

## 2. Rezeptionsstudien

Lehrkräfte werden als wichtige Akteure in der Praxis vor Ort verstanden (Holtappels 1995, Rolff 1991, Türk 1989) und sind zunehmend in den Evaluationsprozess einbezogen, indem sie Ergebnisrückmeldungen erhalten. Sie müssen die Entwicklungen tragen und umsetzen. Um zu verstehen, was in den Schulen mit Ergebnisrückmeldungen geschieht und um zu klären, ob die Annahme, dass aufgrund von wissenschaftlicher Evaluation Probleme definiert und entsprechende Bereiche verbessert werden können (Fend 1980, Stamm 2003) richtig ist, ist eine junge Forschungsrichtung und eine Untergruppe von Evaluationsstudien entstanden: Rezeptionsstudien wie die hier vorliegende. Hierbei werden Rückmeldung (Bereitstellung von Informationen<sup>6</sup> für die Praxis durch Leistungsmessung), Rezeption (Rückgriff auf Informationen und die Umwandlung in Evaluationswissen durch die Praxis) und die Verwendung dieser Informationen untersucht.<sup>7</sup>

Jede Evaluationsstudie ist zwangsläufig mit dem Problem der Anschlussfähigkeit ihrer Ergebnisse zu den Kontexten praktischen Handelns und Entscheidens konfrontiert (Kuper & Schneewind 2006). Die zentrale Frage ist, ob durch die Bereitstellung der Information die Qualität verbessert werden kann und welche Aspekte auf die Verwendung Einfluss nehmen. Bisher gibt es wenig Forschung zur Frage, ob Lehrkräfte ihr Unterrichtsverhalten aufgrund von Rückmeldungen verändern (Tymms 1995, 127). Die positive Wirkung von Rückmeldungen wurde zwar bereits in Organisationen und auch bei Lehrern nachgewiesen (Coe 2002), in Deutschland verwendete Rückmeldesysteme wurden aber in Hinblick auf Gestaltung, Akzeptanz, Nutzen und Interpretation erst in den letzten Jahren verstärkt untersucht (Rolff 2002a, 98). Internationale Vergleiche zeigen, dass es in Deutschland in diesem Forschungsfeld deutlichen Nachholbedarf gibt (van Ackeren 2003).

Insbesondere in den Niederlanden, Großbritannien und Australien gibt es langjährige Erfahrungen mit Schulleistungsstudien und Rückmeldungen (Bangert-Drowns et al. 1991, Black & William 1998, Butler & Winne 1995,

---

<sup>6</sup> Abgrenzung von Daten und Information (vgl. Kuper 2004 & Willke 2001): Daten sind die kleinsten, nicht weiter zu verkleinernde Einheiten. Sie sind die Beobachtungen (Testergebnis, via Kreuz im Fragebogen geäußerte Einstellung, Beobachtungsprotokoll Unterricht) denen durch anschließende Kategorisierung Sinn zugewiesen wird (richtig/falsch, wenig/viel Zustimmung, Meldehäufigkeit/ Kommunikationsstruktur): Informationen sind verarbeitete Daten. Daten werden interpretiert und kategorisiert. Dies geschieht in Abhängigkeit des Relevanzkontextes der Person oder Institution und Daten werden dadurch zu Information. Wissen ist die Einbettung der Information in den Erfahrungskontexten und das Handeln der Person.

<sup>7</sup> Die verschiedenen Gegenstände von Evaluation, die jeweils unterschiedlichen Zielgruppen und Interessen von Evaluationen werden von Westermann (2004) ausführlich erläutert.

