 (,50)	3,0 (,53)	3,1 (,57)	2,8 (,44)	3,1 (,47)	3,0 (,45)	2,9 (,50)
<b>Negative Folgen der QA</b>	2,5 (,52)	2,6 (,48)	2,5 (,45)	2,5 (,42)	2,5 (,82)	2,5 (,46)	2,6 (,52)

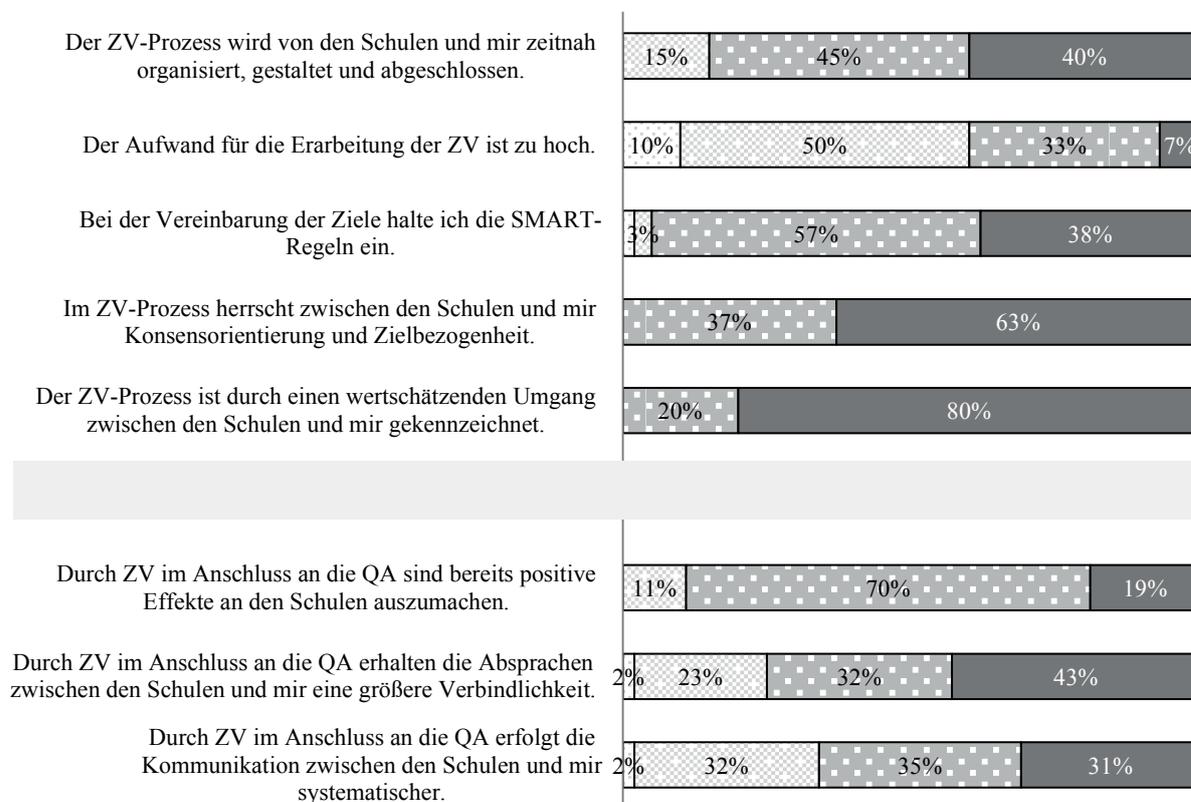
Im Einzelnen zeigt sich, dass die Skalenwerte der *Rezeption\_quantitativ* und *Reflexion* im Gegensatz zur *Rezeption\_qualitativ* nicht kontinuierlich mit der Anzahl gelesener Berichte steigen. Eine höhere Ausprägung der Phasen *Rezeption*, *Reflexion* und *Folgen der Qualitätsanalyse* zeigt sich in Abhängigkeit abgeschlossener Zielvereinbarungen: Bei abgeschlossenen Zielvereinbarungen mit mehr als der Hälfte der Schulen werden die Phasen geringfügig positiver eingeschätzt, als wenn ein Schulaufsichtsmitarbeiter erst mit weniger als der Hälfte der Schulen Zielvereinbarungen abgeschlossen hat. Bei den Rahmenbedingungen lassen sich bzgl. dieses Indikators allerdings keine nennenswerten Unterschiede beobachten. Ein weiteres Ergebnis zeigt sich in der Kontinuität der Skalen *Negative Folgen* und *Akzeptanz der Qualitätsanalyse*, die sich unabhängig von der Hospitation, der Anzahl gelesener Berichte und abgeschlossener Zielvereinbarungen zwischen identischen mittleren Werten bewegen.

### 10.6.5. Zielvereinbarungsprozess

Im Rahmen der Schulinspektion nimmt der Zielvereinbarungsprozess eine zentrale Brückenfunktion zwischen den Ergebnismeldungen und der -nutzung ein. Wie sich der Zielvereinbarungsprozess gestaltet, wird aus der Perspektive der befragten Schulaufsichtsbeamten (N = 72) wie folgt beantwortet (s. Abb. 24):

N = 72; ZV = Zielvereinbarungen; QA = Qualitätsanalyse

□ 1=trifft gar nicht zu □ 2=trifft eher nicht zu □ 3=trifft eher zu □ 4=trifft voll zu



**Abb. 24: Häufigkeitsverteilung der Items zum Zielvereinbarungsprozess aus der Perspektive der Schulaufsicht**

Die Schulaufsichtsbeamten schätzen den Zielvereinbarungsprozess mit den Schulen als paritätische Zusammenarbeit ein. Die inhaltliche Zielbestimmung erfolgt ihrer Ansicht nach überwiegend (zu 61 %) gleichberechtigt von schulischer und schulaufsichtlicher Seite. Etwa ein Drittel (38 %) der befragten Schulaufsichtsbeamten gibt an, dass die inhaltliche Zielbestimmung in erster Linie durch die Schulen gelenkt wird. Der Zielvereinbarungsprozess wird von der Mehrzahl der Schulaufsichtsbeamten zeitnah organisiert, gestaltet und abgeschlossen. Der Aufwand für die Erarbeitung der Zielvereinbarungen wird überwiegend als moderat angesehen, jedoch empfindet ein nicht unerheblicher Anteil der Befragten (40 %) den Aufwand als zu hoch. Nahezu die gesamte Stichprobe gibt an, sich bei der Vereinbarung der Ziele an die SMART-Regel zu halten (38 % trifft voll zu; 57 % trifft eher zu). Das Klima im Zielvereinba-

rungsprozess wird ausschließlich positiv wahrgenommen und ist laut Selbstauskünften der Schulaufsicht durch Konsensorientierung, Zielbezogenheit und einen wertschätzenden Umgang zwischen den Beteiligten gekennzeichnet.

Ausgehend von diesen überwiegend positiven Merkmalen des Zielvereinbarungsprozesses kann auf entsprechend günstige Bedingungen bei der Erarbeitung der Dokumente und auf eine hohe Qualität dieser Dokumente durch die Anwendung der SMART-Regel geschlossen werden.

### 10.6.6. Zielinhalte

Während die inhaltlichen Kategorien der Zielvereinbarungen in Studie II auf der Grundlage der Dokumente inhaltsanalytisch bestimmt wurden und die gemeinsame Perspektive der Schulen und der Schulaufsicht widerspiegeln, werden in Studie III die Zielinhalte über eine Selbstauskunft der Schulaufsichtsbeamten erfragt. Die Zielinhalte wurden durch offene Fragen erhoben und waren nicht verpflichtend zu beantworten. Die Angaben wurden inhaltsanalytisch ausgewertet, indem sie den jeweiligen Qualitätsaspekten des Qualitätstableaus zugeordnet wurden. Nicht immer war eine direkte Zuordnung der genannten Zielinhalte zu den Qualitätsaspekten des Qualitätstableaus möglich, da in einigen Fällen ein globales Ziel, z. B. *Unterricht* angegeben wurde. In solchen Fällen wurde auf die kategoriale Differenzierung der Qualitätsbereiche (QB 1 bis QB 6) zurückgegriffen. In anderen Fällen konnte keine Zuordnung zu den Qualitätsaspekten vorgenommen werden, da der Zielinhalt nicht eindeutig auf das Qualitätstableau zurückzuführen war, z. B. bei dem Zielinhalt *Transparenz*.

Insgesamt wurden von den befragten Schulaufsichtsmitarbeitern 159 Zielinhalte genannt: Dabei machten 58 Personen Angaben zum ersten Ziel, 54 Personen zum zweiten Ziel und 47 Personen zum dritten Ziel, so dass von durchschnittlich 53 Fällen für die weiteren Berechnungen auszugehen ist. Tabelle 33 gibt einen Überblick über die genannten Zielinhalte absteigend sortiert nach der Häufigkeit der Nennungen. Insgesamt konnten 22 unterschiedliche Zielinhalte identifiziert werden, von denen allerdings die Hälfte lediglich jeweils einmal genannt wurde. Dabei zeigt sich ein dominanter Anteil des *QB 2 Lernen und Lehren – Unterricht* von 76 %. Der QB 2 lässt sich auf der Grundlage der Qualitätsaspekte differenziert betrachten: *2.6 Individuelle Förderung und Unterstützung* (18 % auf der Basis der Nennungen; 55 % auf der Basis der Fälle), *2.2*

*Leistungskonzept – Leistungsanforderung und Leistungsbewertung* (17 % auf der Basis der Nennungen; 51 % auf der Basis der Fälle) sowie *2.4 Unterricht – Unterstützung eines aktiven Lernprozesses* (16 % auf der Basis der Nennungen; 49 % auf der Basis der Fälle). Weitere 16 % der Zielinhalte ließen sich keinem Qualitätsaspekt zuordnen, beziehen sich jedoch auf den Qualitätsbereich 2 *Lernen und Lehren – Unterricht*. Auf die Frage, wie viele Ziele in einem Zusammenhang mit den Ergebnissen der Qualitätsberichte stehen, antwortet etwa die Hälfte der befragten Personen (54 %), dass sie diesen Zusammenhang bei allen Zielen sieht. Weitere 36 % geben an, dass mehr als die Hälfte der vereinbarten Ziele in einem Zusammenhang mit den Ergebnissen der Qualitätsberichte steht.

**Tab. 33: Auswertung der offenen Angaben zu den Zielinhalten aus der Perspektive der Schulaufsicht (N = 159 Nennungen; N = 53 Fälle)**

Zielinhalte (Angaben bzgl. einer offenen Frage)	Zuordnung Qualitäts- aspekt	Absolute Häufig- keit Ziel- inhalt	Prozentuale Häufigkeit (Basis Nennun- gen; N = 159)	Prozentuale Häufigkeit (Basis Fälle; N = 53)
Individuelle Förderung	2.6	29	18 %	55 %
Leistungsbewertung	2.2	27	17 %	51 %
Schüleraktivierung	2.4	26	16 %	49 %
Unterrichtsentwicklung/ Unterricht	QB 2	26	16 %	49 %
Schulinterne Curricula	2.1	13	8 %	25 %
(Schulinterne) Evaluation	6.2	11	7 %	21 %
Transparenz	-	4	3 %	8 %
(Schulinterne) Kommunikation	5.3	4	3 %	8 %
Schulorganisation	QB 4	3	2 %	6 %
Fortbildung/Weiterentwicklung beruflicher Kompetenzen	5.2	3	2 %	6 %
Didaktische Jahresplanung	6.3	2	1 %	4 %
Lehrerkooperation	5.3	1	< 1 %	2 %
Partizipation	3.4	1	< 1 %	2 %
Projektgruppen Schulent- wicklung	6.1	1	< 1 %	2 %
Beratungskonzept	-	1	< 1 %	2 %
Methodenkompetenz	1.4	1	< 1 %	2 %
Feedback-Kultur	-	1	< 1 %	2 %
Vertretungskonzept	4.2	1	< 1 %	2 %
Teamentwicklung, -bildung	5.3	1	< 1 %	2 %
Elternarbeit	3.4	1	< 1 %	2 %
Qualitätsmanagement	-	1	< 1 %	2 %
Deutsch als Zweitsprache	-	1	< 1 %	2 %
<b>Gesamt <math>\Sigma</math></b>		<b>159</b>	<b>100 %</b>	<b>304 %</b>

### 10.6.7. Beratungs- und Unterstützungsleistungen

Die Ergebnisse der Qualitätsanalysen gewähren dem Großteil der befragten Schulaufsichtsbeamten Entscheidungsgrundlagen über die Unterstützungsbedarfe von Schulen (in 92 % der Fälle zutreffend). Dabei sind Beratungen im Bereich *Schulentwicklung* die am häufigsten bereitgestellten Unterstützungsleistungen, gefolgt von Beratungen bei alltäglichen schulischen Fragen und Problemen (s. Tab. 34). Vergleichsweise seltener werden dagegen als Unterstützungsleistungen im Anschluss an die Qualitätsanalyse die Vermittlung und Auswahl von Fortbildungen bereitgestellt. Die Aussage, den Schulen überhaupt keine Unterstützungsleistungen angeboten zu haben, wird in 88 % der Fälle verneint. Inwiefern diese Beratungs- und Unterstützungsleistungen im direkten Zusammenhang mit der Qualitätsanalyse stehen und sich von den herkömmlichen Unterstützungsleistungen der Schulaufsicht unterscheiden, kann nicht beantwortet werden, da hierzu keine Angaben aus der eigenen Befragung und der Forschungsliteratur vorliegen.

**Tab. 34: Bereitgestellte Unterstützungsleistungen durch die Schulaufsicht**

<b>Unterstützungsleistungen</b>	<b>Mittelwert</b>	<b>Standardabweichung</b>
Beratung zur Schulentwicklung	4,2	,96
Beratung bei alltäglichen Fragen und Problemen	3,6	1,4
Auswahl von Fortbildungen	2,9	1,2
Vermittlung von Fortbildungen	2,7	1,2
Keine Unterstützungsleistungen	1,2	,72

N = 72; Skalierung: 1 = keiner Schule; 2 = weniger als der Hälfte der Schulen; 3 = der Hälfte der Schulen; 4 = mehr als der Hälfte der Schulen; 5 = allen Schulen

Eine wichtige Komponente der schulaufsichtlichen Beratung und Unterstützung der Schulen stellen zusätzliche Ressourcen dar. Wie allerdings die Befragungsergebnisse belegen, werden den Schulen von 92 % der befragten Schulaufsichtsbeamten keine zusätzlichen Ressourcen im Anschluss an die Qualitätsanalyse zur Verfügung gestellt. Konträr zu dieser realen Position werden zur optimalen Unterstützung der Schulen aus der Schulaufsichtsperspektive jedoch mehr Ressourcen benötigt (in 88 % der Fälle zutreffend).

Die Einhaltung der Zielvereinbarungen wird von den befragten schulaufsichtlichen Mitarbeitern nahezu in allen Fällen (in 96 % der Fälle zutreffend) überprüft. Die Überprüfung der Zielvereinbarungseinhaltung erfolgt insbesondere durch fest vereinbarte Termine mit der Schulleitung (in 98 % der Fälle zutreffend) sowie durch Schulbesuche (in 80 % der Fälle zutreffend). Zudem wurden folgende offenen Angaben zu weiteren Formen der Überprüfung gemacht: Teilnahme an Lehrerkonferenzen und sonstige Dienstgeschäfte; Einsicht in Unterlagen, z. B. in Dokumentationen, Protokolle (Arbeits- und Beschlussprotokolle), Arbeitsergebnisse, Konzepte etc. (3 Nennungen); regelmäßige Thematisierung im Rahmen der regionalen Schulleiterdienstbesprechung.

### **10.6.8. Zusammenfassung der offenen Antworten im Kommentarfeld**

Ein Viertel der Befragten nutzte das offene Kommentarfeld, um durch den Fragebogen nicht erfasste Inhalte mitzuteilen. Die Kommentare stellen eine reflektierte und kritische Sicht der Schulaufsicht auf die Qualitätsanalyse und ihre Anschlussfähigkeit heraus und fungieren damit als eine wichtige Ergänzung der standardisierten Angaben im Fragebogen.

Die Qualitätsanalyse wird von den befragten Schulaufsichtsbeamten überwiegend als sehr hilfreich und die Zusammenarbeit mit dem Dezernat 4 Q als regelmäßig und gut bezeichnet. Allerdings könnte sie noch ausgebaut werden. Dabei wird der Wunsch genannt, als Gast in der Zuhörerrolle an Portfoliokonferenzen teilnehmen zu können. Dadurch würde eine viel breitere Basis für eine optimierte Zusammenarbeit mit den Schulen, vor allem im Hinblick auf Zielvereinbarungen und abgestimmte Unterstützungsangebote im weiteren Entwicklungsprozess entstehen. In mehreren Kommentaren werden die unzureichenden und wenig konkreten Unterstützungsangebote nach Abschluss der Qualitätsanalyse genannt. Von Seiten der Kompetenzteams müssten mehr und sinnvolle Angebote gemacht werden. Darüber hinaus werden größere zeitliche und personelle Ressourcen sowie eine nachhaltige, wirksame Begleitung durch die Schulaufsicht im Nachgang der Qualitätsanalyse gefordert.

In weiteren Kommentaren werden die über alle Schulformen hinweg identischen Ergebnisse der Qualitätsanalyse sowie der Entwicklungsbedarfe der Schulen thematisiert. Dadurch werde der Aufwand der Qualitätsanalyse für

nicht mehr gerechtfertigt eingeschätzt. Zudem würden die Ergebnisse der Qualitätsanalyse die eigene Kenntnis bestätigen, die die Schulaufsicht über die Schulen und ihre Entwicklungsbedarfe habe. Mit entsprechenden Unterstützungsangeboten könne Schulentwicklung ohne weitere Qualitätsanalysen organisiert werden. Das folgende Zitat fasst diese Position sehr treffend zusammen: „*Entscheidend ist also nicht die Fortführung der Qualitätsanalyse, sondern die Bereitstellung der Ressourcen für die nachhaltige Unterrichtsentwicklung, z. B. anhand von Best-Practice-Beispielen (...)*“. Auf der anderen Seite werden kurze Analyseintervalle gefordert, da die Qualitätsanalyse sonst von den Schulen im Sinne eines Entwicklungsauftrags nicht mehr ernstgenommen würde.

## 10.7. Resümee

Im Folgenden werden die Befunde der Studie III zusammengefasst und in einer kritischen Reflexion verortet. Zusammenfassend zeigt sich das Bild einer Schulaufsichts-Stichprobe, welche größtenteils in den viereinhalb Jahren seit der Implementierung der Qualitätsanalyse durchaus das Verfahren praxisnah durch Hospitationen kennengelernt hat, bereits mehrere Qualitätsberichte (im Durchschnitt zwischen sechs und zehn Berichte) gelesen hat und mit durchschnittlich der Hälfte der Schulen Zielvereinbarungen geschlossen hat. Die Ergebnisse der Qualitätsanalyse bestätigen weitestgehend bereits vorhandene Erfahrungen und Informationen der Schulaufsicht über die Schulen. Die über Selbstauskunft der Schulaufsichtsbeamten erfragten Zielinhalte decken sich mit den in Studie II inhaltsanalytisch bestimmten Zielinhalten in Anlehnung an das Qualitätstableau. Der dominante Anteil der Aspekte des *Qualitätsbereichs 2 Lernen und Lehren – Unterricht* wird in Studie III bestätigt (insgesamt 75 % der Nennungen).

Mit den erzielten Erkenntnissen ist es gelungen, das Rahmenmodell von Helmke (2004) zu erweitern und aus der Perspektive der Schulaufsicht darzustellen. Auf der Grundlage einer explorativen Faktorenanalyse konnten die Phasen *Rezeption*, *Reflexion* und *Folgen der Inspektion* sowie die beiden Rahmenbedingungen *Akzeptanz der Qualitätsanalyse* und *Wahrgenommene Professionalität der Qualitätsprüfer* des postulierten Modells identifiziert werden. Dabei wird die Phase *Rezeption* differenziert anhand der Facetten *quantitativ*

und *qualitativ* betrachtet. Aus methodischer Sicht ist die Skala *Rezeption\_ qualitativ* kritisch zu bewerten, da sie aus lediglich zwei Items zusammengesetzt ist. Sie wurde dennoch beibehalten, da sie wichtige Inhalte enthält und eine Ergänzung zur quantitativ orientierten Skala darstellt. Insgesamt wurden die Phasen *Rezeption* und *Reflexion* von den Schulaufsichtsbeamten überaus positiv eingeschätzt, während die Folgen der Qualitätsanalyse und die Rahmenbedingungen dieses Prozesses kritischer bewertet wurden. Die Rahmenbedingungen stehen im positiven, mittleren bis hohen und signifikanten Zusammenhang zu den beiden Phasen *Rezeption* und *Reflexion* und stellen damit relevante Bedingungen für die Nutzung der Ergebnisse aus Qualitätsanalysen dar. Der Umfang bereits gelesener Qualitätsberichte und abgeschlossener Zielvereinbarungen zeigt keine signifikante Auswirkung auf die *Rezeption* und *Reflexion* der Ergebnisse aus der Perspektive der Schulaufsicht.