Caldas & Bangston 1999, Dihoff et al. 2003, Fitz-Gibbon 2002, Hancock, Thurmann & Hubbard 1995, Kulhavy & Stock 1989, Rowe 2005). Für Deutschland stellen Helmke & Schrader fest, dass

*„noch wenig darüber bekannt (ist), welchen Wert solche Informationen für die Lehrkräfte tatsächlich haben, wie man Rückmeldungen zur Erzielung höchstmöglicher Verständlichkeit und Brauchbarkeit gestalten müsste, wie Lehrkräfte damit umgehen und welche Wirkungen sie zeigen. ... Zwischen dem Erhalt einer solchen Leistungsrückmeldung, einer dadurch veranlassten Verbesserung des Unterrichts und letztendlich effizienterem Lernen und besseren Leistungen liegt ein weiter Weg, der noch gezielter erforscht werden muss: Erforderlich ist unter anderem, dass die Rückmeldungen überhaupt zur Kenntnis genommen, verstanden und genauer analysiert werden.“*

(Schrader & Helmke 2002, 56ff)

Schrader und Helmke machen deutlich, dass sich hier ein neues Forschungsfeld ergibt. Eine Übersicht über Leistungsstudien, die Rückmeldungen generiert haben, stellt die Tabelle 1 auf den folgenden Seiten dar. Hier finden sich auch Hinweise zu den jeweiligen Rezeptionsstudien, sofern welche durchgeführt wurden. Einige der Untersuchungen, die in Deutschland zur Rezeption von Ergebnismeldungen durchgeführt wurden, möchte ich hier kurz darstellen, verbunden mit der Darstellung der Studie, in dessen Rahmen die Ergebnismeldungen vorgenommen wurden. Es wurden diese Rezeptionsstudien ausgewählt, weil sie einen Überblick über den aktuellen Stand der Forschung in *Deutschland* vermitteln. Es handelt sich bei den ausgewählten Studien sowohl um in Deutschland durchgeführte bundeslandübergreifende (VERA) oder bundeslandspezifische Studien (Markus/Walzer, QuaSUM, Kompetenztests) als auch um zwei Beispiele aus dem Ausland: Die Rezeption von Check 5 in der Schweiz und die Rezeption von in Großbritannien genutzten Rückmeldungen. Die Studien werden nach folgender Struktur kurz erläutert:

- 1) Studie in der die Daten erhoben wurden,
- 2) Form der Ergebnismeldung,
- 3) Rezeptionsstudie.

Da eine ausführliche Darstellung der Untersuchungen den Rahmen dieser Arbeit sprengt, verweise ich für weitergehendes Interesse auf die im Text angegebenen Quellen.

**Tabelle 1: Übersicht über verschiedene Schulleistungsstudien und Rezeptionsstudien**

	Name	Funktion	Zeitraum Erhebung		Klassen-/Altersstufe	Schulleistung	Kontextvariablen	Rückmeldeform	Rezeptionsstudie
Internationale Studien	TIMSS Third International Mathematics and Science Study	Bildungssystemmonitoring	1994, 1995	bundesweite Stichprobe	7., 8. & 12. bzw. 13. Klasse	Mathematik, Naturwissenschaft	schulische und außerschulische Kontextbedingungen	Nur allgemeine Informationen.	Kohler, B. (2005). Rezeption internationaler Schulleistungsstudien. Münster: Waxmann.
	PISA Programme for International Student Assessment	Bildungssystemmonitoring	a) 2000, b) 2003, c) 2006	bundesweite Stichprobe	15-Jährige	a) Lesen, b) Mathe, c) Naturwissenschaft d) Cross-Curricular Competencies (Lernstrategien, motivationale Aspekte, Kooperationsfähigkeit)	sozio-ökonomische Hintergrundmerkmale; Variablen zu schulischen Lernbedingungen	Broschüren mit schulspezifischen Informationen (Kombination von Daten zu schulischen Kontext-, Prozess- und Wirkungsvariablen)	Watermann, R. & Stanat, P. (2004): Schülerrückmeldungen in PISA 2000. Sozialnorm- und kriteriumsorientierte Rückmeldeverfahren. In: Empirische Pädagogik, 18, 40-61.
	PIRLS/IGLU Die Internationale-Grundschul-Lese-Untersuchung	Bildungssystemmonitoring	2001	bundesweite Stichprobe	9. bis 10. Jährige	Lesen	Schüler- und Lehrerfragebögen zur schulischen und häuslichen Situation	Freiwillige Rückmeldung an die beteiligten Lehrpersonen	Schwippert, Knut: Leistungsrückmeldungen an Grundschulen im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU). In: Empirische Pädagogik, 2004, 18(1), 62-81.
Bundesland-übergreifende Studien (Deutschland)	VERA Vergleichsarbeiten (1*	Unterrichtsentwicklung, Diagnostik, Qualitätssicherung	jährlich seit 2003 bzw. 2004	flächendeckend Berlin, Brandenburg, Bremen, NRW, M.-Vorpommern, Rheinland-Pfalz, Schleswig-Holstein	Ende der 3. Klasse oder Anfang der 4. Klasse	Deutsch, Mathe	Lehrerfragebogen: relevante Merkmale des Klassen- und Schulkontextes; SchülerFB: Geschlecht, Sprachhintergrund, Zeugnisnote	online, Klassenspezifische Informationen, Fähigkeitsniveaus, Fehlerprofile,	Groß Ophoff, J., Hosenfeld, I. & Koch, U. (2006). Ergebnissrückmeldungen und ihre Rezeption im Projekt VERA. In: Kuper, H. & Schneewind, J. (Hrsg.), Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster: Waxmann, 19-40.
	DESI Deutsch Englisch Schülerleistung International (2*	Bildungssystemmonitoring	2001	bundesweite Stichprobe	15. Jährige bzw. 9. Klasse	Englisch, Deutsch	individuelle Schülermerkmale; Merkmale des Unterrichts; Selbstkonzepte; Unterrichtsziele; Lernverhalten, Klassenklima, Lernstrategien; Schülererebene: Geschlecht, Migrationsstatus, Bildungsstatus der Eltern; Klassenebene: Migrantenanteil, Bilingualität, durchschnittliche Bildungsnähe der Eltern, Schulform und Bundesland	Ja, nicht bekannt	von der Gathen, J. (2006). Grenzen und Möglichkeiten der innerschulischen Rezeption von Rückmeldungen aus Large-Scale-Assessments. In: Kuper, H. & Schneewind, J. (Hrsg.), Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster: Waxmann. 77-88 sowie Helmke, A.; Goebel, K.; Hosenfeld, I.; Schrader, F.-W.; Vo, T.; Wagener, W. (2003): Zur Rolle des Unterrichts im Projekt DESI. In: Empirische Pädagogik 17(3), 396-411.
Bundesland-spezifische Studien (Deutschland)	QuaSum Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik Brandenburg	Schulmonitoring, Diagnostik	1999	flächendeckend Brandenburg	5. & 9. Klassen	Mathematik	schulische und außerschulische Bedingungen durch Schüler-, Lehrer-, Eltern-, Schulleiter FB's	Klassen/Kursspezifische und Landesweite Ergebnisse, Schulformvergleich	Peek, R. & Nilshon, I. (2004) Schülerrückmeldungen von Schulleistungsstudien am Beispiel des quasum-Projektes. Zwei Untersuchungen zur Wirksamkeit. Ministerium für Bildung, Jugend und Sport Land Brandenburg. Potsdam
	MARKUS Mathematik – Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontexte	Schulmonitoring, Diagnostik	2000	flächendeckend Rheinland-Pfalz	8. Klassen	Mathematik	Befragung von Schülern, Lehrern, Schulleitern zu schulischem und außerschulischem Kontext	Klassenspezifische Ergebnisse	Schrader, F.-W., Helmke, A. (2003). Evaluation – und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. Schweizerische Zeitschrift für Bildungswissenschaften 25 (1), 79 – 110
	LAU Lern-Ausgangslagen-Untersuchung Hamburg	Bildungssystemmonitoring	1996, 1998, 2000, 2005	flächendeckend, längsschnittlich	5., 7., 9., 13.Klasse	Deutsch, Mathe, 1. Fremdsprache	schul- und unterrichtsbezogene Einstellungen	Klassenbezogene Rückmeldungen in absoluten Zahlen	Ja, Leifadengestützte Interviews (seit 2005 jährlich)
	BeLesen Berliner Längsschnittstudie zur Lesekompetenzentwicklung bei Grundschulkindern (3*	Schulevaluation, Diagnostik	2002 - 2006	Stichprobe	1. bis 4. Klasse	Deutsch, Mathe, Sachkunde	Kognitive Leistungsfähigkeit der Schülerschaft, Unterrichtsgestaltung per Lehrerfragebogen	Klassen- und schülerbezogene Rückmeldung	Schneewind