Einige Items bzw. Skalen wurden abgewandelt und inhaltlich an die Qualitätsanalyse angepasst aus der Untersuchung von Gärtner und Wurster (2009) zur Wirkung der Schulinspektion in Berlin adaptiert und lassen sich den vorliegenden Ergebnissen gegenüberstellen. Ein direkter Vergleich ist insofern eingeschränkt, als in der Untersuchung von Gärtner und Wurster Schulleitungen (N = 132) und Lehrkräfte (N = 389) befragt wurden und ein anderes Antwortformat zum Einsatz kam. Dennoch zeigen sich in einem eingeschränkten Vergleich des Zusammenhangs der Ziele mit dem Inspektionsbericht folgende Ergebnisse: Während in Berlin ein Viertel (26 %) der befragten Schulleitungen einen Zusammenhang aller Ziele mit dem Inspektionsbericht wahrnimmt, beträgt dieser Anteil aus der Perspektive der Schulaufsicht in NRW 54 % (vgl. Gärtner/Wurster 2009, S. 27f.). Die erfragten negativen und erwünschten Folgen der Schulinspektion sind aus Sicht der Lehrkräfte und Schulleitungen in Berlin und der Schulaufsicht in NRW vergleichbar, da sie in beiden Fällen auf oder nahe der jeweiligen theoretischen Skalenmitte liegen und somit weder eindeutig dem positiven noch dem negativen Meinungsbild entsprechen (vgl. ebd., S. 33). Auch die Einschätzung der Professionalität und Arbeit der Inspektoren weicht aus der Perspektive der Schulaufsicht in NRW nicht wesentlich von den erzielten Befragungsergebnissen im Rahmen der Wirkungsuntersuchung der Schulinspektion Berlin ab (vgl. ebd., S. 36).

Bezüglich der Diskussion methodischer Aspekte ist anzuführen, dass der Pretest auf eine technische und inhaltliche Überprüfung beschränkt blieb und ein Pretest im Feld maßgeblich zur Qualität der Befragung beigetragen hätte. Hinsichtlich der gewählten Befragungsform kann festgehalten werden, dass typische methodenbedingte Nachteile von Onlinebefragungen z. B. in Form der unzureichenden Internetabdeckung für die vorliegende Stichprobe nicht zutreffen. In der Befragung wurde eine Rücklaufquote von 41 % erreicht. Wird die Tatsache berücksichtigt, dass kein personalisierter Zugang mit entsprechender Möglichkeit einer persönlichen Kontaktierung der Teilnehmer gegeben war und die Feldphase zudem in eine bedingt durch mehrere Feiertage typische Urlaubszeit fiel, ist die Teilnahmequote auch unter Berücksichtigung grober Referenzwerte als akzeptabel zu kennzeichnen.

## 11. Triangulation der Studien I, II und III

Dieser Arbeit liegt das Ziel zugrunde, den komplexen Rezeptions-Nutzungs-Prozess von Ergebnissen aus Schulinspektionen aus der Perspektive der Schulen und der Schulaufsicht zu untersuchen. Zur Erreichung dieses Ziel wurde ein Triangulationsdesign mit drei eigenständigen Studien entworfen, in welchen jeweils anhand unterschiedlicher Perspektiven und Methoden verschiedene Facetten des Forschungsgegenstands beleuchtet wurden:

Studie I: Sekundäranalyse schulischer Evaluationsbögen

Studie II: Inhaltsanalyse von Zielvereinbarungen

Studie III: Onlinebefragung der Schulaufsicht

Als Grundlage der triangulativen Verknüpfung dieser Studien gilt das adaptierte und erweiterte Rahmenmodell in Anlehnung an Helmke (2004; s. Kapitel 5). Dieses Modell bildet einen Rahmen, in welchem die Einzelergebnisse der drei Studien verortet und zusammengeführt werden können. Gleichzeitig gelingt mit den neuen Erkenntnissen eine Erweiterung des postulierten Modells.

Während in Studie I auf der Grundlage der Evaluationsbögen die einzelnen Phasen und Rahmenbedingungen des Modells aus der Perspektive der Schulen identifiziert wurden, liefert Studie II anhand der Inhaltsanalyse der Zielvereinbarungsdokumente Anhaltspunkte zur (geplanten) Nutzung der Ergebnisse gleichzeitig aus schulischer als auch aus schulaufsichtlicher Perspektive, da es sich dabei um gemeinsam erarbeitete Dokumente handelt. Als ergänzende Perspektive wird in Studie III die Schulaufsicht berücksichtigt und einzelne Elemente des Modells aus deren Sicht beleuchtet. Die Kombination der drei Studien setzt zum einen auf der Ebene der Ergebnisse, zum anderen am Einzelfall an, indem einem Großteil der Zielvereinbarungen ( $N = 87$ ) der entsprechende Evaluationsbogen einer Einzelschule zugeordnet werden kann. Die Verknüpfung verschiedener Datenquellen oder Methoden am Einzelfall stellt die konsequenteste Form der Triangulation dar. Anhand der „between-method-Triangulation“ gelingt es, die begrenzten Erkenntnismöglichkeiten der einzelnen Studien zu überwinden und die jeweiligen Schwächen und Verzerrungspotentiale der Einzelmethoden zu minimieren. Den Schwächen der Studie I bezüglich der fehlenden Kontrolle der Dateneingabe und -bereinigung sowie der

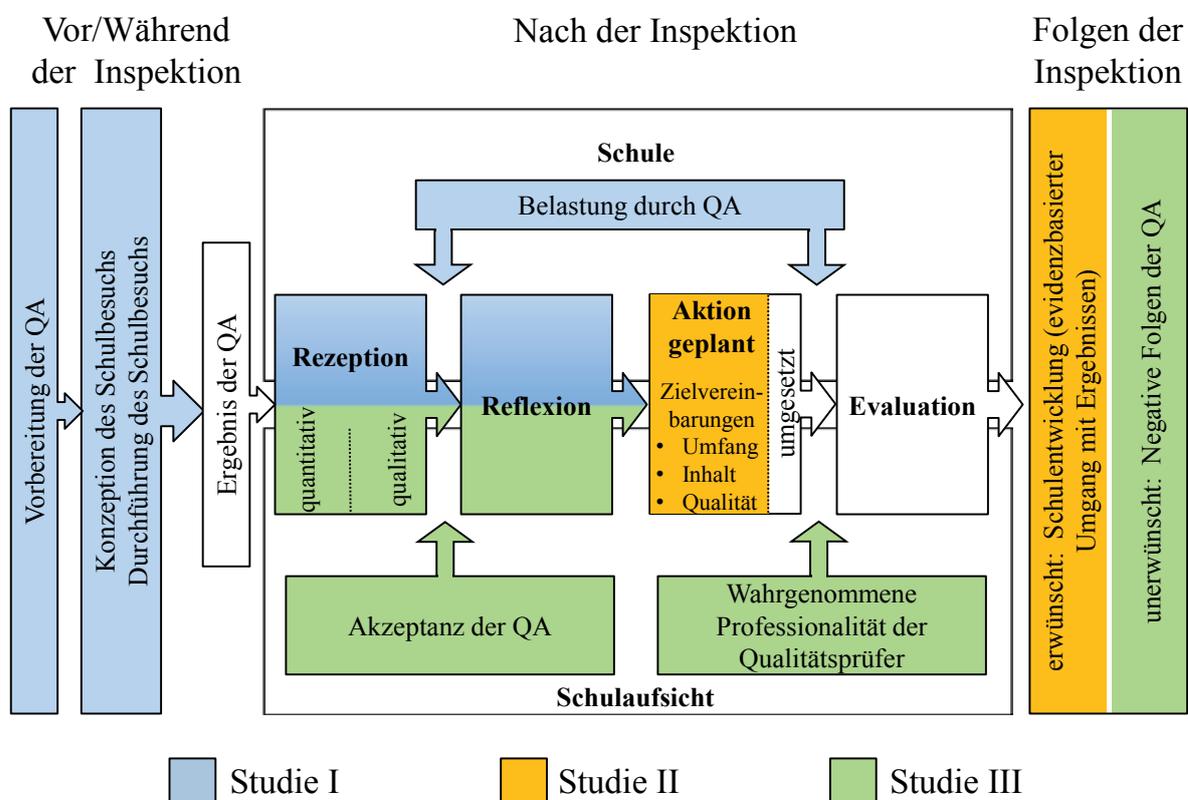
fehlenden Möglichkeit, eine eigens entwickelte Befragung zu konstruieren, wird mit der nichtreaktiven inhaltsanalytischen Auswertung in Studie II und der eigens entwickelten Online-Befragung in Studie III begegnet. Insgesamt erzeugt das triangulative Vorgehen in dieser Arbeit einen Mehrwert, indem damit in die Breite und Tiefe des Forschungsgegenstands vorgedrungen werden konnte, wie es mit einer einzelnen Methode nicht zu erreichen gewesen wäre.

### **11.1. Erweitertes Rezeptions-Nutzungs-Modell von Schulinspektionsergebnissen**

Die bislang separat ausgewerteten Studien I, II und III leisteten bereits einen eigenständigen Beitrag zur Beantwortung der Forschungsfragen aus ihrer jeweiligen methodischen und akteursspezifischen Perspektive. Darauf aufbauend werden im Folgenden die modellbezogenen Ergebnisse der Einzelstudien zusammengeführt und das Rahmenmodell auf der Basis der Ergebnisse ergänzt.

Insbesondere ist es durch die einzelnen Studien gelungen, die unterschiedlichen Prozessphasen *vor*, *während* und *nach der Schulinspektion* abzubilden (vgl. Scheerens 2003, S. 326). In Studie I und III konnten die Phasen *Nach der Schulinspektion (Rezeption und Reflexion)* aus zwei verschiedenen Perspektiven nachgezeichnet werden. Dabei lässt sich die Phase *Rezeption* aus der Perspektive der Schulaufsicht sowohl quantitativ als auch qualitativ differenzieren. Durch die Identifikation der den Rückmeldungen der Schulinspektionsergebnissen vorgelagerten Phasen *Vorbereitung der Qualitätsanalyse, Konzeption* und *Durchführung des Schulbesuchs* konnte das herkömmliche Modell von Helmke (2004) erweitert werden. Diese Phasen stehen in einem mittleren bis hohen Zusammenhang zu den Phasen *Rezeption* und *Reflexion*. Damit geben sie der Annahme Gärtners (2011, S. 158; vgl. auch Gärtner/Pant 2011, S. 26f.) Nahrung, dass die Vorbereitung auf die Schulinspektion bereits Impulse zur Schulentwicklung initiiert. Weiterhin ist es auf der Dokumentengrundlage der Zielvereinbarungen abgeleitet von ihrem Umfang, Inhalt und ihrer Qualität gelungen, die Phase *Aktion (geplant)* zu rekonstruieren. Eine weitere Ergänzung des Helmke-Modells wird mit der Skizzierung der „postaktionalen Phasen“, d. h. den *Folgen der Schulinspektion* erzielt, welche auch im Rahmenmodell zur Wirkung von Schulinspektionen von Ehren und Visscher (2006) Berück-

sichtigung finden. Dabei liefert zum einen Studie II Anhaltspunkte zu den intendierten Folgen der Schulinspektion im Sinne einer evidenzbasierten Nutzung der Rückmeldungen, indem die in den Zielvereinbarungen verschriftlichten Maßnahmenplanungen tatsächlich als erste direkte Schritte der Nutzung interpretiert werden können. Neben den intendierten Folgen der Schulinspektion wurden konträr die nicht-intendierten negativen und unerwünschten Folgen in Studie III anhand der von Gärtner und Wurster (2009, S. 33) adaptierten Items beleuchtet.



**Abb. 25: Auf der Grundlage der Ergebnisse der Studie I, II und III erweitertes Rahmenmodell zum Rezeptions-Nutzungs-Prozess von Schulinspektionsergebnissen**

Darüber hinaus kann als weitere Erkenntnis der durchgeführten Untersuchungen ihr Beitrag zur Benennung konkreter Rahmenbedingungen des Modells herausgestellt werden. Dabei handelt es sich um die *empfundene Belastung* der Schulen durch die Qualitätsanalyse, um die *Akzeptanz der Qualitätsanalyse* und *wahrgenommene Professionalität der Qualitätsprüfer* aus der Perspektive der Schulaufsicht. Schließlich wurde als weitere Modifizierung des Modells die

Beschränkung auf den Akteur Schule aufgehoben und die *Schulaufsicht* mitaufgenommen, da der Abschluss von Zielvereinbarungen durch eine kooperative Zusammenarbeit beider Akteure gekennzeichnet ist. All diese in den drei Studien separat erhobenen Erkenntnisse bzgl. der Phasen, Rahmenbedingungen und Perspektiven werden in das erweiterte Rahmenmodell integriert. Darin wird der Beitrag der einzelnen Studien farblich gekennzeichnet (s. Abb. 25).

## 11.2. Triangulation auf der Ebene der Ergebnisse

Eingangs sei darauf hingewiesen, dass die einzelnen Skalen durch die unterschiedliche Operationalisierung und die verschiedenen inhaltlichen Facetten die direkte Vergleichbarkeit der Ergebnisse aus schulischer und schulaufsichtlicher Perspektive erschweren. Vielmehr können sie auf der Ebene ihrer Ergebnisse zueinander in Beziehung gesetzt werden.

Auf der Grundlage der Befunde der Studie I wurde deutlich, dass den Rückmeldungen der Schulinspektionsergebnisse Prozessphasen vorangehen, die Auswirkungen auf die Rezeption und Reflexion der Ergebnisse haben. Es handelt sich dabei um die Phasen *Vorbereitung der Qualitätsanalyse*, *Konzeption des Schulbesuchs* und *Durchführung des Schulbesuchs*. Diese Phasen beinhalten sowohl die Abstimmung, Planung und Kommunikation zwischen der Schule und dem Qualitätsteam im Vorfeld der Schulbesuchstage, die logistischen und organisatorischen Aspekte des Schulbesuchs als auch das wahrgenommene Verhalten des Qualitätsteams bei den Unterrichtsbesuchen und den Interviews. Dabei wird insbesondere die *Vorbereitung der Qualitätsanalyse* (MW = 3,7; vierstufiges Antwortformat) und die personellen Ressourcen der Qualitätsanalyse (MW = 3,8) von den Schulen übereinstimmend äußerst positiv bewertet. Die Phasen während der Schulinspektion – *Konzeption* und *Durchführung des Schulbesuchs* – werden hingegen von den Schulen vergleichsweise weniger optimistisch bewertet (MW = 3,1 und 3,3). Unter Umständen erweist sich eine mit Ängsten und Widerstand behaftete Reaktion der Schulen bzgl. der Qualitätsanalyse als Erklärung für dieses geteilte Stimmungsbild. Erst nach dem erfolgten Schulbesuch zeigt sich wieder eine positivere Wahrnehmung der Schulen bei der *Rezeption* und *Reflexion* der Ergebnisse: Die Qualitätsberichte werden als überaus prägnant, verständlich, detailliert und aktuell wahrgenommen

(MW = 3,6). Des Weiteren stimmen die Ergebnisse mit den mündlichen Aussagen der Qualitätsprüfer überein und erweisen sich als anschlussfähig. Aus der Perspektive der Schulaufsicht lässt sich die Phase *Rezeption* in eine quantitative und qualitative Komponente differenzieren. Bei der qualitativ orientierten Phasenausprägung, welche die Genauigkeit und Intensität der Auseinandersetzung der Schulaufsicht mit den Ergebnissen erfasst, wird analog zur schulischen Perspektive ein Mittelwert von 3,6 erreicht. Die quantitativ orientierte Rezeptionserfassung enthält die Merkmale der Qualitätsberichte, welche die Rezeption der Ergebnisse beeinflussen. Diese Merkmale wurden über eine fünfstufige Antwortskala erhoben und weisen einen Mittelwert von 4,3 auf. Auch wenn ein direkter Vergleich der schulischen und schulaufsichtlichen Perspektive entfällt, lässt sich in beiden Fällen eine positive Antworttendenz und eine hohe Zustimmung der Ergebnisrezeption konstatieren. Auch bzgl. der *Reflexion der Ergebnisse* zeigen sich aus schulischer und schulaufsichtlicher Perspektive vergleichbare Tendenzen: In Studie I wird die *Reflexion* der Ergebnisse anhand von sechs Items mit einem Mittelwert von 3,3 bewertet, während der Mittelwert in Studie III bei 3,2 liegt (8 Items; in beiden Fällen vierstufige Antwortskala). Aus schulischer Perspektive werden insbesondere diejenigen Merkmale der Qualitätsberichte bewertet, welche maßgeblich die *Reflexion der Ergebnisse* auslösen, während aus schulaufsichtlicher Perspektive zusätzlich die Reflexion über Gelingensbedingungen von Zielvereinbarungen bzgl. ihres Beitrags zur Schulentwicklung erfasst wird. In Studie I hängen die Phasen *Rezeption* und *Reflexion* mit einem Korrelationskoeffizienten von  $r = ,56$  signifikant und stark zusammen. In Studie III ergeben sich bei der Prüfung desselben Zusammenhangs identische signifikante und hohe Korrelationen, welche bei  $r = ,59$  (*Rezeption\_quantitativ* und *Reflexion*) und  $r = ,64$  (*Rezeption\_qualitativ* und *Reflexion*) liegen.

Den Rezeption-Nutzungs-Prozess begleiten unterschiedliche Rahmenbedingungen. Hierbei bringt Studie I die Rahmenbedingung *Belastung durch die Qualitätsanalyse* hervor, während Studie III zusätzlich aus schulaufsichtlicher Perspektive die *Akzeptanz der Qualitätsanalyse* sowie die *Professionalität der Qualitätsprüfer* beleuchtet. Die Rahmenbedingungen werden im Vergleich zu den Prozessphasen sowohl von den Schulen als auch der Schulaufsicht deutlich kritischer bewertet. Inhaltlich liefern die Rahmenbedingungen wichtige An-

haltspunkte über die Stimmungslage der Schulen und der Schulaufsicht im Kontext der Qualitätsanalyse. Hervorzuheben ist, dass aus der Perspektive der Schulaufsicht sowohl die *Akzeptanz der Qualitätsanalyse* als auch die *Professionalität der Qualitätsprüfer* in einer positiven Ausprägung gewährleistet sind, allerdings aus schulischer Sicht auch Belastungen durch die Zusammenstellung der Daten und Dokumente für die Qualitätsanalyse entstehen.

Insgesamt machen diese o. g. Befunde deutlich, dass all jene Merkmale des Rezeptions-Nutzungs-Prozesses, die die Schulen und die Schulaufsicht aktiv gestalten können (Rezeption, Reflexion, Aktion), positiver bewertet werden im Vergleich zu denjenigen Faktoren, welche nicht von diesen Akteuren beeinflussbar sind. Dieser Befund lässt sich dahingehend interpretieren, dass Selbsteinschätzungen i. d. R. immer positiver ausfallen als Fremdeinschätzungen.

Der Umfang gelesener Qualitätsberichte und abgeschlossener Zielvereinbarungen wirkt sich nicht signifikant auf die Wahrnehmung der Phasen und Rahmenbedingungen aus. Die in Studie I und III errechneten Varianzaufklärungsanteile bestimmter Skalen hinsichtlich der Phasen *Rezeption* und *Reflexion* zeigen, dass die Rahmenbedingungen nur wenig zur Varianzaufklärung der Phase *Rezeption* und deutlich mehr zur Phase *Reflexion* beitragen. In Studie I haben sowohl die Phasen vor und während der Qualitätsanalyse als auch die Rahmenbedingung *Belastung durch die Qualitätsanalyse* ein vergleichsweise geringes Aufklärungspotential bzgl. des Varianzanteils der *Rezeption*. Hingegen erweisen sich insbesondere die *Durchführung des Schulbesuchs* und die *Rezeption der Ergebnisse* für die Erklärung der Varianzen im Bereich der *Reflexion* durchaus als äußerst einflussreich ( $R^2 = ,50$ ).

Im Folgenden erfolgt eine triangulative Betrachtung der vereinbarten Ziele auf der Ebene der Ergebnisse. Während die inhaltlichen Kategorien der Zielvereinbarungen in Studie II auf der Grundlage der Dokumente inhaltsanalytisch bestimmt wurden, sind in Studie III die Zielinhalte über Selbstauskünfte der Schulaufsicht durch die Abfrage der drei am häufigsten vereinbarten Zielinhalte erhoben worden. In beiden Zugängen erfolgt die Bestimmung der Zielinhalte anhand der Aspekte des Qualitätstableaus NRW. Dabei zeigt sich ein identisches Bild der vereinbarten Zielinhalte aus der Perspektive der Schulaufsicht und auf der Grundlage der Dokumente. In beiden unterschiedlichen methodischen und perspektivischen Zugängen bei der Zielbestimmung wird überein-

stimmend der dominante Qualitätsbereich 2 *Lernen und Lehren – Unterricht* mit den Qualitätsaspekten 2.6 *Individuelle Förderung und Unterstützung*, 2.2 *Leistungsanforderung und -bewertung* und 2.4 *Unterstützung eines aktiven Lernprozesses* diagnostiziert. Dabei zeigen sich geringe Unterschiede in der Gewichtung der einzelnen Aspekte (s. Tab. 35).

**Tab. 35: Gegenüberstellung der drei häufigsten Zielinhalte in den Studien II und III**

Studie II: Inhaltsanalyse Zielvereinbarungen (N = 339 Zielinhalte)		Studie III: Onlinebefragung der Schulaufsicht (N = 159 Zielinhalte)	
2.4 Unterstützung eines aktiven Lernprozesses	25 %	2.6 Individuelle Förderung und Unterstützung	18 %
2.6 Individuelle Förderung und Unterstützung	13 %	2.2 Leistungsanforderung und -bewertung	17 %
2.2 Leistungsanforderung und -bewertung	12 %	2.4 Unterstützung eines aktiven Lernprozesses	16 %

Der Qualitätsbereich 2 – *Lernen und Lehren – Unterricht*, der mit Abstand am häufigsten in 62 % der Ziele in Studie II und in 76 % der Angaben in Studie III vertreten ist, umfasst das Kerngeschäft der Schulen, d. h. den Unterricht bzw. das Lernen und eine Lernumgebung, in der das Lernen unterstützt und ermöglicht wird. Auf der Ebene der stärker ausdifferenzierten Qualitätsaspekte zeigt sich in Studie II der Qualitätsaspekt 2.4 (25 %), gefolgt vom Qualitätsaspekt 2.6 (13 %) und 2.2 (12 %) in der Wahl der Zielbestimmung den übrigen Qualitätsaspekten deutlich überlegen. Diese Konzentration der Zielinhalte auf spezifische und nur wenige dominante Aspekte des gesamten Qualitätstableaus zeigt sich auch auf der Grundlage der Befunde der Studie III. Laut Angaben der befragten Schulaufsichtsmitarbeiter entfallen 18 % aller genannten Zielinhalte auf den Qualitätsaspekt 2.6 und jeweils 17 % und 16 % der Nennungen auf die Qualitätsaspekte 2.2 und 2.4.

Weiterhin zeigt sich, dass zum überwiegenden Anteil (58 %) Ziele auf der Grundlage derjenigen Qualitätsaspekte definiert werden, in denen die Schulen von der Qualitätsanalyse als *eher schwach als stark* (Bewertungsstufe 2) bewertet werden. Dabei weisen die entsprechenden Bewertungsstufen im negativen Bereich (Stufe 1 und 2) eine signifikante Häufung auf und bei den Bewertungsstufen im positiven Bereich (Stufe 3 und 4) liegt eine signifikante Unterrepräsentation vor. Resümierend ist eine starke Relation zwischen eher

schwach bewerteten Aspekten und den schulischen Handlungsintentionen nach der Qualitätsanalyse erkennbar, was sich als deutlicher Hinweis einer evidenzbasierten Reaktion auf die Ergebnisrückmeldungen interpretieren lässt.

Als weitere Erkenntnis der Studie II gilt die Bestimmung der Zielqualität. Die Spezifität von Zielen ist ein Bestandteil der SMART-Regel, welche vom MSW NRW zur Formulierung von Zielvereinbarungen empfohlen wird. Die vorliegenden Ziele weisen eine mittlere Zielspezifität von 2,3 auf, d. h. sie sind überwiegend *spezifisch* mit einer Tendenz zum Ausprägungsmerkmal *sehr spezifisch* verfasst und beeinflussen damit in positiver Weise die Zielerreichung. Dabei gehen sehr spezifische Ziele tendenziell mit einem höheren Zielvereinbarungsumfang (> 3 Seiten) und mit einer geringeren Anzahl von Zielen (< 3) innerhalb eines Zielvereinbarungsdokuments einher.

Ergänzend liefert die Studie III Anhaltspunkte zum Zielvereinbarungsprozess: Nach Ansicht der befragten Schulaufsichtsbeamten erfolgt die inhaltliche Zielbestimmung überwiegend (zu 61 %) gleichberechtigt von schulischer und schulaufsichtlicher Seite und wird in geringerem Ausmaß (zu 38 %) durch die Schulen gesteuert. Die Rolle der Schulaufsicht in diesem Prozess ist als beratend und unterstützend zu charakterisieren. Die Einleitung und der Abschluss des Zielvereinbarungsprozesses erfolgt zeitnah im Anschluss an die Rückmeldung der Ergebnisse und orientiert sich an der empfohlenen SMART-Regel. Insgesamt wird aus Sicht der Schulaufsicht der Aufwand für die Erarbeitung der Zielvereinbarungen als moderat, jedoch in 40 % der Fälle auch als zu hoch angesehen. Im Prozess der Zielvereinbarung herrscht laut Angaben der befragten Schulaufsicht ein positives und durch Konsensorientierung, Zielbezogenheit und einen wertschätzenden Umgang geprägtes Klima.

### **11.3. Triangulation auf der Ebene des Einzelfalls**

Im Folgenden werden die Ergebnisse der Triangulation auf der Ebene des Einzelfalls präsentiert. In Studie I wurden die modellgeleiteten Phasen *Vorbereitung der Qualitätsanalyse*, *Konzeption des Schulbesuchs*, *Durchführung des Schulbesuchs*, *Rezeption* und *Reflexion der Ergebnisse* faktorenanalytisch identifiziert und in Studie II wurde die Phase *Aktion* anhand des Merkmals Zielqualität inhaltsanalytisch bestimmt. Durch die Verknüpfung beider Datenquellen

auf Einzelschulebene ist die Möglichkeit der Überprüfung der Phasenfolge und der Zusammenhänge zwischen den Phasen gegeben.

Die Fallauswahl für die Triangulation erfolgte auf der Basis der vorliegenden Zielvereinbarungsdokumente, zu denen die entsprechenden Evaluationsbögen zugeordnet wurden. Insgesamt ist es gelungen, 87 der 114 Schulen aus der Zielvereinbarungsstichprobe die entsprechenden Evaluationsbögen zuzuordnen. Da die Evaluationsbögen vom MSW NRW kontinuierlich in das Statistikprogramm SPSS eingetragen werden, erfasst der vorliegende Datensatz nicht zwangsläufig alle aktuell befragten Schulen und somit liegen nicht zu allen 114 Schulen die entsprechenden Evaluationsbögen vor. In Tabelle 36 wird ein Überblick über die Verteilung der Triangulationsstichprobe auf die einzelnen Schulformen und die Schuljahre gegeben. Die am häufigsten vertretenen Schulformen sind Hauptschulen (21,9 %), Realschulen (19,5 %) und Grundschulen (17,2 %). Die Schulformverteilung der Triangulationsstichprobe (n = 87) verhält sich nicht repräsentativ zur Stichprobe der Evaluationsbögen (n = 1155) und aller nordrhein-westfälischen Schulen (N = 6064), da einige Schulformen über- oder unterrepräsentiert sind. In etwa der Hälfte der Stichprobe wurden Zielvereinbarungen im Schuljahr 2009/10 (52,9 %) geschlossen, gefolgt von 20,7 % im Schuljahr 2007/08. Dies entspricht ziemlich genau der Verteilung in der Zielvereinbarungsstichprobe (N = 114 Schulen; s. Studie II).

**Tab. 36: Verteilung der Triangulationsstichprobe nach Schulform und -jahr**

Triangulationsstichprobe (Einzelfall) nach Schulform			Triangulationsstichprobe (Einzelfall) nach Schuljahr		
	(absolut und prozentual)			(absolut und prozentual)	
Grundschule	15	17,2 %	2006/07*	-	-
Förderschule	10	11,5 %	2007/08	18	20,7 %
Hauptschule	19	21,9 %	2008/09	10	11,5 %
Realschule	17	19,5 %	2009/10	46	52,9 %
Gymnasium	10	11,5 %	2010/11**	13	14,9 %
Gesamtschule	2	2,3 %			
Berufskolleg	14	16,1 %			
<b>Gesamt</b>	<b>87</b>	<b>100,0 %</b>	<b>Gesamt</b>	<b>87</b>	<b>100,0 %</b>

\* Im Schuljahr 2006/07 wurde die QA NRW implementiert. Daher erfolgte kein Abschluss von Zielvereinbarungen.

\*\* Es wurden nur Zielvereinbarungen aus dem 1. Schulhalbjahr 2010/11 berücksichtigt. N = 87

Insgesamt wurden die Qualitätsaspekte der 87 Schulen zu 2,4 % mit der Bewertungsstufe 1, zu 12 % mit der Stufe 2, zu 64 % mit der Stufe 3 und zu 23 % mit der Stufe 4 bewertet. Diese Verteilung entspricht beinahe übereinstimmend der Verteilung der Bewertungsstufen in allen 114 Schulen der Zielvereinbarungsstichprobe. Auch der Umfang (Seitenanzahl) der Zielvereinbarungen sowie die Verteilung der Zielqualität in der Triangulationsstichprobe verhalten sich nahezu identisch zur Zielvereinbarungsstichprobe. Werden die mittleren Skalenwerte der Triangulationsstichprobe der Gesamtstichprobe der Evaluationsbögen (N = 1155) gegenübergestellt, so zeigen sich signifikante Unterschiede. Während die Skalen *Rezeption*, *Reflexion*, *Vorbereitung* und *Durchführung* in der Triangulationsstichprobe um 0,2 Skalenpunkte über den Skalenmittelwerten der Evaluationsstichprobe liegen, wird die Skala *Konzeption des Schulbesuchs* um 0,1 Skalenpunkte höher und die Belastung um denselben Wert geringer als in der Referenzstichprobe bewertet. Dies deutet darauf hin, dass in der Triangulationsstichprobe die einzelnen Phasen des Rezeptions-Nutzungs-Prozesses positiver wahrgenommen werden und diese somit als positiv verzerrt gelten könnte.

Zur Bestimmung des Zusammenhangs zwischen den einzelnen präaktionalen Prozessphasen und der Ergebnisnutzung wurden unterschiedliche Berechnungen durchgeführt: Bei der Gegenüberstellung der tatsächlichen und erwarteten Ausprägung der Zielqualität als Indikator der Phase *Aktion* in Abhängigkeit der präaktionalen Phasen ergaben sich keine signifikanten Unterschiede. Überprüft wurde dies anhand von standardisierten Residuen und des Chi-Quadrat-Tests. Da die beobachtete Häufigkeit von den bei Unabhängigkeit zu erwartenden Werten nicht wesentlich abweicht, ist die Abhängigkeit zwischen den präaktionalen Phasen und der Phase *Aktion* nicht gegeben. Damit muss auf der Berechnungsgrundlage von Kreuztabellen eine Auswirkung der präaktionalen Phasen auf die Ergebnisnutzung widerlegt werden. Als weitere Überprüfungsmöglichkeit wurde die Korrelationsanalyse herangezogen. Hierbei wurden die bereits in Studie I aufgezeigten positiven signifikanten Zusammenhänge zwischen den einzelnen präaktionalen Phasen repliziert, während sie mit der neu hinzugefügten Phase *Aktion (geplant)* nicht sinnvoll, d. h. weder positiv noch stark zusammenhängen (s. Tab. 37). Zum einen zeigen die Kendall-Tau-b-Werte äußerst geringe und nicht signifikante und zum anderen sogar negative Korrelationen auf.

**Tab. 37: Skalen-Interkorrelationen auf Einzelfallebene**

Skalen	Konzeption des Schulbesuchs	Durchführung des Schulbesuchs	Rezeption der Ergebnisse	Reflexion der Ergebnisse	Aktion (Zielspezifität)
Vorbereitung der QA	,15	,27*	,43**	,24**	,02
Konzeption des Schulbesuchs		,26*	,18	,45**	-,07
Durchführung des Schulbesuchs			,29**	,33**	-,07
Rezeption der Ergebnisse				,30**	,12
Reflexion der Ergebnisse					-,05

\*\* Die Korrelation ist auf dem Niveau von ,01 (2-seitig) signifikant.

\* Die Korrelation ist auf dem Niveau von ,05 (2-seitig) signifikant.

Korrelationsmaß : Kendall-Tau-b; N = 87

Darüber hinaus ergeben sich weder signifikante noch hohe Korrelationen zwischen den präaktionalen Phasen und der Seiten- und Zielanzahl der Zielvereinbarungen als weitere Indikatoren der Phase *Aktion*. Dies bedeutet nicht, dass es keine Zusammenhänge zwischen den Phasen gibt, sondern dass sie im Rahmen der empirischen Möglichkeiten nicht nachgewiesen werden konnten. Damit erscheint die Verbindung der verschiedenen Datenquellen auf Einzelfallebene nicht praktikabel und bleibt zunächst auf die Ergebnisebene beschränkt.

Als mögliche Erklärungen für die nicht gelungene lineare Verknüpfung der Phasen können zum einen die verschiedenen Datenquellen angeführt werden. Diese messen methodenspezifisch jeweils unterschiedliche Facetten des Rezeption-Nutzungs-Prozesses. Zudem können keine identischen Protagonisten vorausgesetzt werden, welche sowohl den Evaluationsbogen ausgefüllt als auch die Zielvereinbarung mit der Schulaufsicht getroffen haben. Gleichwohl ist davon auszugehen, dass in beiden Fällen die Schulleitung federführend war.

Zum anderen kann der herangezogene Indikator *Zielqualität* für die Phase (*geplante*) *Aktion* in Frage gestellt werden. Weder wird damit die *Aktion* direkt erfasst, noch ist es fraglich, inwiefern dieser Indikator allein eine Phase repräsentieren kann. Es wird unterstellt, dass von der Zielqualität, im Speziellen von der Zielspezifität, unmittelbar auf die Nutzung der Ergebnisse geschlossen werden kann. Damit wird nur ein Merkmal von vielen der im Theorieteil darge-

legten Komponenten der Zielsetzungstheorie erfasst. Unter Berücksichtigung der Komplexität des Übergangs von der präaktionalen zur aktionalen Phase und der Zielsetzungstheorie erscheint die vorgenommene Operationalisierung der Phase *Aktion* beschränkt. Es müsste ein Zugang geschaffen werden, weitere Komponenten der Zielsetzungstheorie zusätzlich zu erfassen, da nur so die von Zielmerkmalen beeinflusste *Aktion* zuverlässig bestimmt werden könnte.

Des Weiteren lassen sich die Phasenzusammenhänge vor dem Hintergrund der Komplexität des Forschungsgegenstands hinterfragen. Auch wenn die *Vorbereitung* und *Durchführung der Qualitätsanalyse* sowie die Ergebnisrezeption und -reflexion positiv verlaufen, so können Hindernisse, welche im schulischen Alltag zahlreich vorliegen, z. B. ausgelöst durch fehlende Zeit oder Priorisierung, die Ergebnisnutzung reduzieren oder gar bedrohen. In der Alltagswelt werden die einzelnen Phasen nicht modellgeleitet vollzogen, sondern es kann an manchen Stellen zu Brüchen kommen. Insofern ist eine Auswirkung der präaktionalen Phasen auf die Nutzung der Ergebnisse modellbasiert intendiert und logisch schlüssig, allerdings wird dies der Komplexität des Prozesses nicht gerecht und sollte nicht simplifiziert werden. Schließlich sind auch methodische Kritikpunkte zur Erklärung der nicht nachweisbaren Phasenzusammenhänge auf Einzelfallebene anzuführen. Diese werden aus der bestehenden Methodenkritik der Triangulation abgeleitet: Flick (2009, S. 30) merkt diesbezüglich an: „Jede Methode konstituiert den Gegenstand, der mit ihr erforscht wird, auf unterschiedliche Weise“. Die einzelnen Beiträge einer triangulativen Zusammenschau erzeugen ihrem jeweiligen Radius entsprechende Erkenntnisse, welche sich nicht zwangsläufig mit den anderen überschneiden.

Als Resümee ergibt sich hinsichtlich der Triangulation der Studien I, II und III folgendes Bild: Die Triangulation konnte auf zwei Forschungsfragen – bzgl. der Erweiterung des Rahmenmodells und der Verknüpfung der Studien auf der Ebene der Ergebnisse – zufriedenstellende Antworten geben. Hingegen erwies sich die Triangulation auf der Ebene des Einzelfalls anhand der unterschiedlichen Methoden und Operationalisierungen nur begrenzt aussagekräftig. Auf der Grundlage dieser Befunde sollten in weiteren Untersuchungen andere Zugangsmöglichkeiten eruiert werden, welche den Rezeptions-Nutzungs-Prozess breiter und tiefer erforschen.

## IV. SCHLUSS

---

### **12. Zusammenfassende Bewertung der Ergebnisse und Ausblick**

In dem abschließenden Kapitel werden der Ausgangspunkt, die Ziele und Fragestellungen der Arbeit rekapituliert. Daran schließt sich eine kurze Wiederholung der theoretischen Einbettung der Arbeit, des methodischen Designs und die abschließende Zusammenfassung der Ergebnisse an. Auf dieser Basis werden die erzielten Ergebnisse reflektiert, Grenzen der Arbeit aufgezeigt und ein Ausblick auf Forschungsdesiderate gegeben. Da die Ergebnisse und die Methoden der Studien I, II und III jeweils in den entsprechenden Kapiteln einer kritischen Reflexion unterzogen wurden, erfolgt dieser für das abschließende Kapitel jeder Arbeit unentbehrliche Bestandteil im vorliegenden Fall nur komprimiert, um unnötige Redundanzen zu vermeiden.

#### **12.1. Ausgangspunkt der Arbeit**

Ausgangspunkt dieser Arbeit war die in der deutschen Bildungslandschaft aktuelle Implementierung der Schulinspektion als externes Evaluationsinstrument mit dem Potential als Impulsgeber für einzelschulische Qualitätsentwicklung. Dies setzt eine Auseinandersetzung der Schulen mit den Ergebnissen und ein Erkennen der Handlungsbedarfe voraus ebenso wie die Transformation der Ergebnisse in Reformbemühungen und konkrete Entwicklungsvorhaben. Bislang steht dieser Prozess dem Unwissen gegenüber, ob und wenn ja, wie die Ergebnisse der Schulinspektion rezipiert und – in Form von Zielvereinbarungen zwischen Einzelschule und Schulaufsicht – für die Weiterentwicklung der einzelnen Schule genutzt werden. Übergreifendes Ziel der vorliegenden Arbeit war es damit, den Umgang von Schulen und der Schulaufsicht mit Ergebnissen der Qualitätsanalyse NRW zu erörtern. Dazu wurde als Grundlage für die empirischen Studien ein bereits bestehendes Modell zur pädagogischen Nutzung von Lernstandserhebungen (vgl. Helmke 2004) weiterentwickelt und dem Untersuchungsfeld Schulinspektion angepasst.

Als theoretische Grundlegung der Fragestellung dienten neben der Aufarbeitung der empirischen Forschungslage der beiden relevanten Forschungsstränge – der Schulinspektions- und Rezeptionsforschung – die Einbettung des Forschungsthemas in den Schulqualitäts- und Schulreformkontext. Im Zuge der sog. Neuen Steuerung wurden eine Reihe von Reformmaßnahmen der internen und externen Evaluation implementiert, von denen sich die Schulverwaltung Impulse für die Schul- und Unterrichtsentwicklung in der Einzelschule verspricht, deren Wirksamkeit im Hinblick darauf aber keineswegs empirisch untersucht und abgesichert ist. Die Schulinspektion als ein Instrument externer Evaluation tangiert unmittelbar die großen Themenbereiche Schulqualität und Schulreform, an denen heute kaum eine wissenschaftliche Untersuchung aus dem bildungswissenschaftlichen Fachbereich vorbeiführt. Schulqualität stellt im Rahmen der Schulinspektion sowohl einen Ausgangspunkt als auch eine Zielgröße dar (vgl. Dederling/Müller/Bos 2008, S. 10). Die Präsentation der Schulinspektion erfolgte begründet durch den Rückgriff auf *policy borrowing* im Kontext ihrer unterschiedlichen Ebenen: Ausgehend von der Schulinspektion im europäischen Ausland wurde die Implementierung dieses Instruments in Deutschland beleuchtet, woran sich auf der dritten Gliederungsebene die Qualitätsanalyse NRW anschloss. In den deutschen Bundesländern zeigt sich neben unterschiedlichen Begrifflichkeiten und vielfältigen Funktionszuweisungen der Schulinspektion eine Pluralität der Schulinspektionsverfahren und -konzepte. Neben der Skizzierung der schulaufsichtlichen Kontextbedingungen in NRW wurde das an die Schulinspektion anknüpfende Instrument Zielvereinbarung mit Exkursen in ökonomische und psychologische Theorieansätze und dem diesbezüglichen Diskussions- und Forschungsstand veranschaulicht.