	Name	Funktion	Zeitraum Erhebung	Stichprobe	Klassen-/Altersstufe	Schulleistung	Kontextvariablen	Rückmeldeform	Rezeptionsstudie
Bundesland-spezifische Studien (Deutschland)	KESS 4 Kompetenzen und Einstellungen von Schülerinnen und Schülern, Klasse 4 (4*)	Lernstands-erhebung, Schulmonitoring	seit 2003	flächendeckend Hamburg, wird als Längsschnitt weitergeführt	Beginn in der 4. Klasse	Deutsch, Mathe, Englisch, Sachkunde	Befragung von Schülern, Eltern, Lehrern, Schulleitern zu unterrichtsrelevanten und individuellen Bedingungen schulischen Lernens, schulischer Leistung; Bewertung des Unterrichts (Deutsch, Mathe, Sachunterricht) durch Schülerfragebogen; S: Freizeitverhalten von Schülern, erlebtes Erziehungsverhalten; E: kulturelles Kapital, Bildungshintergrund; L: Unterrichtsaktivitäten, Kooperation, Einstellungen; SL: Ausstattung, Kooperation, Entwicklung, Umsetzung amt. Vorgaben, Schulstandort	klassenbezogene Rückmeldung, kriteriums- und sozialnormorientiert, Lernstand und Kontextvariablen	Es liegen keine Angaben vor
	Lernstand 9 NRW (5*)	Schulmonitoring, Diagnostik	seit 2004	flächendeckend	9. Klasse	Deutsch, Mathe, Englisch	Schülermerkmale, Elternmerkmale (kulturelles Kapital, Bildungsnähe, sozialer Status, Ausbildung, Beruf, Migrationsstatus), Unterrichtsmerkmale, schulische Rahmenbedingungen	Klassen/Kurspezifische und Landesweite Ergebnisse// Über Ergebnisse u. notwendige Konsequenzen wird in Schulkonferenz berichtet	Ja, Fallstudien Universität Wuppertal Viola Hartung
	Orientierungsarbeiten in der Grundschule Bayern	Schulevaluation, Diagnostik, Bildungs-monitoring	seit SJ 02/03	flächendeckend	2. und 3. Klasse	Deutsch (Lesen in 2 und 3, richtig schreiben nur in 3), Mathe	sozioökonomischer Hintergrund der Schüler nur bei Pilotierungserhebung	Mittelwerte der Klassen und der Schule; I Schulen erhalten Durchschnittsergebnisse spätestens in der letzten Schulwoche, gleichzeitig Veröffentlichung im Internet; keine Aufbereitung für verschiedene Gruppen; Individualdaten für folgende Rückmeldung geplant	Es liegen keine Angaben vor
	Vergleichsarbeiten in der Grundschule Berlin	Vergleichsarbeiten, Evaluation, Individualdiagnostik	seit 2003	flächendeckend	2. Klasse	Deutsch, Mathe	sozioökonomischer Hintergrund der Schüler	schriftlich auf Anfrage und im Internet; Ergebnisse Gesamtstichprobe und Hinweise zu Ziel und Durchführung der Studie, Mittelwerte, Individualdaten	Es liegen keine Angaben vor
	Orientierungsarbeiten in Sachsen	Landesweite Vergleichsarbeiten	seit 2003/04	flächendeckend	3., 6., 8. Klasse	Deutsch, Mathe, Englisch (nur Klasse 6 u. 8)	Lehrerfragebogen: anvisierte Kompetenzstufen, Konsequenzen für Unterrichtspraxis, Vorschläge für weitere Arbeiten	landesweite Auswertung im Internet	Es liegen keine Angaben vor
	Kompetenztests Thüringen	Landesweite Vergleichsarbeiten	2001/02 Leistungstests Deutsch & Mathe, seit 2002/03 landesweite Vollerhebung mit Rückmeldung	flächendeckend	3. und 6. Klasse	Deutsch, Mathematik	Geschlecht, Muttersprache, diagnostizierter sonderpäd. Förderbedarf	Rückmeldung per Internet klassenbezogene Rückmeldung, Punktwerte	Nachtigall, C. et al. (2005). Landesbericht Thüringer Kompetenztests <a href="http://www.kompetenztest.de/downloads1.html">http://www.kompetenztest.de/downloads1.html</a>