Bei der Aufarbeitung des empirischen Forschungsstands zu Schulinspektionen wurde auf bereits bestehende Systematisierungen zurückgegriffen: Es wurden Studien zu Einstellungen und Akzeptanz der Schulinspektion, zu Auswirkungen der Schulinspektion auf Schulentwicklung und Schülerleistungen sowie zur Messgüte der Schulinspektionsinstrumente gesichtet. Dabei wurde der Schwerpunkt auf deutschsprachige Studien gelegt. Die Befunde sind z. T. als heterogen und widersprüchlich zu kennzeichnen und machen deutlich, dass Schulleitungen und Lehrkräfte überwiegend positive Akzeptanzbekundungen gegenüber der Schulinspektion äußern, wobei bei Schulleitungen eine ver-

gleichsweise höhere Akzeptanzwahrnehmung zu beobachten ist. Zudem wird die Verfahrensakzeptanz der Schulinspektion positiver bewertet als die Ergebnisaakzeptanz. Bezüglich der Auswirkungen von Schulinspektionen auf Schulentwicklung kann festgehalten werden, dass ein Zusammenhang zwischen der Schulinspektion und der Schulentwicklung besteht, allerdings in eher geringem Ausmaß. Vornehmlich wird diese Erkenntnis bislang über die Anzahl eingeleiteter Maßnahmen als Indikator für Schulentwicklung erfasst und in geringerem Ausmaß über die Qualität und den Erfolg der Maßnahmen. Der Forschungsstand zur Schulinspektion macht insgesamt deutlich, dass sich eine „empirische Forschungslandschaft i. e. S.“ (Brauckmann u.a. 2010, S. 133) aktuell noch entfaltet und diesbezüglich noch viele Desiderate bestehen.

Schließlich wurde im theoretischen Teil dieser Arbeit der Stand der Rezeptionsforschung nicht nur zur Schulinspektion, sondern ganz allgemein zu Ergebnismrückmeldungen aufgearbeitet und dabei insbesondere auch theoretische Modelle zur Ergebnisrezeption und -nutzung berücksichtigt. Auf dieser Grundlage wurde das Helmkesche Rahmenmodell zur Ergebnisnutzung (2004) auf die Analyse der Rezeption und Nutzung von Schulinspektionsergebnissen adaptiert und erweitert. Vor dem Hintergrund der identifizierten Forschungsdefizite wurden folgende Forschungsfragen formuliert, die auf der Basis des modifizierten Rahmenmodells einer Prüfung unterzogen wurden:

1. Wie rezipieren und reflektieren die Schulen die Ergebnismrückmeldungen der Schulinspektion?
2. In welcher Weise spiegeln sich diese Reflexionen in den erarbeiteten Zielvereinbarungen wider, d. h. wie viele und welche Entwicklungsmaßnahmen werden geplant und in welchem Zusammenhang stehen sie zu den Ergebnissen der Schulinspektion?
3. Wie nimmt die Schulaufsicht die Ergebnisse der Schulinspektion wahr, in welcher Weise unterstützt und berät sie die Schulen bei der Maßnahmenplanung?
4. Lässt sich aus den verschiedenen Perspektiven auf den Rückmeldeprozess ein allgemeines Modell der Rezeption und Nutzung von Ergebnissen aus Schulinspektionen ableiten?

## 12.2. Zusammenfassende Bewertung der Studien I, II und III und ihrer Triangulation

Um ein möglichst breites und tiefes Verständnis des Forschungsgegenstands dieser Arbeit zu erzielen, wurde ein Triangulationsdesign entworfen, welches unterschiedliche Perspektiven und differenzierte methodische Zugänge vereint: In a) Studie I wurde eine sekundärstatistische Auswertung ministeriell administrierter Evaluationsbögen für Schulen im Anschluss an die Qualitätsanalyse, in b) Studie II eine Inhaltsanalyse von Zielvereinbarungen und in c) Studie III eine eigens konzipierte Onlinebefragung der Schulaufsicht durchgeführt.

Zu a) Mit Studie I wurde intendiert, einerseits die Gültigkeit des postulierten Rahmenmodells zu prüfen, andererseits erste Hinweise auf die Nutzung von Schulinspektionsergebnissen zu liefern. Die ministeriell administrierten Evaluationsbögen (N = 1155) ließen für dieses Ziel großes Potential erkennen, da die Prozesse der Schulinspektion – Vorgespräch – Besuch – Bericht – im Fragebogen thematisiert wurden und somit eine inhaltliche Parallele zwischen der Primär- und Sekundäranalyse bestand. Mit Hilfe von explorativen und konfirmatorischen Faktorenanalysen ließen sich verschiedene Faktoren identifizieren, mit denen verschiedene Elemente des theoretischen Rahmenmodells indikativiert werden konnten. Insofern bestätigen die identifizierten Faktoren die Abgrenzbarkeit unterschiedlicher Rezeptionsphasen und liefern Hinweise für die im Rahmenmodell spezifizierte zeitliche Abfolge der Phasen. Während sich die Skalen *Rezeption* und *Reflexion* eindeutig im Modell verorten lassen, können die weiteren Skalen *Vorbereitung der Qualitätsanalyse*, *Konzeption* und *Durchführung des Schulbesuchs* sowie die *Belastung durch die Qualitätsanalyse* zusätzlichen Phasen und einer Rahmenbedingung des Modells zugeordnet werden. In inhaltlicher Sicht zeigt sich, dass die Schulen die Vorbereitungs- sowie Ergebnisrückmeldephase als eher positiv wahrnehmen, der Durchführung des Schulbesuchs selbst allerdings deutlich skeptischer gegenüber stehen. Als Erklärung könnten der Widerstand oder auch die Ängste der Lehrkräfte angeführt werden. Anhand der berechneten Zusammenhangsmuster zeigt sich, dass eine positive Wahrnehmung der Vorbereitungs- und Durchführungsphase der Schulinspektion die Auseinandersetzung mit ihren Ergebnissen fördert. Bezüglich der kritischen Reflexion der Methode und Datenlage muss angeführt

werden, dass die Daten typische methodenbedingte Schwächen der Sekundäranalyse, z. B. die fehlende Kontrolle der Datenerhebung und -bereinigung, aufweisen sowie die sozial erwünschte Beantwortung der Evaluationsfragebögen nicht ausgeschlossen werden kann. Zudem wurden die Evaluationsbögen überwiegend von Schulleitungen ausgefüllt und spiegeln nicht zwangsläufig auch die Perspektive der Lehrkräfte wider. In diesem Kontext ist aus der Forschung bekannt, dass Schulleitungen im Vergleich zu Lehrkräften die einzelnen Aspekte der Schulinspektion positiver einschätzen.

Zu b) Studie II wurde mit dem Ziel durchgeführt, formale und inhaltliche Aspekte der zwischen Schulen und Schulaufsicht im Anschluss an die Qualitätsanalyse geschlossenen Zielvereinbarungen zu untersuchen sowie die Zielqualität zu bewerten, um Aufschluss über die Verarbeitung und Nutzung der Ergebnisse zu erhalten. Neben den Zielvereinbarungen (N = 114) standen auch die entsprechenden Qualitätsberichte der Schulen zur Verfügung, so dass eine Analyse der Konkordanz zwischen den durch die Qualitätsanalyse benannten relativen Schwächen einer Schule und den in der Zielvereinbarung genannten Zielen möglich war. Die inhaltsanalytische Auswertung der Dokumente macht deutlich, dass mit Abstand am häufigsten Ziele aus dem *Qualitätsbereich 2 Lernen und Lehren – Unterricht* Gegenstand der Zielvereinbarungen sind: Jede Zielvereinbarung enthält durchschnittlich zwei solcher Ziele und etwa zwei Drittel aller genannten Ziele entfallen auf diesen Bereich. Besonders häufig innerhalb dieses Qualitätsbereichs wird die Förderung einer aktiven Teilnahme am Unterricht genannt (Qualitätsaspekt 2.4). Die vergleichende Betrachtung der in den Qualitätsberichten erzielten Bewertungsstufen mit jenen bei den Zielinhalten zeigt, dass in den Zielvereinbarungen vornehmlich jene Qualitätsaspekte fokussiert werden, die als *erheblich entwicklungsbedürftig* bzw. *eher schwach als stark* bewertet wurden: Der Anteil dieser Bewertungen in den Zielinhalten ist um knapp das Sechsfache höher als in den Qualitätsberichten (58 % vs. 12 %). Dieser Befund kann als unmittelbare, evidenzbasierte Reaktion auf gemessene schulische Prozesse gedeutet werden. Der größte Teil der Zielvereinbarungen ist im Umfang überschaubar und beinhaltet auf bis zu vier Seiten durchschnittlich drei Ziele. Bezüglich der Zielqualität sind die Ziele und Zielvereinbarungen *spezifisch* mit einer Tendenz zu *sehr spezifisch* verfasst. Dies ist auf der Grundlage der theoretisch abgeleiteten Annahmen ein günstiger In-

dikator dafür, Handlungen in positiver Weise zu beeinflussen. Dabei deckt die Zielspezifität ein wichtiges, jedoch nicht das alleinige Kriterium für Zielqualität ab (vgl. SMART-Regel).

Werden die Methode und die Ergebnisse der Studie II einer kritischen Reflexion unterzogen, so zeigt sich zum einen, dass die Verwendung von Dokumenten den Vorteil hat, dass sie direkt verfügbar sind, Messfehler vermieden werden können und sie durch einen nichtreaktiven Charakter gekennzeichnet sind. Die Ergebnisse verweisen auf eine evidenzbasierte Nutzung der Schulinspektionsergebnisse, allerdings nur auf der Ebene der schriftlich formulierten Handlungsabsichten. Über die tatsächliche Umsetzung der geplanten Maßnahmen können auf der Grundlage dieser Arbeit keine Angaben gemacht werden.

Zu c) Schließlich wurden in Studie III die Schulaufsichtsbeamten in NRW zur Rezeption und Nutzung der Ergebnisse aus Qualitätsanalysen online befragt (N = 72). Zusammenfassend zeigt sich das Bild einer befragten Schulaufsicht, welche größtenteils in den vergangenen fünf Jahren seit der Implementierung der Qualitätsanalyse durchaus das Verfahren praxisnah durch Hospitationen kennengelernt hat, mehrere Qualitätsberichte gelesen hat und mit durchschnittlich der Hälfte der inspizierten Schulen Zielvereinbarungen geschlossen hat. Die Ergebnisse bestätigen weitestgehend bereits vorhandene Erfahrungen und Informationen der Schulaufsicht über die Schulen. Analog zu Studie I wurde eine explorative Faktorenanalyse durchgeführt, um einzelne Komponenten des modifizierten Rahmenmodells zu identifizieren. Auf der Basis der Onlinebefragung lassen sich verschiedene Elemente des Modells voneinander abgrenzen und mit den Ergebnissen der Korrelations- und Regressionsanalysen wurden Hinweise auf die sequentielle Phasenstruktur des Modells geliefert. Somit kann das postulierte Modell auch aus Schulaufsichtsperspektive Gültigkeit beanspruchen. In inhaltlicher Sicht schätzen die Schulaufsichtsbeamten die Rezeption und Reflexion der Schulinspektionsergebnisse sowie ihre Umsetzung in Zielvereinbarungen als sehr positiv ein. Dabei hat die Intensität der Auseinandersetzung mit den Ergebnissen der Qualitätsanalyse keine Auswirkung auf die Wahrnehmung der Phasen und Rahmenbedingungen des Rahmenmodells. Kritisch werden stattdessen negative Folgen der Qualitätsanalyse sowie die verfügbaren Ressourcen für Unterstützungsangebote beurteilt.

Am Ende des empirischen Teils wurden die aus den unterschiedlichen Datenquellen gewonnenen Informationen trianguliert, um auf diese Weise ein mehrperspektivisches Bild der Rezeption und Nutzung von Schulinspektionsergebnissen zeichnen zu können. Dabei zeigte sich ein gelungenes Bild auf der Ebene der Modellbestätigung und -erweiterung und der Verknüpfung der drei Studien auf Ergebnisebene. Allerdings kam eine sinnvolle Verbindung auf Einzelschulebene nicht zustande, da sich bei der Verknüpfung der Daten aus Studie I und II keine Zusammenhänge zwischen der Qualität der vereinbarten Ziele und der Wahrnehmung der Rezeption und Reflexion der Inspektionsergebnisse ermitteln ließen. Als Begründung können die unterschiedlichen methodischen Herangehensweisen angeführt werden, welche unterschiedliche Facetten des Rezeptions-Nutzungs-Prozesses beleuchten.

Auf der Grundlage dieser Arbeit kann die eher pessimistische Sicht Böttchers (2009, S. 207), dass durch Evaluationen im besten Fall nichts verbessert werde, nicht bestätigt werden. Auch seine und Keunes (2011, S. 139) Einschätzung bezüglich der Schulinspektion, sie hätte sich „bislang eher nicht als ein Instrument der Schulentwicklung erwiesen“ (s. ebd.) und, dass die Einleitung und Umsetzung konkreter Maßnahmen kein zwingender Effekt von Schulinspektion seien, ebenso wie die Qualität von Maßnahmen als ungesichert gelte, erweist sich auf der Grundlage dieser Arbeit als nicht zutreffend. Auch wenn die tatsächliche Umsetzung der geplanten Maßnahmen in der vorliegenden Arbeit nicht untersucht wurde, offenbaren die Befunde auf der Konzeptionsebene eindeutig eine evidenzbasierte Nutzung der Ergebnisse aus Schulinspektionen durch die Schulen und die Schulaufsicht.

Die Grenzen der Arbeit können insbesondere in den verwendeten Datenquellen gesehen werden. Der Beurteilung der Rezeptions- und Nutzungsphase liegen überwiegend Selbsteinschätzungen der beteiligten Akteure und den analysierten Zielvereinbarungen verschriftlichte Handlungsabsichten zugrunde. Damit ist zwar eine wichtige und notwendige, aber noch nicht hinreichende Voraussetzung für Schul- und Unterrichtsentwicklung gegeben. Im Folgenden können verschiedene Forschungsszenarien benannt werden, die notwendig wären, um das tatsächliche Reform- und Entwicklungspotential der Schulinspektion zu untersuchen.

### 12.3. Ausblick auf zukünftige Forschung

Inwieweit die untersuchten Handlungsabsichten in konkrete Entwicklungsmaßnahmen münden, muss erst in zukünftigen Studien untersucht werden. Weitere bzw. spätere Studien hätten zu ermitteln, ob in denjenigen Schulen, die konkrete Zielvereinbarungen zur Beseitigung erkannter Schwächen getroffen haben, diese nach ein bis drei Jahren tatsächlich behoben sind. Erst dann ließe sich die Nachhaltigkeit der Schulinspektion tatsächlich nachweisen. Dazu müssten die Prozesse von der Ergebnisrezeption über die Entscheidungsfindung bis zur Umsetzung der Maßnahmen wissenschaftlich begleitet werden und dabei die Handlungslogiken der Schulen und der Schulaufsicht rekonstruiert werden, z. B. anhand von Fallanalysen oder Interviews. Ohne qualitative Methoden wird dieser Prozess nicht zufriedenstellend erörtert werden können.

Neben diesen Fragen kann auf weitere Forschungsdesiderate verwiesen werden, die aus der Präsentation und Systematisierung des Forschungsstands resultieren. Es ist deutlich geworden, dass bisher die Schulentwicklungsfunktion der Schulinspektion im Fokus der Forschungsaktivitäten stand. Die bisher noch unzureichend berücksichtigten weiteren Funktionen der Schulinspektion – Wissensgewinnung, Rechenschaftslegung und Normendurchsetzung (vgl. Landwehr 2011) – sollten in zukünftigen Forschungen auch aufgegriffen werden, um der Komplexität der Schulinspektion gerecht zu werden.

Auf der Basis internationaler Forschungsbefunde und theoretischer Überlegungen liegt ein Wirkungsmodell zur Schulinspektion vor, welches neben erwünschten Wirkungen hinsichtlich der Schulentwicklung auch unerwünschte Wirkungen berücksichtigt (vgl. Ehren/Visscher 2006). Es fehlen nach wie vor konkrete Wirkmechanismen, welche in zukünftigen Forschungsbemühungen zu erörtern wären. Daneben ist eine multiperspektivische Einbeziehung aller beteiligten Akteure auf den Prozess gefordert, da bisher die Perspektive der Schüler und Eltern noch zu wenig Berücksichtigung gefunden hat. Ebenso müssten die nicht-intendierten Effekte der Schulinspektion mehr beleuchtet werden.

In diesem Zusammenhang stellt sich auch die Frage nach einem Zukunftsszenario der Schulinspektion im Allgemeinen. Neben einigen anderen Bundesländern ist die Mehrzahl der nordrhein-westfälischen Schulen nach fünf Jahren Praxis der Qualitätsanalyse noch nicht inspiziert worden. Durch die lange Dau-

er der ersten Inspektionswelle ist anzuzweifeln, dass sich ein zweiter Inspektionszyklus nach gleichem Schema anschließt. Bereits seit der Implementierung befindet sich das Instrument in einem ständigen Prozess der Revision und Optimierung. Es ist zu vermuten, dass es im Falle eines zweiten Inspektionsturnus zu einer bedarfsgerechten Schulinspektion und zu einer stärkeren Einbeziehung der Schulen nach niederländischem Vorbild kommen könnte. Dies würde dem internationalen Trend der Schulinspektion entsprechen, welcher sich mit einer zunehmenden Flexibilisierung des Schulinspektionskonzepts und -verfahrens und einer stärkeren Berücksichtigung der Selbstevaluation der Schulen vollzieht. Die Entscheidung, wie ein zweiter Inspektionsturnus aussehen könnte, sollte auch von der Wissenschaft mitgetragen werden. Es liegen bereits umfassende Datenbestände der Schulinspektion in den Bundesländern vor und jedes Bundesland evaluiert das jeweilige Schulinspektionsverfahren. Diese bedeutenden Quellen eignen sich dazu, von der Administration, Wissenschaft und den Schulen kooperativ aufgegriffen zu werden, um so eine optimierte zweite Schulinspektionswelle voranzutreiben. Zurzeit besteht die einmalige historische Chance eines quasiexperimentellen Designs im Rahmen der Schulinspektionsforschung, da bundesweit verschiedene Entwicklungsstände der Schulinspektion zu verzeichnen sind: Von bereits inspizierten, im zweiten Turnus inspizierten und noch nicht inspizierten Schulen. Diese Chance sollte nicht ungenutzt bleiben, da sich dieses Fenster bald schließen wird.

Zuletzt sei angemerkt, dass die Schulinspektion eine Vielzahl an Ergebnissen und Datenmaterial generiert, welche großes Potential für weiterführende wissenschaftliche Aktivitäten haben und eine Erkenntniserweiterung bezüglich guter Schulen und guten Unterrichts und ihrer Entwicklung vermuten lassen. Trotz der hohen Erwartungsstimmung müssen die möglichen Wirkungen der Schulinspektion vergegenwärtigt werden. Dieses Kapitel schließt somit mit einem entsprechenden Zitat von Nils Berkemeyer (2011, S. 125): „Schulinspektionen und Leistungsmessung bieten keine Lösungen für Probleme an. Wer das erwartet, wird zwangsläufig enttäuscht. Sie können aber wichtiges Diagnosewissen liefern.“ Dieses Zitat kann auf der Grundlage der vorliegenden Arbeit wie folgt ergänzt werden: ..., welches bereits von den Schulen auf konzeptioneller Ebene evidenzbasiert genutzt wird, während die Untersuchung des tatsächlichen Reformpotentials der Schulinspektion noch aussteht.



## V. LITERATUR

---

Achtziger, Anja/Gollwitzer, Peter M. (2006): Motivation und Volition im Handlungsverlauf. In: Heckhausen, Jutta/Heckhausen, Heinz (Hrsg.): Motivation und Handeln. Berlin und Heidelberg: Springer (3. Aufl.), S. 277-302.

Ackeren, Isabell van (2003): Evaluation, Rückmeldung und Schulentwicklung. Münster: Waxmann.

Ackeren, Isabell van/Bellenberg, Gabriele (2004): Parallelarbeiten, Vergleichsarbeiten und Zentrale Abschlussprüfungen. Bestandsaufnahme und Perspektiven. In: Holtappels, Heinz Günter/Klemm, Klaus/Pfeiffer, Hermann/Rolff, Hans-Günter/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 13. Weinheim: Juventa, S. 125-159.

Ackeren, Isabell van/Klemm, Klaus (2009): Entstehung, Struktur und Steuerung des deutschen Schulsystems. Eine Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Ackeren, Isabell van/Heinrich, Martin/Thiel, Felicitas (Hrsg.) (2013): Evidenzbasierte Steuerung im Bildungssystem? Befunde aus dem BMBF-SteBis-Verbund. Münster: Waxmann.

Altrichter, Herbert (2009): Datenfeedback und Unterrichtsentwicklung. Probleme eines Kernelements im „neuen Steuerungsmodell“ für das Schulwesen. In: Böttcher, Wolfgang/Dicke, Jan Nikolas/Ziegler, Holger (Hrsg.): Evidenzbasierte Bildung. Wirkungsevaluation in Bildungspolitik und pädagogischer Praxis. Münster: Waxmann, S. 211-226.

Altrichter, Herbert (2010): Konzepte der Fremdevaluation. In: Bohl, Thorsten/Helsper, Werner/Holtappels, Heinz Günter/Schelle, Carla (Hrsg.): Handbuch Schulentwicklung. Bad Heilbrunn: Klinkhardt, S. 170-175.

Altrichter, Herbert (2010): Schul- und Unterrichtsentwicklung durch Datenrückmeldung. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 219-254.

Altrichter, Herbert (2011): Wie steuert sich ein Schulsystem? Annäherungen an einen Begriff mit Konjunktur. In: Knoke, Andreas/Durdel, Anja (Hrsg.): Steuerung im Bildungswesen. Zur Zusammenarbeit von Ministerien, Schulaufsicht und Schulleitungen. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 121-132.

Altrichter, Herbert/Heinrich, Martin (2006): Evaluation als Steuerungsinstrument im Rahmen eines „neuen Steuerungsmodells“ im Schulwesen. In: Böttcher, Wolfgang/Holtappels, Heinz Günter/Brohm, Michaela (Hrsg.): Evaluation im Bildungswesen. Eine Einführung in Grundlagen und Praxisbeispiele. Weinheim: Juventa, S. 51-64.

Altrichter, Herbert/Maag Merki, Katharina (2010): Steuerung der Entwicklung des Schulwesens. In: Dies. (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 15-39.

Altrichter, Herbert/Messner, Elgrid/Posch, Peter (2006): Schulen evaluieren sich selbst. Seelze: Klett/Kallmeyer (2. Aufl.).

Altrichter, Herbert/Rürup, Matthias (2010): Schulautonomie und die Folgen. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 111-144.

Amonat, Haymo (2008): Qualitätsanalyse in Nordrhein-Westfalen: Ablauf und Instrumente. In: Müller, Sabine/Dederling, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 53-69.

Backhaus, Klaus/Erichson, Bernd/Plinke, Wulff/Weiber, Rolf (2010): Multivariate Analysemethoden. Eine anwendungsorientierte Einführung. Heidelberg u.a.: Springer (13. Aufl.).

Backhaus, Klaus/Erichson, Bernd/Weiber, Rolf (2011): Fortgeschrittene multivariate Analysemethoden. Berlin: Springer.

Baltes-Götz, Bernhard (2008): Behandlung fehlender Werte in SPSS und Amos. Universitäts-Rechenzentrum Trier. Verfügbar unter: <http://www.uni-trier.de/fileadmin/urt/doku/bfw/bfw.pdf> [18.09.2011].

Bauer, Karl-Oswald (2004): Evaluation – via regia zu besserer Qualität in Schule und Unterricht? In: Holtappels, Heinz Günter/Klemm, Klaus/Pfeiffer, Hermann/Rolff, Hans-Günter/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 13, Weinheim: Juventa, S. 161-185.

Baumert, Jürgen (2001): Evaluationsmaßnahmen im Bildungsbereich. Eine kritische Sicht auf mögliche Zugänge. In: ÖFEB, 1. Jg., H. 2. Verfügbar unter: <http://www.oefeb.at/newsletter/0102.pdf?PHPSESSID=4b39d2b73fe09b39cc85e8199ca01b48> [18.10.2011].

Baur, Nina/Florian, Michael J. (2009): Stichprobenprobleme bei Online-Umfragen. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 109-128.

Bellenberg, Gabriele (2008): Zur Nutzung von zentralen Abschlussprüfungen als Bausteine eines umfassenden Qualitätssicherungs- und -entwicklungskonzepts – ein Baustellenbericht. In: Böttcher, Wolfgang/Bos, Wilfried/Döbert, Hans/Holtappels, Heinz Günter (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 223-233.

Bellmann, Johannes/Weiß, Manfred (2009): Risiken und Nebenwirkungen Neuer Steuerung im Schulsystem. Theoretische Konzeptualisierung und Erklärungsmodelle. In: Zeitschrift für Pädagogik, 55. Jg., H. 2, S. 286-308.

Benz, Arthur (2004): Multilevel Governance – Governance in Mehrebenen-Systemen. In: Ders. (Hrsg.): Governance – Regieren in komplexen Systemen. Eine Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 125-146.

Berkemeyer, Nils (2008): Selbstevaluation in Schulen – Grundlagen und Beispiele. In: Berkemeyer, Nils/Bos, Wilfried/Manitius, Veronika/Müthing, Kathrin (Hrsg.): Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven. Münster: Waxmann, S. 257-270.

Berkemeyer, Nils (2009): Die Steuerung des Schulsystems. Theoretische und praktische Explorationen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Berkemeyer, Nils (2011): Unterstützungssysteme der Schulentwicklung zwischen Konkurrenz, Kooperation und Kontrolle. In: Altrichter, Herbert/Helm, Christoph (Hrsg.): Akteure und Instrumente der Schulentwicklung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 115-130.

Berkemeyer, Nils/Bos, Wilfried (2009): Professionalisierung im Spannungsfeld externer und interner Evaluation. In: Zlatkin-Troitschanskaia, Olga/Beck, Klaus/Sembill, Detlef/Nickolaus, Reinhold/Mulder, Regina (Hrsg.): Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung. Weinheim/Basel: Beltz, S. 529-541.

Berkemeyer, Nils/Müller, Sabine (2010): Schulinterne Evaluation – nur ein Instrument zur Selbststeuerung von Schulen? In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 195-218.

BMBF (2010a): Rahmenprogramm des BMBF zur Förderung der empirischen Bildungsforschung. Schulinspektion als Steuerungsimpuls zur Schulentwicklung und seine Realisierungsbedingungen auf einzelschulischer Ebene (Verbundvorhaben). Verfügbar unter: <http://www.empirische-bildungsforschung-bmbf.de/zeigen.html?seite=8838> [27.10.2011].

BMBF (2010b): Rahmenprogramm des BMBF zur Förderung der empirischen Bildungsforschung. Externe Evaluation/Schulinspektion und Qualitätsentwicklung von Schulen. Verfügbar unter: <http://www.empirische-bildungsforschung-bmbf.de/zeigen.html?seite=8836> [27.10.2011].

Böhm-Kasper, Oliver/Weishaupt, Horst (2008): Quantitative Ansätze und Methoden der Schulforschung. In: Helsper, Werner/Böhme, Jeanette (Hrsg.): Handbuch der Schulforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 91-123 (2. Aufl.).

Bonsen, Martin/Bos, Wilfried/Rolff, Hans-Günter (2008): Zur Fusion von Schulleffektivitäts- und Schulentwicklungsforschung. In: Bos, Wilfried/Holtappels, Heinz Günter/Pfeiffer, Hermann/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 15, Weinheim: Juventa, S. 11-39.

Bonsen, Martin/Büchter, Andreas/Peek, Rainer (2006): Datengestützte Schul- und Unterrichtsentwicklung. Bewertungen der Lernstandserhebungen in NRW durch Lehrerinnen und Lehrer. In: Bos, Wilfried/Holtappels, Heinz Günter/Pfeiffer, Hermann/Rolff, Hans-Günter/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 14, Weinheim: Juventa, S. 125-148.

Bonsen, Martin/Gathen, Jan von der (2004): Schulentwicklung und Testdaten. Die innerschulische Verarbeitung von Leistungsrückmeldungen. In: Holtappels, Heinz Günter/Klemm, Klaus/Pfeiffer, Hermann/Rolff, Hans-Günter/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 13, Weinheim: Juventa, S. 225-252.

Bortz, Jürgen/Döring, Nicola (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Heidelberg u.a.: Springer (4. Aufl.).

Bos, Wilfried/Bonsen, Martin/Berkemeyer, Nils (2010): Einzelschule und Schülerleistung. In: Bohl, Thorsten/Helsper, Werner/Holtappels, Heinz Günter/Schelle, Carla (Hrsg.): Handbuch Schulentwicklung. Bad Heilbrunn: Klinkhardt, S. 62-65.

Bos, Wilfried/Dedering, Kathrin/Holtappels, Heinz Günter/Müller, Sabine/Rösner, Ernst (2007): Schulinspektion in Deutschland. Eine kritische Bestandsaufnahme. In: Buer, Jürgen van/Wagner, Cornelia (Hrsg.): Qualität von Schule. Ein kritisches Handbuch. Frankfurt am Main: Lang, S. 241-257.

Bos, Wilfried/Holtappels, Heinz Günter/Rösner, Ernst (2006): Schulinspektion in den deutschen Bundesländern – eine Baustellenbeschreibung. In: Bos, Wilfried/Holtappels, Heinz Günter/Pfeiffer, Hermann/Rolff, Hans-Günter/Schulz-Zander, Renate (Hrsg.): Jahrbuch der Schulentwicklung, Band 14, Weinheim: Juventa, S. 81-123.

Böttcher, Wolfgang (2002): Kann eine ökonomische Schule auch eine pädagogische sein? Schulentwicklung zwischen Neuer Steuerung, Organisation, Leistungsevaluation und Bildung. Weinheim: Juventa.

Böttcher, Wolfgang (2009): Was leisten Evaluationen für die Qualitätsentwicklung? In: Bohl, Thorsten/Kiper, Hanna (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 207-217.

Böttcher, Wolfgang/Bos, Wilfried/Döbert, Hans/Holtappels, Heinz Günter (2008): Bildung und Beobachtung. In: Dies. (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 7-12.

Böttcher, Wolfgang/Keune, Miriam (2010): Funktionen und Effekte der Schulinspektion. Ausgewählte nationale und internationale Forschungsbefunde. In: Böttcher, Wolfgang/Dicke, Jan Nikolas/Hogrebe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft. Münster: Waxmann, S. 151-164.

Böttcher, Wolfgang/Keune, Miriam (2011): Wirkungsanalyse der Schulinspektion. Empirische Ergebnisse aus Deutschland. In: Quesel, Carsten/Husfeldt, Vera/Landwehr, Norbert/Steiner, Peter (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep Verlag, S. 123-144.

Böttcher, Wolfgang/Keune, Miriam/Neiwert, Pia (2010): Evaluationsbericht zum Projekt „Schulinspektion in Hessen – Wirkungen auf die Qualitätsentwicklung von Schulen und die Arbeit der Schulaufsicht. Münster: Westfälische Wilhelms-Universität.

Böttcher, Wolfgang/Kotthoff, Hans-Georg (2007a): Schulinspektion zwischen Rechenschaftslegung und schulischer Qualitätsentwicklung: Internationale Erfahrungen. In: Dies. (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann, S. 9-20.

Böttcher, Wolfgang/Kotthoff, Hans-Georg (Hrsg.) (2007b): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann.

Böttcher, Wolfgang/Rürup, Matthias (2010): Landesspezifische Steuerungskonzepte. In: Bohl, Thorsten/Helsper, Werner/Holtappels, Heinz Günter/Schelle, Carla (Hrsg.): Handbuch Schulentwicklung. Bad Heilbrunn: Klinkhardt, S. 56-62.

Böttger-Beer, Manuela/Vaccaro, Didier/Koch, Erik (2010): Wirkmodell zur externen Evaluation. In: Böttcher, Wolfgang/Dicke, Jan Nikolas/Hogrebe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft. Steuerungsinstrumente zwischen Anspruch und Wirklichkeit. Münster: Waxmann, S. 319-335.

Böttger-Beer, Michaela/Koch, Erik (2008): Externe Schulevaluation in Sachsen – ein Dialog zwischen Wissenschaft und Praxis. In: Böttcher, Wolfgang/Bos, Wilfried/Holtappels, Heinz Günter (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 253-264.

Botz, Lieselotta/Heese, Sarah/Schirow, Olga (2011): Qualitätskriterien bei Schulinspektionen. In: Saldern, Matthias von (Hrsg.): Schulinspektion. Fluch und Segen externer Evaluation. Norderstedt: Books on Demand, S. 113-133.

Brauckmann, Stefan/Kühne, Stefan/Stäsche, Uta/Tarazona, Mareike/Weishaupt, Horst/Wittmann, Eveline (2010): Steuerung und Verwaltung des Bildungswesens. In: Jäger, Reinhold S./Nenniger, Peter/Petillon, Hanns/Schwarz, Bernd/Wolf, Bernhard (Hrsg.): Empirische Pädagogik 1990 - 2010. Eine Bestandsaufnahme der Forschung in der Bundesrepublik Deutschland, Band 1, Grundlegende empirische pädagogische Forschung. Landau: Verlag Empirische Pädagogik, S. 119-160.

Breisig, Thomas (2006): Entlohnen und Führen mit Zielvereinbarungen. Methoden, Chancen und Risiken – Wissen für Betriebs- und Personalräte. Frankfurt am Main: Bund Verlag (3. Aufl.).

Brimblecombe, Nicola/Shaw, Marian/Ormston, Michael (1996): Teachers' intention to change practice as a result of OFSTED school inspections. In: Educational Management Administration and Leadership, 24 (4), S. 339-354.

Brimblecombe, Nicola/Ormston, Michael/Shaw, Marian (1995): Teachers' perceptions of school inspection: a stressful experience. In: Cambridge Journal of Education, 25 (1). S. 53-61.

Brosius, Felix (2008): SPSS 16. Das mitp-Standardwerk. Heidelberg: Mitp.

Brückner, Yvonne/Tarazona, Mareike (2010): Finanzierungsformen, Zielvereinbarung, New Public Management, Globalbudgets. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 81-109.

Bruggen, Johan C. van (2006): Schulinspektion in den Niederlanden. Metaevaluation und punktuelle Tiefe auf Wunsch der Schule. In: Buchen, Herbert/Horster, Leonhard/Rolff, Hans-Günter (Hrsg.): Schulinspektion und Schulleitung. Berlin: Raabe, S. 107-124.

Brüsemeister, Thomas/Newiadomsky, Martina (2008): Schulverwaltung – Ein unbekannter Akteur? In: Langer, Roman (Hrsg.): Warum tun die das? Governanceanalysen zum Steuerungshandeln in der Schulentwicklung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 73-93.

Bühl, Achim (2008): SPSS 16. Einführung in die moderne Datenanalyse. München u.a.: Pearson Studium.

Bühner, Markus (2011). Einführung in die Test- und Fragebogenkonstruktion. München u.a.: Pearson Studium (3. Aufl.).

Buhren, Claus G./Rolff, Hans-Günter (2009): Personalmanagement für die Schule. Ein Handbuch für Schulleitung und Kollegium. Weinheim und Basel: Beltz (2. Aufl.).

Bungard, Walter (2000): Zielvereinbarungen – Renaissance eines „alten“ Führungskonzeptes auf Gruppen- und Organisationsebene. In: Bungard, Walter/Kohnke, Oliver (Hrsg.): Zielvereinbarungen erfolgreich umsetzen. Konzepte, Ideen und Praxisbeispiele auf Gruppen- und Organisationsebene. Wiesbaden: Gabler, S. 15-33.

Bungard, Walter/Kohnke, Oliver (Hrsg.) (2000): Zielvereinbarungen erfolgreich umsetzen. Konzepte, Ideen und Praxisbeispiele auf Gruppen- und Organisationsebene. Wiesbaden: Gabler, S. 35-65.

Bürger, Thomas (2007): Plädoyer für die Notwendigkeit einer unterstützenden Schulaufsicht. Eine Auswertung von vorhandenen Studien zur Schulaufsicht. In: Recht der Jugend und des Bildungswesens, 55, Heft 3, S. 381-388.

Burkard, Christoph (2001): Steuerung von Schule durch Schulaufsicht. In: Hofmann, Jan (2001): Schulaufsicht im Umbruch. Neue Aufgaben der Schulaufsicht bei der Qualitätssicherung und -entwicklung von Schule. Kronach u.a.: Carl Link Verlag, S. 51-66.

Burkard, Christoph/Eikenbusch, Gerhard (2000): Praxishandbuch Evaluation in der Schule. Berlin: Cornelsen.

Cheesman, Robert (2007): School Inspection in England. In: Böttcher, Wolfgang/Kotthoff, Hans-Georg (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann, S. 33-41.

Couper, Mick P./Coutts, Elisabeth (2006): Online-Befragung. Probleme und Chancen verschiedener Arten von Online-Erhebungen. In: Diekmann, Andreas (Hrsg.): Methoden der Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 217-243.

Cullingford, Cedric/Daniels, Sandra (1999): Effects of Ofsted inspections on school performance. In: Cullingford, Cedric (Hrsg.): An inspector calls – Ofsted and its effect on school standards, London: Kogan, S. 59-69.

Davies, John D./Lee, John/Postlethwaite, Keith/Tarr, Jane/Thomas, Gary/Yee, Wan Ching (1999): After inspection in special schools: action planning and making progress. In: British Journal of Special Education, Vol. 26, No. 3, S. 130-135.

Dedering, Kathrin (2012): Schulinspektion als wirksamer Weg der Systemsteuerung? In: Zeitschrift für Pädagogik, 58, H.1, S. 69-88.

Dedering, Kathrin/Müller, Sabine (2008): Schulinspektion in Deutschland – Forschungsbereiche und -desiderata. In: Böttcher, Wolfgang/Bos, Wilfried/Döbert, Hans/Holtappels, Heinz Günter (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 241-252.

DeGEval (2012): Standards für Evaluation. Verfügbar unter: <http://www.degeval.de/degeval-standards/standards> [02.01.2012].

Denzin, Norman K. (1970): The research act in sociology: A theoretical introduction to sociological methods. London: Butterworths.

Denzin, Norman K. (1989): The research act in sociology: A theoretical introduction to sociological methods. London/New York: McGraw-Hill (3. Aufl.).

Diekmann, Andreas (2009): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt (20. Aufl.).

Diemer, Tobias/Kuper, Harm (2010): Formen, Grenzen und Perspektiven der innerschulischen Nutzung zentraler Lernstandserhebungen als Instrument neuer Steuerung. In: Böttcher, Wolfgang/Dicke, Jan Nikolas/Hogrebe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft. Münster: Waxmann, S. 259-267.

Diemer, Tobias/Kuper, Harm (2011): Formen innerschulischer Steuerung mittels zentraler Lernstandserhebungen. In: Zeitschrift für Pädagogik, 57, H. 4, S. 554-571.

Ditton, Hartmut (2000): Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In: Zeitschrift für Pädagogik, 41. Beiheft, S. 73-92.

Ditton, Hartmut (2007): Schulqualität – Modelle zwischen Konstruktion, empirischen Befunden und Implementierung. In: Buer, Jürgen van/Wagner, Cornelia (Hrsg.): Qualität von Schule. Ein kritisches Handbuch. Frankfurt am Main: Lang, S. 83-93.

Ditton, Hartmut (2010): Evaluation und Qualitätssicherung. In: Tippelt, Rudolf/Schmidt, Bernhard (Hrsg.): Handbuch Bildungsforschung. Wiesbaden: VS Verlag für Sozialwissenschaften (3. Aufl.), S. 607-623.

Döbert, Hans/Rürup, Matthias/Dedering, Kathrin (2008): Externe Evaluation von Schulen in Deutschland – die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede. In: Döbert, Hans/Dedering, Kathrin (Hrsg.): Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte. Münster: Waxmann, S. 63-151.

Döbrich, Peter / Schnell, Herbert / Sroka, Wendelin (2008): Schulinspektion in ausgewählten Ländern. In: Döbert, Hans / Dederich, Kathrin (Hrsg.): Externe Evaluation von Schulen. Münster: Waxmann, S. 165-237.

Dorn, Michael (2008): Auswertung der Daten – das Programm ELBA. In: Müller, Sabine / Dederich, Kathrin / Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLichterhand, S. 93-102.

Dubs, Rolf (2005): Innovationen, Schulentwicklung und Qualitätsmanagement – eine ganzheitliche Sicht. In: Senatsverwaltung für Bildung, Jugend und Sport Berlin / Friedrich-Ebert-Stiftung (Hrsg.): Innovationen lohnen sich. Fachtagung in der Friedrich-Ebert-Stiftung Berlin, 19. August 2004. Berlin, S. 9-20. Verfügbar unter: <http://library.fes.de/pdf-files/bueros/berlin/50272.pdf> [17.12.2011].