(1\* 50% der Testaufgaben werden von den Lehrkräften aus einem Pool ausgewählt.

(2\* Im Fach Englisch ist an einigen Schulen zusätzlich eine Videostudie durchgeführt worden. Ziele: Beschreibung Unterrichtspraxis; Erklärung des Lernerfolgs durch Merkmale des Unterrichts, Lehrer-Schüler-Interaktion, Klassenführung, mündliche Kommunikationsfähigkeit der SchülerInnen

(3\* Längsschnittliche Betrachtung der Schülerinnen und Schüler

(4\* Rückmeldung ging auch an die Schulaufsicht.

(5\* Schulaufsicht erhält nach schulinterner Aufarbeitung Bericht über Ergebnisse und schulintern vereinbarte Konsequenzen

Neben den in der Rubrik Rezeptionsstudie angegebenen Quellen, wurde genutzt: Ackeren, I. van & Klemm, K. (2000). TIMSS, PISA, LAU, MARKUS und so weiter – Ein aktueller Überblick über Typen und Varianten von Schulleistungstudien. PÄDAGOGIK 12/00, 10-15

## **2.1 Vergleichsarbeiten in der Grundschule (VERA)**

### *1) Studie, in der die zurückgemeldeten Daten erhoben wurden*

Im Projekt VERA (vgl. Groß Ophoff 2006) werden seit 2004 jährlich in sieben Bundesländern (Berlin, Brandenburg, Bremen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz und Schleswig-Holstein) zu Beginn der 4. Klassenstufe flächendeckend alle Grundschulkin-der in den Fächern Mathematik und Deutsch getestet. Untersucht wird der Lernstand. Die jeweiligen Bildungsministerien sind Auftraggeber der Studie. Die Ergebnisse der Vergleichsarbeiten dienen der externen Evaluation. Die so generierten Daten sollen unter anderem Prozesse der Unterrichtsentwicklung, die Verbesserung der diagnostischen Kompetenzen, die Nutzung des Internets für die Qualitätssicherung sowie die Sicherung der Bildungsstandards unterstützen. Ein Teil der in den Tests gestellten Aufgaben wird zentral vorgegeben, ein anderer Teil wird von den teilnehmenden Lehrpersonen selbst zusammengestellt. Die Tests werden von den Lehrpersonen erhoben und eingegeben. Die Auswertung erfolgt zentral.