Dubs, Rolf (2009): Schulaufsicht. In: Blömeke, Sigrid / Bohl, Thorsten / Haag, Ludwig / Lang-Wojtasik, Gregor / Sacher, Werner (Hrsg.): Handbuch Schule. Bad Heilbrunn: Klinkhardt, S. 511-517.

Ehren, Melanie C. M. / Visscher, Adrie J. (2006): Towards a theory on the impact of school inspections. In: British Journal of Educational Studies, Vol. 54, No. 1, S. 51-72.

Ehren, Melanie C. M. / Visscher, Adrie J. (2008): The relationships between school inspections, school characteristics and school improvement. In: British Journal of Educational studies, Vol. 56, No. 2, S. 205-227.

Ernst, Christian / Döbert, Hans (2006): Evaluation und dann? In: PädForum: Unterrichten und Erziehen, 34. Jg., H. 1, S. 54-56.

Exekutivagentur Bildung, Audiovisuelles und Kultur (P9 EURYDICE) (2009): Schlüsselzahlen zum Bildungswesen in Europa 2009. Verfügbar unter: [http://eacea.ec.europa.eu/education/eurydice/documents/key\\_data\\_series/105DE.pdf](http://eacea.ec.europa.eu/education/eurydice/documents/key_data_series/105DE.pdf) [19.09.2011].

Fabrigar, Leandre R. / Wegener, Duane T. / MacCallum, Robert C. / Strahan, Erin J. (1999): Evaluating the Use of Exploratory Factor Analysis in Psychological Research. In: Psychological Methods, Vol. 4, No. 3, S. 272-299.

Fend, Helmut (2000): Qualität und Qualitätssicherung im Bildungswesen. Wohlfahrtsstaatliche Modelle und Marktmodelle. In: Zeitschrift für Pädagogik, 41. Beiheft, S. 55-72.

Fend, Helmut (2011): Die Wirksamkeit der Neuen Steuerung – theoretische und methodische Probleme ihrer Evaluation. In: Zeitschrift für Bildungsforschung, 1, S. 5-24.

Flick, Uwe (2007): *Qualitative Sozialforschung. Eine Einführung*. Reinbek: Rowohlt (5. Aufl.).

Flick, Uwe (2008): *Triangulation. Eine Einführung*. VS Verlag für Sozialwissenschaften: Wiesbaden (2. Aufl.).

Flick, Uwe (2009): *Sozialforschung. Methoden und Anwendungen*. Reinbek: Rowohlt.

Flick, Uwe (2010): *Triangulation*. In: Mey, Günter/Mruck, Katja (Hrsg.): *Handbuch Qualitative Forschung in der Psychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 278-289.

Friedrichs, Jürgen (1990): *Methoden empirischer Sozialforschung*. Opladen: Westdeutscher Verlag (14. Aufl.).

Fromm, Sabine (2010): *Datenanalyse mit SPSS für Fortgeschrittene 2: Multivariate Verfahren für Querschnittsdaten*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Früh, Werner (2007): *Inhaltsanalyse*. Konstanz: UVK Verlagsgesellschaft (6. Aufl.).

Fuchs, Hans-Werner (2008): *Educational Governance und neue Steuerung: Grundsätze – Beispiele – Erwartungen*. In: Langer, Roman (Hrsg.): *„Warum tun die das?“ Governanceanalysen zum Steuerungshandeln in der Schulentwicklung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 19-29.

Füssel, Hans-Peter (2008): *Schulinspektion und Schulaufsicht, Schulinspektion oder Schulaufsicht, Schulinspektion versus Schulaufsicht, Schulinspektion als Schulaufsicht?* In: Döbert, Hans/Dedering, Kathrin (Hrsg.): *Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte*. Münster: Waxmann, S. 153-163.

Füssel, Hans-Peter/Kretschmann, Rudolf (2005): *Verträge im Bildungsbereich – Chancen und Grenzen*. In: *Recht der Jugend und des Bildungswesens*, 53, H. 1, S. 56-70.

Gärtner, Holger (2010): *Aktuelle Studien zu Effekten von Schulinspektionsverfahren in Berlin/Brandenburg*. Vortrag auf der ARGEV Fachtagung „Wirkungen und Wirksamkeit der externen Schulevaluation“ am 02.07.2010 in Baden, Schweiz. Verfügbar unter: [http://www.argev.ch/\\_files/Veranstaltungen/10\\_Juli\\_Wirksamkeit/Pr%C3%A4sentation\\_G%C3%A4rtner.pdf](http://www.argev.ch/_files/Veranstaltungen/10_Juli_Wirksamkeit/Pr%C3%A4sentation_G%C3%A4rtner.pdf) [22.10.2011].

Gärtner, Holger (2011): *Die Wirkung von Schulinspektion auf Schulentwicklung. Eine quasi-experimentelle Feldstudie*. In: Quesel, Carsten/Husfeldt, Vera/Landwehr, Norbert/Steiner, Peter (Hrsg.): *Wirkungen und Wirksamkeit der externen Schulevaluation*. Bern: hep Verlag, S. 145-161.

Gärtner, Holger/Pant, Hans Anand (2011): Validierungsstrategien für Verfahren und Ergebnisse von Schulinspektion. In: Müller, Sabine/Pietsch, Marcus/Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 9-31.

Gärtner, Holger/Wurster, Sebastian (2009): Befragung zur Wirkung von Schulinspektion in Berlin. Ergebnisbericht. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg. Verfügbar unter: [http://www.isq-bb.de/uploads/media/ISQ\\_Ergebnisbericht\\_Schulinspektion\\_Berlin\\_final.pdf](http://www.isq-bb.de/uploads/media/ISQ_Ergebnisbericht_Schulinspektion_Berlin_final.pdf) [15.04.2010].

Geiser, Christian (2010): Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Geißler, Gert (2008): Geschichtliches zur „Vornahme periodischer Untersuchungen“ der Schulen. In: Döbert, Hans/Dedering, Kathrin (Hrsg.): Externe Evaluation von Schulen. Historische, rechtliche und vergleichende Aspekte. Münster: Waxmann, S. 23-61.

Glaser, Edith (2010): Dokumentenanalyse und Quellenkritik. In: Friebertshäuser, Barbara/Langer, Antje/Prenzel, Annedore (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim und München: Juventa, S. 365-375 (3. Aufl.).

Gläser-Zikuda, Michaela (2010): Kombination qualitativer und quantitativer Methoden in der Schulentwicklungsforschung. In: Bohl, Thorsten/Helsper, Werner/Holtappels, Heinz Günter/Schelle, Carla (Hrsg.): Handbuch Schulentwicklung. Bad Heilbrunn: Klinkhardt, S. 154-156.

Globalpark AG (2011): Handbuch EFS Survey 8.0. Verfügbar unter: [http://www.unipark.info/files/efssurvey80manual\\_ger\\_2011-06-07.pdf](http://www.unipark.info/files/efssurvey80manual_ger_2011-06-07.pdf) [25.07.2011].

Gräf, Lorenz (2010): Online-Befragung. Eine praktische Einführung für Anfänger. Münster: LIT Verlag.

Gray, John/Wilcox, Brian (1995): In the aftermath of inspection: the nature and fate of inspection report recommendations. In: Research Papers in Education, vol. 10, no 1, S. 1-18.

Gröhlich, Carola (2012): Bildungsqualität. Strukturen und Prozesse in Schule und Unterricht und ihre Bedeutung für den Kompetenzerwerb. Münster: Waxmann.

Groß Ophoff, Jana/Hosenfeld, Ingmar/Koch, Ursula (2007): Formen der Ergebnisrezeption und damit verbundene Schul- und Unterrichtsentwicklung. In: Empirische Pädagogik, 21, H. 4, S. 411-427.

Gruehn, Sabine (2000): Unterricht und schulisches Lernen. Münster: Waxmann.

Gruschka, Andreas (2010): Die Schulinspektion war da und hinterließ einen Bericht. In: Pädagogische Korrespondenz, 41, S. 75-92.

Gruschka, Andreas/Heinrich, Martin/Köck, Nicole/Martin, Ellen/Pollmanns, Maroin/Tiedtke, Michael (2003): Innere Schulreform durch Kriseninduktion? Fallkonstruktionen und Strukturanalysen zu den Wirkungen administrativ verordneter Schulprogrammarbeit. Projektdesign und Zwischenbericht. Frankfurt am Main: Fachbereich Erziehungswissenschaften der Johann Wolfgang Goethe-Universität.

Hartung-Beck, Viola (2009): Schulische Organisationsentwicklung und Professionalisierung. Folgen von Lernstandserhebungen an Gesamtschulen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Harvey, Lee/Green, Diana (Übersetzung von Terhart, Ewald) (2000): Qualität definieren. Fünf unterschiedliche Ansätze. In: Zeitschrift für Pädagogik. 41. Beiheft, S. 17-39.

Heckhausen, Heinz/Gollwitzer, Peter M. (1987): Thought contents and cognitive functioning in motivational versus volitional states of mind. In: Motivation and Emotion, 11, S. 101-120.

Heckhausen, Jutta/Heckhausen, Heinz (Hrsg.) (2006): Motivation und Handeln. Berlin und Heidelberg: Springer (3. Aufl.).

Heinrich, Martin (2007): Governance in der Schulentwicklung. Von der Autonomie zur evaluationsbasierten Steuerung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Heinrich, Martin (2008): Von der Steuerung zu Aushandlungsprozessen als neue Form der Handlungskoordination. In: Langer, Roman (Hrsg.): ‚Warum tun die das?‘ Governanceanalysen zum Steuerungshandeln in der Schulentwicklung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 31-46.

Heinrich, Martin (2010): Testen, prüfen, vergleichen – und dann? Über die Auswirkungen von Lernstandserhebungen und Vergleichsarbeiten auf die Lehrerverarbeit. In: Friedrich Jahresheft XXVIII, Lehrerverarbeit – Lehrer sein. Friedrich-Verlag: Seelze, S. 116-119.

Helmke, Andreas (2004). Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. In: Das Seminar – Lehrerbildung und Schule, 2, S. 90-112.

Helmke, Andreas (2009): Unterrichtsqualität und Lehrerverprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze-Velber: Kallmeyer/Klett (2. Aufl.).

Herrgesell, Peter (2004): Schulische Qualitätsprogramme – arbeitsökonomisch, kurzfristig und objektiv analysieren. Einsatz von MAXqda in der Praxis pädagogischer Schulaufsicht. In: Kuckartz, Udo/Grünenberg, Heiko/Lauterbach, Andreas (Hrsg.): Qualitative Datenanalyse: computergestützt. Methodische Hintergründe und Beispiele aus der Forschungspraxis. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 139-153.

Hessisches Kultusministerium (2007): Zielvereinbarungen nach Schulinspektion. Leitfaden für die Entwicklung von Zielvereinbarungen zwischen Schule und Schulaufsicht nach der Schulinspektion. Wiesbaden. Verfügbar unter: [http://dms-schule.bildung.hessen.de/allgemeines/begabung/broschueren/Zielvereinbarungen\\_nach\\_Schulinspektion\\_1\\_.pdf](http://dms-schule.bildung.hessen.de/allgemeines/begabung/broschueren/Zielvereinbarungen_nach_Schulinspektion_1_.pdf) [22.11.2011].

Hofmann, Jan (Hrsg.) (2001): Schulaufsicht im Umbruch. Neue Aufgaben der Schulaufsicht bei der Qualitätssicherung und -entwicklung von Schule. Kronach u.a.: Carl Link Verlag.

Holtappels, Heinz Günter (2004): Schulprogramm – ein Instrument zur systematischen Entwicklung der Schule. In: Ders. (Hrsg.): Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen. Weinheim und München: Juventa, S. 11-28.

Holtappels, Heinz Günter (2011): Schulinterne Steuerungsinstrumente der Schulentwicklung. In: Altrichter, Herbert/Helm, Christoph (Hrsg.): Akteure und Instrumente der Schulentwicklung. Baltmannsweiler: Schneider-Verl. Hohengehren, S. 131-149.

Holtappels, Heinz Günter (Hrsg.) (2004): Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen. Weinheim und München: Juventa.

Holtappels, Heinz Günter/Müller, Sabine (2002): Inhalte und Struktur von Schulprogrammen. Inhaltsanalyse der Schulprogramm-Texte Hamburger Schulen. In: Jahrbuch der Schulentwicklung, 12, S. 209-231.

Holtappels, Heinz Günter/Müller, Sabine (2004): Inhalte von Schulprogrammen – Ergebnisse einer Inhaltsanalyse Hamburger Schulprogrammtexte. In: Holtappels, Heinz Günter (Hrsg.): Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen. Weinheim und München: Juventa, S. 79-116.

Hölzle, Philipp (2000): Die Anwendung des Führungsinstruments „Zielvereinbarung“ in der deutschen Wirtschaft. In: Braun, Ottmar L. (Hrsg.): Zielvereinbarungen im Kontext strategischer Organisationsentwicklung, Landau: Verlag Empirische Pädagogik, S. 75-100.

Homeier, Wulf (2008): Zielvereinbarungen – was nach der Qualitätsanalyse folgt. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Köln: Wolters Kluwer, S. 129-137.

Hosenfeld, Annette (2010): Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung? Die Wirkung von videographischer und schriftlicher Rückmeldung bei Lehrkräften der vierten Jahrgangsstufe. Münster: Waxmann.

Hosenfeld, Ingmar/Zimmer-Müller, Michael (2010): Vergleichende Leistungsstudien. In: Jäger, Reinhold S./Nenniger, Peter/Petillon, Hanns/Schwarz, Bernd/Wolf, Bernhard (Hrsg.): Empirische Pädagogik 1990 - 2010. Eine Bestandsaufnahme der Forschung in der Bundesrepublik Deutschland, Band 1, Grundlegende empirische pädagogische Forschung. Landau: Verlag Empirische Pädagogik, S. 89-118.

Hovestadt, Gertrud (2009): Externe Evaluation und datengestützte Schulentwicklung – Eine Bestandsaufnahme in den Bundesländern. In: Bohl, Thorsten/Kiper, Hanna (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 29-41.

Huber, Franz (2006): Konsequenzen aus der externen Evaluation an Bayerns Schulen. Auswertung einer Befragung von Schulleiterinnen und Schulleitern zu den Entwicklungen an ihren Schulen. München: ISB QA. Verfügbar unter: <http://www.isb.bayern.de/isb/download.aspx?DownloadFileID=7e7f68ed30a8d4371beb6f76725a7bc4> [22.10.2009].

Huber, Franz (2008): Konsequenzen aus der externen Evaluation an Bayerns Schulen. Auswertung einer Befragung von Schulleiterinnen und Schulleitern zu den Entwicklungen an ihren Schulen. In: Böttcher, Wolfgang/Bos, Wilfried/Döbert, Hans/Holtappels, Heinz Günter (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 265-278.

Huber, Günther L. (1989): Qualität versus Quantität in der Inhaltsanalyse. In: Bos, Wilfried/Tarnai, Christian (Hrsg.): Angewandte Inhaltsanalyse in Empirischer Pädagogik und Psychologie. Münster: Waxmann, S. 32-47.

Huber, Stephan Gerhard/Skedsmo, Guri/Durrer, Luzia (2011): EU-Project: Impact of School Inspection on Teaching and Learning (ISI-TL). Poster auf dem Schulleitungssymposium 2011. Verfügbar unter: <http://www.schulleitungssymposium.net/2011/pdf/vortraege/poster/ISI-TL-SLS-Poster-11-09-03-end.pdf> [03.01.2012].

Husfeldt, Vera (2011a): Wirkungen und Wirksamkeit der externen Schulevaluation. Überblick zum Stand der Forschung. In: Zeitschrift für Erziehungswissenschaft, 14. Jg., H. 2, S. 259-282.

Husfeldt, Vera (2011b): Externe Schulevaluation. Ein länderübergreifender Blick auf Forschungen und Modelle. In: Quesel, Carsten/Husfeldt, Vera/Landwehr, Norbert/Steiner, Peter (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep Verlag, S. 13-33.

Jaccard, James/Wan, Choi K. (1996): LISREL Approaches to Interaction Effects in Multiple Regression. Sage University series in Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage.

Jackob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.) (2009): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Jacob, Rüdiger/Heinz, Andreas/Décieux, Jean Philippe/Eirmbter, Willy H. (2011): Umfrage. Einführung in die Methoden der Umfrageforschung. München: Oldenbourg (2. Aufl.).

Jäger, Michael (2004): Transfer in Schulentwicklungsprojekten. Wiesbaden: VS Verlag für Sozialwissenschaften.

Jäger, Hildegard (2013): Neuausrichtung der Qualitätsanalyse. In: Schule NRW, 10/2013, S. 492-494.

Janssens, Frans J. G. (2007): Supervising the Quality of Education. In: Böttcher, Wolfgang/Kotthoff, Hans-Georg (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann, S. 113-132.

Kelle, Udo (2009): Computergestützte Analyse qualitativer Daten. In: Flick, Uwe/Kardorff, Ernst von/Steinke, Ines (Hrsg.): Qualitative Forschung: Ein Handbuch. Reinbek: Rowohlt, S. 485-501 (7. Aufl.).

Kelle, Udo/Erzberger, Christian (2009): Qualitative und quantitative Methoden: kein Gegensatz. In: Flick, Uwe/Kardorff, Ernst von/Steinke, Ines (Hrsg.): Qualitative Forschung: Ein Handbuch. Reinbek: Rowohlt, S. 299-308 (7. Aufl.).

Keune, Miriam Sharon (2014): Schulinspektion unter besonderer Berücksichtigung externer Zielvereinbarungen. Eine explorative ‚mixed methods‘-Studie am Beispiel der hessischen Schulinspektion. Münster: Monsenstein und Vannerdat.

Key, Tim (2007a): School Inspection in England: Evaluation, Accountability and Quality Development. In: Böttcher, Wolfgang/Kotthoff, Hans-Georg (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann, S. 21-32.

Key, Timothy (2007b): Vom Versagen zum Erfolg: Die Rolle der Schulinspektionen bei der Verbesserung von Schulen. In: Reusch, Christiane (Hrsg.): Lehrer unter Druck. Gütersloh: Verl. Bertelsmann-Stiftung, S. 17-35.

Kiper, Hanna (2008): Diskurse zur Unterrichtsentwicklung – eine kritische Betrachtung. In: Berkemeyer, Nils/Bos, Wilfried/Manitius, Veronika/Müthing, Kathrin (Hrsg.): Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven. Münster: Waxmann, S. 95-119.

Kirchhoff, Sabine/Kuhnt, Sonja/Lipp, Peter/Schlawin, Siegfried (2010): Der Fragebogen. Datenbasis, Konstruktion und Auswertung. Wiesbaden: VS Verlag für Sozialwissenschaften (5. Aufl.).

Klein, Esther Dominique/Kühn, Svenja Mareike/Ackeren, Isabell van/Block, Rainer (2009): Wie zentral sind zentrale Prüfungen? Abschlussprüfungen am Ende der Sekundarstufe II im nationalen und internationalen Vergleich. In: Zeitschrift für Pädagogik, 55. Jg., H. 4, S. 596-621.

Kleinbeck, Uwe (2006): Handlungsziele. In: Heckhausen, Jutta/Heckhausen, Heinz (Hrsg.): Motivation und Handeln. Berlin und Heidelberg: Springer (3. Aufl.), S. 255-276.

Klieme, Eckhard/Tippelt, Rudolf (2008): Qualitätssicherung im Bildungswesen. Eine aktuelle Zwischenbilanz. In: Zeitschrift für Pädagogik, 53. Beiheft, S. 7-13.

Klingemann, Hans-Dieter/Mochmann, Ekkehard (1975): Sekundäranalyse. In: Koolwijk, Jürgen van/Wieken-Mayser, Maria (Hrsg.): Techniken der empirischen Sozialforschung. Band 2. München: Oldenbourg, S. 178-195.

Klippert, Heinz (2010): Schulinspektion – und was dann? Problemanzeigen zur gängigen Evaluationspraxis. In: Pädagogik, H. 10, S. 28-31.

Kluchert, Gerhard/Leschinsky, Achim (2001): Woher die Karawane kommt. Schulaufsicht zwischen Kontrolle und Beratung – historisch betrachtet. In: Hofmann, Jan (Hrsg.): Schulaufsicht im Umbruch. Neue Aufgaben der Schulaufsicht bei der Qualitätssicherung und -entwicklung von Schule. Kronach u.a.: Carl Link Verlag, S. 13-32.

KMK (Hrsg.) (2006): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. München: LinkLuchterhand in Wolters Kluwer.

KMK (Hrsg.) (2010): Das Bildungswesen in der Bundesrepublik Deutschland 2009. Darstellung der Kompetenzen, Strukturen und bildungspolitischen Entwicklungen für den Informationsaustausch in Europa. Verfügbar unter: [http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen\\_pdfs/dossier\\_dt\\_ebook.pdf](http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen_pdfs/dossier_dt_ebook.pdf) [28.12.2011].

Koch, Stefan/Gräsel, Cornelia (2004): Schulreformen und Neue Steuerung – erziehungs- und verwaltungswissenschaftliche Perspektiven. In: Koch, Stefan/Fisch, Rudolf (Hrsg.): Schulen für die Zukunft. Neue Steuerung im Bildungswesen. Baltmannsweiler: Schneider Verlag Hohengehren, S. 3-24.

Koch, Ursula (2011): Verstehen Lehrkräfte Rückmeldungen aus Vergleichsarbeiten? Datenkompetenz von Lehrkräften und die Nutzung von Ergebnissrückmeldungen aus Vergleichsarbeiten. Münster: Waxmann.

Koch, Ursula/Groß Ophoff, Jana/Hosenfeld, Ingmar/Helmke, Andreas (2006): Qualitätssicherung: Von der Evaluation zur Schul- und Unterrichtsentwicklung – Ergebnisse der Lehrerbefragungen zur Auseinandersetzung mit den VERA-Rückmeldungen. In: Eder, Ferdinand (Hrsg.): Qualität durch Standards? Münster: Waxmann, S. 187-199.

Koch-Riotte, Barbara (2008): Kommunikationstraining: Wichtiger Ausbildungsbaustein zur Vorbereitung der Qualitätsprüferinnen und -prüfer. In: Müller, Sabine/Dederling, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 122-128.

Kohler, Britta (2004): Zur Rezeption externer Evaluation durch Lehrkräfte, Eltern sowie Beamte der Schulaufsicht. In: Empirische Pädagogik, 18, 1, S. 18-39.

Kohler, Britta (2005): Rezeption internationaler Schulleistungstudien. Wie gehen Lehrkräfte, Eltern und die Schulaufsicht mit Ergebnisse schulischer Evaluationsstudien um? Münster: Waxmann.

Kohler, Britta (2009): Umgang von Lehrer/innen, Eltern und Schulaufsicht mit Ergebnissen internationaler Schulleistungstudien. In: Bohl, Thorsten/Kiper, Hanna (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 81-96.

Kohnke, Oliver (2000): Die Anwendung der Zielsetzungstheorie zur Mitarbeitermotivation und -steuerung. In: Bungard, Walter/Kohnke, Oliver (Hrsg.): Zielvereinbarungen erfolgreich umsetzen. Konzepte, Ideen und Praxisbeispiele auf Gruppen- und Organisationsebene. Wiesbaden: Gabler, S. 35-65.

Köller, Michaela (2009): Konstruktion und Implementierung von Schulprogrammen – Ein triangulativer Forschungsansatz. Frankfurt am Main: Peter Lang.

Kotthoff, Hans-Georg (2003): Bessere Schulen durch Evaluation? Internationale Erfahrungen. Münster: Waxmann.

Kotthoff, Hans-Georg/Böttcher, Wolfgang (2010): Neue Formen der „Schulin-spektion“: Wirkungshoffnungen und Wirksamkeit im Spiegel empirischer Bildungsforschung. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 295-325.

Krems, Burkhardt (2012): Zielvereinbarung. In: Online-Verwaltungslexikon. Verfügbar unter: [www.olev.de](http://www.olev.de); Version 2.13) [11.01.2012].

Kromrey, Helmut (2001): Evaluation – ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis. In: Sozialwissenschaften und Berufspraxis, 24. Jg., H. 2, S. 105-130.

Kromrey, Helmut (2002): Empirische Sozialforschung. Opladen: UTB (10. Aufl.).

Kruft, Alfred (2008): Ausbildung der Qualitätsprüfer und -prüferinnen. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLichterhand, S. 116-121.

Krüger, Heinz-Hermann/Pfaff, Nicolle (2008): Triangulation quantitativer und qualitativer Zugänge in der Schulforschung. In: Helsper, Werner/Böhme, Jeanette (Hrsg.): Handbuch der Schulforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 157-179 (2. Aufl.).

KSNW – Konferenz Schulaufsicht Nordrhein-Westfalen e.V. (2008): Länderbericht. Delegiertenversammlung 2008 in Speyer. Verfügbar unter: [http://www.konferenz-schulaufsicht-nrw.de/Jahresbericht\\_NRW/2008/2008.html](http://www.konferenz-schulaufsicht-nrw.de/Jahresbericht_NRW/2008/2008.html) [27.12.2011].

Kuckartz, Udo (2010): Einführung in die computergestützte Analyse qualitativer Daten. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kuckartz, Udo/Ebert, Thomas/Rädiker, Stefan/Stefer, Claus (2009): Evaluation online. Internetgestützte Befragung in der Praxis. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kuckartz, Udo/Grunenberg, Heiko (2010): Qualitative Daten computergestützt auswerten: Methoden, Techniken, Software. In: Friebertshäuser, Barbara/Langer, Antje/Prenzel, Annedore (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim und München: Juventa, S. 501-514 (3. Aufl.).

Kuckartz, Udo/Rädiker, Stefan (2010): Computergestützte Analyse (CAQDAS). In: Mey, Günter/Mruck, Katja (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 734-750.

Kussau, Jürgen/Brüsemeister, Thomas (2007): Educational Governance: Zur Analyse der Handlungskoordination im Mehrebenensystem Schule. In: Altrichter, Herbert/Brüsemeister, Thomas/Wissinger, Jochen (Hrsg.): Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 15-54.

Kussau, Jürgen/Brüsemeister, Thomas (2007): Governance, Schule und Politik. Zwischen Antagonismus und Kooperation. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lambrecht, Maike / Kotthoff, Hans Georg / Maag Merki, Katharina (2008): Taktieren oder Öffnen? Die Pilotphase Fremdevaluation in Baden-Württemberg zwischen Entwicklung und Kontrolle – eine mikropolitische Prozess- und Ergebnisanalyse. In: Böttcher, Wolfgang / Bos, Wilfried / Döbert, Hans (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 279-291.

Lambrecht, Maike / Kotthoff, Hans-Georg / Maag Merki, Katharina (2011): Rekontextualisierungen: Schulinspektionsprozesse als indirekte Kommunikation zwischen Schulleitung und Kollegium. In: Müller, Sabine / Pietsch, Marcus / Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 165-191.

Landesregierung NRW (2010): Antwort auf die Große Anfrage 40 der Fraktion der SPD. Drucksache 14/10639. Lage der Schulen in NRW. Verfügbar unter: [http://www.gew-nrw.de/uploads/tx\\_files/ltd\\_14\\_10639.pdf](http://www.gew-nrw.de/uploads/tx_files/ltd_14_10639.pdf) [28.12.2011].

Landwehr, Norbert (2011): Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In: Qesel, Carsten / Husfeldt, Vera / Landwehr, Norbert / Steiner, Peter (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep Verlag, S. 35-69.

Lange, Hermann (2008): Vom Messen zum Handeln: "empirische Wende" der Bildungspolitik? In: Recht der Jugend und des Bildungswesens, 56. Jg., H. 1, S. 7-15.

Lee, Thomas W. / Locke, Edwin A. / Latham, Gary P. (1989): Goal Setting Theory and Job Performance. In: Pervin, L. A. (Hrsg.): Goal Concepts in Personality and Social Psychology. Hillsdale, NJ: Erlbaum, S. 291-326.

Lennartz, Alice (2010): Unterrichtsbeobachtungen – Ein Instrument zur Qualitätsentwicklung. In: Fischer, Christian / Schilmöller, Reinhard (Hrsg.): Was ist guter Unterricht? Qualitätskriterien auf dem Prüfstand. Münster: Aschendorff, S. 160-168.

Leschinsky, Achim / Cortina, Kai S. (2008): Zur sozialen Einbettung bildungspolitischer Trends in der Bundesrepublik. In: Cortina, Kai S. / Baumert, Jürgen / Leschinsky, Achim / Mayer, Karl Ulrich / Trommer, Luitgard (Hrsg.): Das Bildungswesen in der Bundesrepublik Deutschland. Reinbek bei Hamburg: Rowohlt, S. 21-51.

Locke, Edwin A. / Latham, Gary, P. (2002): Building a Practically Useful Theory of Goal Setting and Task Motivation. A 35-Year Odyssey. *American Psychologist*, 57, S. 705-717.

Luginbuhl, Rob / Webbink, Dinand / Wolf, Inge de (2009): Do inspections improve primary school performance? In: *Educational Evaluation and Policy Analysis*, Vol. 31, No. 3, S. 221-237.

Maag Merki, Katharina (2010): Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 145-169.

Maag Merki, Katharina/Emmerich, Marcus (2011): Schulexterne Steuerungsinstrumente der Schulentwicklung. In: Altrichter, Herbert/Helm, Christoph (Hrsg.): Akteure und Instrumente der Schulentwicklung. Baltmannsweiler: Schneider Verlag Hohengehren, S. 151-168.

Maag Merki, Katharina/Holmeier, Monika (2008): Die Implementation zentraler Abiturprüfungen. Erste Ergebnisse zu den Effekten der Einführung auf das schulische Handeln der Lehrpersonen. In: Lankes, Eva Maria (Hrsg.): Pädagogische Professionalität als Gegenstand empirischer Forschung. Münster: Waxmann, S. 233-243.

Maier, Uwe (2008a): Vergleichsarbeiten im Vergleich – Akzeptanz und wahrgenommener Nutzen standardbasierter Leistungsmessungen in Baden-Württemberg und Thüringen. In: Zeitschrift für Erziehungswissenschaft 11, H. 3, S. 453-474.

Maier, Uwe (2008b): Rezeption und Nutzung von Vergleichsarbeiten aus der Perspektive von Lehrkräften. In: Zeitschrift für Pädagogik, 59 Jg., H. 1, S. 95-117.

Maier, Uwe (2009): Testen und dann? – Ergebnisse einer qualitativen Lehrerbefragung zur diagnostischen Funktion von Vergleichsarbeiten. In: Empirische Pädagogik, 23, H. 2, S. 191-207.

Maritzen, Norbert (2006): Eine Trendanalyse. Schulinspektion zwischen Aufsicht und Draufsicht. In: Buchen, Herbert/Horster, Leonhard/Rolff, Hans-Günter (2006): Schulinspektion und Schulleitung. Stuttgart: Raabe, S. 7-25.

Maritzen, Norbert (2008): Schulinspektionen. Zur Transformation von Governance-Strukturen im Schulwesen. In: Die Deutsche Schule, 100, H. 1, S. 85-96.

Maritzen, Norbert (2011): Externe Schulevaluation als Element des Systemmonitorings. In: Quesel, Carsten/Husfeldt, Vera/Landwehr, Norbert/Steiner, Peter (Hrsg.): Wirkungen und Wirksamkeit der externen Schulevaluation. Bern: hep, S. 71-92.

Matthews, Pam/Sammons, Peter (2004): Improvement through inspection: An evaluation of the impact of Ofsted's work. London: Ofsted. Verfügbar unter: <http://dera.ioe.ac.uk/4969/3/3696.pdf> [29.11.2011].

Maurer, Marcus/Jandura, Olaf (2009): Masse statt Klasse? Einige kritische Anmerkungen zu Repräsentativität und Validität von Online-Befragungen. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 61-73.

Mayring, Philipp (2002): Einführung in die qualitative Sozialforschung. Weinheim und Basel: Beltz.

Mayring, Philipp (2008): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim und Basel: Beltz.

Mayring, Philipp (2010): Qualitative Inhaltsanalyse. In: Mey, Günter/Mruck, Katja (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 601-613.

Mayring, Philipp/Brunner, Eva (2010): Qualitative Inhaltsanalyse. In: Friebertshäuser, Barbara/Langer, Antje/Prenzel, Annedore (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim und München: Juventa, S. 323-333 (3. Aufl.).

Medjedović, Irena (2010): Sekundäranalyse. In: Mey, Günter/Mruck, Katja (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 304-315.

Medjedović, Irena/Witzel, Andreas (2010): Wiederverwendung qualitativer Daten. Archivierung und Sekundärnutzung qualitativer Interviewtranskripte. Wiesbaden: VS Verlag für Sozialwissenschaften.

Merten, Klaus (1995): Inhaltsanalyse. Einführung in Theorie, Methode und Praxis. Opladen: Westdeutscher Verlag (2. Aufl.).

Meyer, Hilbert (2005): Was ist guter Unterricht? Berlin: Scriptor (2. Aufl.).

Meyer, Hilbert (2010): Merkmale guten Unterrichts – Ein Kriterienmix. In: Jürgens, Eiko/Standop, Jutta (Hrsg.): Was ist „guter“ Unterricht? Namhafte Expertinnen und Experten geben Antwort. Bad Heilbrunn: Klinkhardt, S. 159-174.

Meyer, Hilbert (2010): Was ist guter Unterricht? Berlin: Cornelsen Scriptor (7. Aufl.).

Middendorf, William (2010): Externe Schulevaluation des Staates als Herausforderung für das spezifische Profil freier Schulen. Möglichkeiten und Grenzen für Handlungsstrategien aus Trägersicht. In: Böttcher, Wolfgang/Dicke, Jan Niklas/Hogrebe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft. Münster: Waxmann, S. 270-281.

Mittag, Sandra (2006): Qualitätssicherung an Hochschulen. Eine Untersuchung zu den Folgen der Evaluation von Studium und Lehre. Münster: Waxmann.

Mohr, Ingola (2006): Analyse von Schulprogrammen. Eine Arbeit im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU). Münster: Waxmann.

Moritz, Wolfram von (2011): Qualitätsanalyse NRW an Evangelischen Schulen. In: Schulverwaltung, 22. Jg., H. 1, Köln: Wolters Kluwer, S. 12-13.

MSW NRW (2006): Qualitätstableau für die Qualitätsanalyse an Schulen in Nordrhein-Westfalen. Zugriff unter: [http://www.schulministerium.nrw.de/QA/Tableau/A\\_Grundlagen/1\\_Recht/index.html](http://www.schulministerium.nrw.de/QA/Tableau/A_Grundlagen/1_Recht/index.html) [23.01.2012].

MSW NRW (2008a): Qualitätstableau für die Qualitätsanalyse an Schulen in Nordrhein-Westfalen. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 183-194.

MSW NRW (2008b): Schulportfolio – Rahmendaten und Dokumente schulischer Arbeit (Datenerhebung zur Vorbereitung der Schulbesuche – Version für Sekundarschulen). In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 195-207.

MSW NRW (2008c): Beobachtungsbogen für Unterricht und Lernprozesse. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 208-212.

MSW NRW (2008d): Qualitätsbericht. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 213-222.

MSW NRW (2008e): Evaluation der Qualitätsanalyse. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 223-226.

MSW NRW (2009a): Qualitätsanalyse in Nordrhein-Westfalen. Impulse für die Weiterentwicklung von Schulen (Jahresbericht). Düsseldorf: MSW.

MSW NRW (2009b): Das Schulwesen in Nordrhein-Westfalen aus quantitativer Sicht 2008/09. Zugriff unter: [http://www.schulministerium.nrw.de/BP/Schulsystem/Statistik/2008\\_09/StatUebers.pdf](http://www.schulministerium.nrw.de/BP/Schulsystem/Statistik/2008_09/StatUebers.pdf) [04.10.2011].

MSW NRW (2011b): Bewertung. Zugriff unter: [http://www.schulministerium.nrw.de/QA/Tableau/C\\_Schulbesuch/5\\_Bewertung/index.html](http://www.schulministerium.nrw.de/QA/Tableau/C_Schulbesuch/5_Bewertung/index.html) [05.10.2011].

MSW NRW (2015a): Das Qualitätstableau NRW. Zugriff unter: <http://www.schulministerium.nrw.de/docs/Schulentwicklung/Qualitaetsanalyse/Tableau/index.html> [17.01.2015].

MSW NRW (2015b): Handreichung zum institutionellen Zielvereinbarungsprozess zwischen Schulen und Schulaufsicht in NRW. Zugriff unter: [https://www.schulministerium.nrw.de/docs/Schulentwicklung/Qualitaetsanalyse/Rechtliche-Grundlagen/5\\_Zielvereinbarung.pdf](https://www.schulministerium.nrw.de/docs/Schulentwicklung/Qualitaetsanalyse/Rechtliche-Grundlagen/5_Zielvereinbarung.pdf) [17.01.2015].

Müller, Sabine (2008): Ausgewählte Ergebnisse zur Qualitätsanalyse aus der Pilotphase 2006. In: Müller, Sabine / Dederling, Kathrin / Bos, Wilfried (Hrsg.): Schuli-sche Qualitätsanalyse in Nordrhein-Westfalen. Köln: Wolters Kluwer, S. 165-167.

Müller, Sabine (2008a): Einschätzungen von Schulen zur Qualitätsanalyse. Erste Ergebnisse der Schulrückmeldungen an die Qualitätsteams. In: SchulVewaltung NRW, Jg. 19, S. 134-136.

Müller, Sabine (2008b): Zielvereinbarungen als Steuerungsinstrument für Ent-wicklungsprozesse nach der Schulinspektion. In: Schulverwaltung, Jg. 9, S. 235-237.

Müller, Sabine (2009): Was halten Schulen von der Qualitätsanalyse? Ein Feed-back. In: Schule NRW, Amtsblatt des Ministeriums für Schule und Weiterbildung, Jg. 4, Düsseldorf, S. 172-174.

Müller, Sabine (2010): Erste Effekte von Schulinspektionen – eine Zwischenbi-lanz. In: Berkemeyer, Nils / Bos, Wilfried / Holtappels, Heinz Günter / McElvany, Ne-le (Hrsg.): Jahrbuch der Schulentwicklung Band 16. Weinheim: Basel, S. 289-308.

Müller, Sabine / Pietsch, Marcus (2011): Was wir messen, wenn wir Unterrichts-qualität messen. Inter-Beurteilerübereinstimmung und -reliabilität bei Unterrichtsbe-obachtungen im Rahmen von Schulinspektion. In: Müller, Sabine / Pietsch, Marcus / Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus em-pirischer Sicht. Münster: Waxmann, S. 33-55.

Muthén, Linda K. / Muthén, Bengt O. (1998-2010): Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén / Muthén. Verfügbar unter: <http://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v6.pdf> [06.02.2012].

NRW (2011): Schulgesetz. Verfügbar unter: [http://www.schulministerium.nrw.de/BP/Schulrecht/Gesetze/SchulG\\_Info/](http://www.schulministerium.nrw.de/BP/Schulrecht/Gesetze/SchulG_Info/) [19.10.2011].

Pant, Hans Anand / Vock, Miriam / Pöhlmann, Claudia / Köller, Olaf (2008): Eine modellbasierte Erfassung der Auseinandersetzung von Lehrkräften mit den länder-übergreifenden Bildungsstandards. In: Lankes, Eva-Maria (Hrsg.): Pädagogische Pro-fessionalität als Gegenstand empirischer Forschung. Münster: Waxmann, S. 245-260.

Patton, Michael Q. (2008): Utilization-focused evaluation. Thousand Oaks, CA: Sage (4. Aufl.).

Peek, Rainer (2002): Die Bedeutung vergleichender Schulleistungsmessungen für die Qualitätskontrolle und Qualitätsentwicklung von Schulen und Schulsystemen. In: Weinert, Franz E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim: Beltz (2. Aufl.), S. 323-335.

Peek, Rainer (2007): Interne Evaluation und einzelschulische Entwicklung – Spagat zwischen Mindeststandards und Machbarem. In: Buer, Jürgen van/Wagner, Cornelia (Hrsg.): Qualität von Schule. Ein kritisches Handbuch. Frankfurt am Main: Lang, S. 141-149.

Peek, Rainer/Dobbelstein, Peter (2006): Zielsetzung: Ergebnisorientierte Schul- und Unterrichtsentwicklung. In: Böttcher, Wolfgang/Holtappels, Hans Günter/Brohm, Michaela (Hrsg.): Evaluation im Bildungswesen. Eine Einführung in Grundlagen und Praxisbeispiele. Weinheim/München: Juventa, S. 177-193.

Peek, Rainer/Pallack, Andreas/Dobbelstein, Peter/Fleischer, Jens/Leutner, Detlev (2006): Lernstandserhebungen 2004 in Nordrhein-Westfalen – zentrale Testergebnisse und Perspektiven für die Schul- und Unterrichtsentwicklung. In: Eder, Ferdinand (Hrsg.): Qualität durch Standards? Münster: Waxmann, S. 219-233.

Perryman, Jane (2006): Panoptic performativity and school inspection regimes: disciplinary mechanisms and life under special measures. In: Journal of Education Policy, Vol. 21, No. 2, S. 147-161.

Perryman, Jane (2007): Inspection and emotion. Cambridge Journal of Education, Vol. 37, No. 2, S. 173-190.

Peuker, Ewald (2008): Die Bedeutung des Schulportfolios in der Qualitätsanalyse. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 103-109.

Pietsch, Marcus (2010): Evaluation von Unterrichtsstandards. In: Zeitschrift für Erziehungswissenschaft, 13, H. 1, S. 121-148.

Pietsch, Marcus (2011): Nutzung und Nützlichkeit der Schulinspektion Hamburg. Ergebnisse der Hamburger Schulinspektionsbefragung. Hamburg: Behörde für Schule und Berufsbildung. Institut für Bildungsmonitoring – Schulinspektion. Verfügbar unter: <http://www.schulinspektion.hamburg.de/index.php/file/download/1744> [16.03.2012].

Pietsch, Marcus/Schulze, Peter/Schnack, Jochen/Krause, Mareile (2011): Elaborierte Rückmeldungen zur Qualität von Unterricht. Über empirisch abgesicherte Bezugsnormen als Grundlage für die Weiterentwicklung von Unterricht und Schule. In: Müller, Sabine/Pietsch, Marcus/Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 193-216.

Pietsch, Marcus/Tosana, Simone (2008): Beurteilereffekte bei der Messung von Unterrichtsqualität. Das Multifacetten-Rasch-Modell und die Generalisierbarkeitstheorie als Methoden der Qualitätssicherung in der externen Evaluation von Schulen. In: Zeitschrift für Erziehungswissenschaft, H. 3, S. 430-452.

Porst, Rolf (2011): Fragebogen. Ein Arbeitsbuch. Wiesbaden: VS Verlag für Sozialwissenschaften (3. Aufl.).

Pötschke, Manuela (2009): Potentiale von Online-Befragungen: Erfahrungen aus der Hochschulforschung. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 75-89.

Recum, Hasso von (2003): Aspekte bildungspolitischer Steuerung. In: Döbert, Hans/Kopp, Botho von/Martini, Renate/Weiß, Manfred (Hrsg.): Bildung vor neuen Herausforderungen. Neuwied: Luchterhand, S. 102-110.

Reimers, Heino (2008): Die Trennung von Zielen und Maßnahmen. Zielvereinbarungen in der Praxis. In: Schulmanagement, 1, S. 21-24.

Reinecke, Jost (2005): Strukturgleichungsmodelle in den Sozialwissenschaften. München: Oldenbourg.

Reiners-Woch, Ingeborg (2002). Zielvereinbarungen. In: Rolff, Hans-Günter/Schmidt, Hans-Joachim (Hrsg.): Brennpunkt Schulleitung und Schulaufsicht. Konzepte und Anregungen für die Praxis. Neuwied: Luchterhand, S. 127-137.

Rheinberg, Falko (2006): Motivation. Stuttgart: Kohlhammer (6. Aufl.).

Rohrmann, Bernd (1979): Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. In: Zeitschrift für Sozialpsychologie, 9, S. 222-245.

Rolff, Hans-Günter (2002): Was bringt die vergleichende Leistungsmessung für die pädagogische Arbeit in Schulen? In: Weinert, Franz E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim: Beltz (2. Aufl.), S. 337-352.

Rolff, Hans-Günter/Schmidt, Hans-Joachim (Hrsg.) (2002): Brennpunkt Schulleitung und Schulaufsicht. Konzepte und Anregungen für die Praxis. Neuwied: Luchterhand.

Rosenbusch, Heinz S. (2002): Vom „private cold war“ zum „Freund in der Not“?! In: Rolff, Hans-Günter/Schmidt, Hans-Joachim (Hrsg.): Brennpunkt Schulleitung und Schulaufsicht. Konzepte und Anregungen für die Praxis. Neuwied: Luchterhand, S. 13-26.

Rosenthal, Leslie (2004): Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. In: Economics of Education Review, Vol. 23, No. 2, S. 143-151.

Rossi, P. H./Freeman, H. E. (1993): Evaluation: A systematic approach. Newbury Park, CA: Sage.

Rost, Detlef H. (2007): Interpretation und Bewertung pädagogisch-psychologischer Studien. Eine Einführung. Weinheim und Basel: Beltz (2. Aufl.).

Ruhwinkel, Walter/Sticka, Jana (2008): Erste Ergebnisse und Erfahrungen aus dem Dezernat 4Q der Bezirksregierung Münster. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 325-334.

Rürup, Matthias (2008): Typen der Schulinspektion in den deutschen Bundesländern. In: Die Deutsche Schule, 100, H. 4, S. 467-477.

Rürup, Matthias/Fuchs, Hans-Werner/Weishaupt, Horst (2010): Bildungsberichterstattung – Bildungsmonitoring. In: Altrichter, Herbert/Maag Merki, Katharina (Hrsg.): Handbuch Neue Steuerung im Schulsystem. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 377-401.

Sammons, Pamela (2006): Improving School an Raising Standards: the Impact of Educational Reforms in England. In: Eder, Ferdinand/Gastager, Angela/Hofmann, Franz (Hrsg.): Qualität durch Standards? Beiträge zum Schwerpunktthema der 67. Tagung der AEPF. Münster: Waxmann, S. 39-53.

Schäfers, Bernhard/Zapf, Wolfgang (2001): Handwörterbuch zur Gesellschaft Deutschlands. Bonn: Bundeszentrale für politische Bildung (2. Aufl.).

Scheerens, Jaap/Bosker, Roel J. (1997): The foundations of educational effectiveness. Oxford: Pergamon.

Scheerens, Jaap/Glas, Cees/Thomas, Sally M. (2003): Educational Evaluation, Assessment, and Monitoring. A Systemic Approach. Lisse: Swets/Zeitlinger B.V.

Schermelleh-Engel, Karin/Moosbrugger, Helfried/Müller, Hans (2003): Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. In: Methods of Psychological Research Online, Vol. 8, No. 2, S. 23-74.

Schmidt, Klaus-Helmut/Kleinbeck, Uwe (2006): Führen mit Zielvereinbarung. Göttingen: Hogrefe.

Schneewind, Julia (2007a): Wie Lehrkräfte mit Ergebnisrückmeldungen aus Schulleistungsstudien umgehen. Ergebnisse aus Befragungen von Berliner Grundschullehrerinnen. Berlin: Freie Universität. Verfügbar unter: [http://www.diss.fu-berlin.de/diss/receive/FUDISS\\_thesis\\_000000002819](http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000002819) [28.11.2008].

Schneewind, Julia (2007b): Erfahrungen mit Erlebnissrückmeldungen im Projekt BeLesen - Ergebnisse der Interviewstudie. In: Empirische Pädagogik, 21, H. 4, S. 368-382.

Schneewind, Julia/Kuper, Harm (2009): Rückmeldeformate und Verwendungsmöglichkeiten der Ergebnisse aus zentralen Lernstandserhebungen. In: Bohl, Thorsten/Kiper, Hanna (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 113-129.

Schnell, Herbert (2010): Funktionen und Aufgaben der Schulaufsicht. In: Bohl, Thorsten/Helsper, Werner/Holtappels, Heinz Günter/Schelle, Carla (Hrsg.): Handbuch Schulentwicklung. Bad Heilbrunn: Klinkhardt, S. 196-199.

Schnell, Rainer/Hill, Paul B./Esser, Elke (2008): Methoden der empirischen Sozialforschung. München und Wien: Oldenbourg (8. Aufl.).

Schouten, Rob (2008): Schulinspektion in den Niederlanden – Erfahrungen und Ergebnisse zu Verfahren, Instrumenten und Effekten. In: Böttcher, Wolfgang/Döbert, Hans/Holtappels, Hans Günter (Hrsg.): Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive. Münster: Waxmann, S. 235-239.

Schrader, Friedrich-Wilhelm/Helmke, Andreas (2003): Evaluation – und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: Schweizerische Zeitschrift für Bildungswissenschaften, 25, 1, S. 79-110.

Schrader, Friedrich-Wilhelm/Helmke, Andreas (2004): Von der Evaluation zur Innovation? Die Rezeptionsstudie WALZER: Ergebnisse der Lehrerbefragung. In: Empirische Pädagogik, 18 (1), S. 140-161.

Schreier, Margit/Odağ, Özen (2010): Mixed Methods. In: Mey, Günter/Mruck, Katja (Hrsg.): Handbuch Qualitative Forschung in der Psychologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 263-277.

Schründer-Lenzen, Agi (2010): Triangulation – ein Konzept zur Qualitätssicherung von Forschung. In: Friebertshäuser, Barbara/Langer, Antje/Prengel, Annedore (Hrsg.): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft. Weinheim und München: Juventa, S. 149-158 (3. Aufl.).

Shaw, Ian/Newton, David P./Aitkin, Mike/Darnell, Ross (2003): Do OFSTED inspections of schools make a difference to GCSE results? In: British Educational Research Journal, Vol. 29, No. 1, S. 63-75.

SICI (2015): Members. Verfügbar unter: <http://www.sici-inspectorates.eu/Members/Contacts-of-Inspectorates> [10.10.2015].

Sommer, Norbert (2011): Wie beurteilen schulische Gruppen die erlebte Schulinspektion? Ergebnisse einer Befragung. In: Müller, Sabine/Pietsch, Marcus/Bos, Wilfried (Hrsg.). Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 137-164.

Sommer, Norbert / Stöhr, Cora / Thomas, Diana (2010): Schulen mit „gravierenden Mängeln“. Situation in Niedersachsen und Einsatzmöglichkeiten der Schulentwicklungsberatung. In: Böttcher, Wolfgang / Dicke, Jan Nikolas / Hogrebe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft, Münster: Waxmann, S. 209-227.

Sparka, Andrea (2007): Schulautonomie in der Bewährung. Ein Vergleich: das Bundesland Nordrhein-Westfalen und die Niederlande. Verfügbar unter: [https://eldorado.tu-dortmund.de/bitstream/2003/25076/2/Dissertation\\_Andrea\\_Sparka.pdf](https://eldorado.tu-dortmund.de/bitstream/2003/25076/2/Dissertation_Andrea_Sparka.pdf) [15.04.2010].

Sparka, Andrea / Bruggen, Johan van (2009): Schulische Unterstützungssysteme in den Niederlanden. In: Die Deutsche Schule, 101, H. 2, S. 136-148.

Spichal, Dieter (2008): Ansprüche an einen Qualitätsbericht. In: Müller, Sabine / Dederig, Kathrin / Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 110-115.

Sprenger, Reinhard K. (2000): Das Sisyphos-Dilemma. In: Bungard, Walter / Kohnke, Oliver (Hrsg.): Zielvereinbarungen erfolgreich umsetzen. Konzepte, Ideen und Praxisbeispiele auf Gruppen- und Organisationsebene. Wiesbaden: Gabler, S. 101-112.

Stamm, Margrit (2003a): Evaluation und ihre Folgen für die Bildung. Eine unterschätzte pädagogische Herausforderung. Münster: Waxmann.

Stamm, Margrit (2003b): Evaluation im Spiegel ihrer Nutzung: Grande idée oder grande illusion des 21. Jahrhunderts? In: Zeitschrift für Evaluation, H. 2, S. 183-200.

Steins, Gisela (2009): Widerstand von Lehrern gegen Evaluationen aus psychologischer Sicht. In: Bohl, Thorsten / Kiper, Hanna (Hrsg.): Lernen aus Evaluationsergebnissen. Verbesserungen planen und implementieren. Bad Heilbrunn: Klinkhardt, S. 185-195.

Stockmann, Rainhard (2007): Einführung in die Evaluation. Eine praktische Handlungsanleitung. In: Ders. (Hrsg.): Handbuch zur Evaluation. Münster: Waxmann, S. 24-71.

Stryck, Tom (2000): Qualitätssicherung in der Geisterbahn. Was hat die Schulaufsicht mit Schulqualität zu tun? In: Zeitschrift für Pädagogik, 41. Beiheft, S. 111-125.

Terhart, Ewald (2000): Qualität und Qualitätssicherung im Schulsystem. Hintergründe – Konzepte – Probleme. In: Zeitschrift für Pädagogik, 48, H. 1, S. 809-829.

Terhart, Ewald (2002): Wie können die Ergebnisse von vergleichenden Leistungsstudien systematisch zur Qualitätsverbesserung in Schulen genutzt werden? In: Zeitschrift für Pädagogik, 48, S. 91-110.

Thielsch, Meinald T. / Weltzin, Simone (2009): Online-Befragung in der Praxis. In: Brandenburg, Torsten / Thielsch, Meinald T. (Hrsg.): Praxis der Wirtschaftspsychologie: Themen und Fallbeispiele für Studium und Praxis. Münster: Monstein und Vannerdat, S. 69-85.

Thürmann, Eike (2007): Entwicklung im Prozess. Schulinspektion in Nordrhein-Westfalen. In: Böttcher, Wolfgang / Kotthoff, Hans-Georg (Hrsg.): Schulinspektion: Evaluation, Rechenschaftslegung und Qualitätsentwicklung. Münster: Waxmann, S. 175-191.

Thym, Daniel (2009): Zielvereinbarungen im Schulrecht zwischen informeller Verwaltungspraxis und rechtlicher Steuerung. In: Recht der Jugend und des Bildungswesens, 57, H. 2, S. 278-290.

Tillmann, Klaus-Jürgen / Dederich, Kathrin / Kneuper, Daniel / Kuhlmann, Daniel / Nessel, Isa (2008): PISA als bildungspolitisches Ereignis. Fallstudien in vier Bundesländern. Wiesbaden: VS Verlag für Sozialwissenschaften.

Tosana, Simone / Lambrecht, Maike / Perels, Franziska / Bardowiecks, Stefan (2011): Die Protokollierung von Interviews im Rahmen von Schulinspektionen. In: Müller, Sabine / Pietsch, Marcus / Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 57-78.

Ulber, Daniela (2010): Veränderte Steuerung im Schulsystem. Evaluation der Zielvereinbarungsgespräche zwischen Schulleitung und Schulaufsicht in Berlin. In: Böttcher, Wolfgang / Dicke, Jan Nikolas / Högbe, Nina (Hrsg.): Evaluation, Bildung und Gesellschaft. Münster: Waxmann, S. 297-304.

Visscher, Adrie J. / Coe, Robert (2003): School Performance Feedback Systems: Conceptualisation, Analysis, and Reflection. School Effectiveness and School Improvement, 14, 3, S. 321-349.

Visscher, Adrie J. / Coe, Robert (Hrsg.) (2002): School improvement through performance feedback (Contexts of learning, Vol. 10). Lisse: Swets / Zeitlinger.

Vock, Miriam / Pant, Hans Anand (2008): Standardbasierte Aufgaben, Leistungsvergleiche und Schulinspektionen. Potentiale der Ländermaßnahmen für die schulische Qualitätssicherung. In: Weißeno, Georg (Hrsg.): Politikkompetenz. Was Unterricht zu leisten hat. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 310-330.

Voß, Werner / Khlavna, Veronika / Schöneck, Nadine M. (2008): Einführung in die Datenanalyse und Datenmanagement mit SPSS. Skript. Bochum.

Weerts, Fré (2007): Das niederländische Inspektionsmodell. Selbstevaluation und externe Evaluation in Proportion. In: Schönig, Wolfgang (Hrsg.): Spuren der Schulevaluation. Zur Bedeutung und Wirksamkeit von Evaluationskonzepten im Schulalltag. Bad Heilbrunn: Klinkhardt, S. 42-53.

Weitzel, Ulrike (2002): Berufsbild und Rollenverständnis der Schulaufsicht. In: Rolff, Hans-Günter/Schmidt, Hans-Joachim (Hrsg.): Brennpunkt Schulleitung und Schulaufsicht. Konzepte und Anregungen für die Praxis. Neuwied: Luchterhand, S. 323-335.

Welker, Martin/Matzat, Uwe (2009): Online-Forschung: Gegenstände, Entwicklung, Institutionalisierung und Ausdifferenzierung eines neuen Forschungszweiges. In: Jakob, Nikolaus/Schoen, Harald/Zerback, Thomas (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 33-47.

Wendt, Heike/Bos, Wilfried (2011): Indikatoren zur Kontextuierung von Inspektionsergebnissen. In: Müller, Sabine/Pietsch, Marcus/Bos, Wilfried (Hrsg.): Schulinspektion in Deutschland. Eine Zwischenbilanz aus empirischer Sicht. Münster: Waxmann, S. 217-236.

Wendt, Peter/Klein, Werner (2007): Die Inspektion war da – was jetzt? Zielvereinbarungen als Instrument der Schulentwicklung. In: Journal für Schulentwicklung, 3, S. 25-31.

Wey, Klaus-Georg (2008): Normierung im Rahmen der Qualitätsanalyse. In: Müller, Sabine/Dedering, Kathrin/Bos, Wilfried (Hrsg.): Schulische Qualitätsanalyse in Nordrhein-Westfalen. Konzepte, erste Erfahrungen, Perspektiven. Köln: LinkLuchterhand, S. 70-76.

Wilcox, Brian/Gray, John (1996): *Inspecting schools: holding schools to account and helping them to improve*. Buckingham: Open university press.

Wirth, Heike/Müller, Walter (2006): Mikrodaten der amtlichen Statistik in der empirischen Sozialforschung. In: Diekmann, Andreas (Hrsg.): Methoden der Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 93-127.

Wolf, Inge de/Janssens, Frans J. G. (2007): Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, Vol. 33, No. 3, S. 379-396.

Wolff, Stephan (2009): Dokumenten- und Aktenanalyse. In: Flick, Uwe/Kardorff, Ernst von/Steinke, Ines (Hrsg.): *Qualitative Forschung: Ein Handbuch*. Reinbek: Rowohlt, S. 502-513 (7. Aufl.).

Wottawa, Heinrich (2001): Qualitätsmanagement durch Zielvereinbarung. In: Spiel, Christiane (Hrsg.): *Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck*. Münster: Waxmann, S. 151-164.

Wottawa, Heinrich (2006): Evaluation. In: Rost, Detlef H. (Hrsg.): *Handwörterbuch pädagogische Psychologie*. Weinheim: Beltz/PVU, S. 162-168.

Wottawa, Heinrich / Thierau, Heike (2003): Lehrbuch Evaluation. Bern: Huber (3. Aufl.).

Wurster, Sebastian / Gärtner, Holger (2010): Schulen im Umgang mit Schulinspektion und deren Ergebnissen: Schultypen im Vergleich. Poster präsentiert auf dem DGfE-Kongress Bildung in der Demokratie. Verfügbar unter: [http://www.medienpaed.fb02.uni-mainz.de/dgfe2010/index.php?option=com\\_content&task=view&id=94&Itemid=999999](http://www.medienpaed.fb02.uni-mainz.de/dgfe2010/index.php?option=com_content&task=view&id=94&Itemid=999999) [24.09.2010].

Zerback, Thomas / Schoen, Harald / Jakob, Nikolaus / Schlereth, Stefanie (2009): Zehn Jahre Sozialforschung mit dem Internet – eine Analyse zur Nutzung von Online-Befragungen in den Sozialwissenschaften. In: Jakob, Nikolaus / Schoen, Harald / Zerback, Thomas (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 15-31.

Zlatkin-Troitschanskaia, Olga (2006): Steuerbarkeit von Bildungssystemen mittels politischer Reformstrategien. Frankfurt am Main: Peter Lang.

Zlatkin-Troitschanskaia, Olga (2008): Erweiterte Schulautonomie – Segen oder Fluch? Chancen, Risiken und (Neben-)Wirkungen der teilautonomen Schule. In: Rolff, Hans-Günter u.a. (Hrsg.): Schulleitung und Schulentwicklung. Loseblattsammlung. Berlin: Raabe. Verfügbar unter: [http://www.lfi.bremerhaven.de/fachartikel/08-11\\_zlatkin\\_erweiterte\\_schulautonomie.pdf](http://www.lfi.bremerhaven.de/fachartikel/08-11_zlatkin_erweiterte_schulautonomie.pdf) [06.11.2009].

# Rezeption und Nutzung von Ergebnissen aus Schulinspektionen

Jana Kleine

Die Frage nach den Wirkungen von Evaluationen im Schulsystem gerät seit einigen Jahren immer stärker in den Fokus der Bildungsforschung. Im Zuge der sog. Neuen Steuerung wurde eine Reihe von Reformmaßnahmen der internen und externen Evaluation implementiert, von denen sich die Schulverwaltung Impulse für die Schul- und Unterrichtsentwicklung in der Einzelschule verspricht. Die Schulinspektion wurde in den Bundesländern als externe Evaluation eingeführt und gilt als ein solches Reforminstrument, dessen Potential als Impulsgeber für einzelschulische Qualitätsentwicklung jedoch weitgehend unbekannt ist. Vor dem Hintergrund der identifizierten Forschungsdefizite wird in der Arbeit anhand verschiedener empirischer Datenquellen und Methoden in einem triangulativen Zugang untersucht, wie die Ergebnisse der Qualitätsanalyse Nordrhein-Westfalen rezipiert und – in Form von Zielvereinbarungen zwischen Einzelschule und Schulaufsicht – für die Weiterentwicklung der einzelnen Schule genutzt werden.