### *2) Form der Ergebnisrückmeldung*

Die beteiligten Lehrpersonen erhalten vom Auswertungsteam in Landau Ergebnisrückmeldungen in zwei Wellen (vgl. Groß Ophoff 2006): In der ersten Welle werden ihnen die von den einzelnen Schülerinnen und Schülern erreichten Fähigkeitsniveaus in Prozent mitgeteilt, im Vergleich zur Klassen- und Schulebene. Sie erhalten in Form von Abbildungen eine Analyse der Fehlerhäufigkeit, die in Beziehung gesetzt werden zur Fehlerhäufigkeit der Normstichprobe. Darüber hinaus erhalten sie die Analyse der diagnostischen Kompetenz, das heißt eine Gegenüberstellung der eingeschätzten und tatsächlichen Lösungshäufigkeit. In der zweiten Welle erhalten die Lehrpersonen Zusatzinformationen: Die Resultate (Fähigkeitsniveaus) werden im Vergleich zum eigenen Bundesland dargestellt sowie im Vergleich zu einer länderunspezifischen Kontextgruppe. Außerdem werden die Lösungshäufigkeiten der Zentralaufgaben mitgeteilt. Neben den Ergebnissen werden den Lehrkräften Informationen zum Umgang mit den Resultaten an die Hand gegeben. Die Rückmeldung wird von der Projektleitung von VERA explizit als Unterstützungsinstrument für Prozesse der Schul- und Unterrichtsentwicklung angesehen. Es wird von Seiten des Projektteams daher viel Wert darauf gelegt, für die Lehrpersonen verständliche und nutzbare Informationen bereit zu stellen, da davon ausgegangen wird, dass die Akzeptanz der Resultate eine Voraussetzung für die Nutzung der Information für Unterrichtsentwicklung ist.

### 3) Rezeptionsstudie

Alle beteiligten Lehrpersonen (N = 1248) wurden im geschützten Bereich des Internetportals von VERA dazu aufgefordert, einen internetbasierten Fragebogen in Hinblick auf die Rahmenbedingungen der Studie zu beantworten. Es konnte eine Vielzahl von Aussagen zu schulischen und externen Rahmenbedingungen gemacht werden sowie zu den aufgrund der Resultate eingeleiteten Maßnahmen (vgl. Groß Ophoff 2006). Relevant für die vorliegende Arbeit sind jedoch folgende Ergebnisse: Die Lehrkräfte empfinden einen Mangel an organisatorischen und zeitlichen Ressourcen, um die notwendigen Fördermaßnahmen umsetzen zu können. Die Arbeit, die durch die Vergleichsarbeiten entsteht wird als zu zeitintensiv empfunden. Die Reflektion über die Resultate und die daraus abgeleiteten Veränderungen des Unterrichts, also die eigentliche Unterrichtsentwicklung, werden selten als Maßnahmen benannt. Groß Ophoff et al. (2006, 36) führen aus,

*„dass die eigentlichen Elemente der Unterrichtsentwicklung, nämlich die auf empirischen Ergebnissen basierte Reflektion des eigenen Unterrichts sowie ggf. daraus resultierende Änderungen, eher selten angeführt werden.“*

Die Maßnahmen, die von den befragten Lehrpersonen erläutert werden, sind demnach eher gestalterisch-handelnd und weniger evaluierend-reflektiv, die Auseinandersetzung mit dem Unterricht wird wenig thematisiert.

## **2.2 Wirkungsanalyse der Leistungsevaluation: Zielerreichung, Ertrag für die Bildungsqualität der Schule und die Rückmeldung von Ergebnissen (WALZER)**

### *1) Studie, in der die zurückgemeldeten Daten erhoben wurden*

Die Rezeptionsstudie WALZER wurde im Anschluss an die Untersuchung MARKUS (**M**athematik-Gesamterhebung **R**heinland-Pfalz: **K**ompetenzen, **U**nterrichtsmerkmale, **S**chulkontext, Helmke & Jäger 2002) vorgenommen. Alle Schülerinnen und Schüler der 8. Klassen nahmen im Mai 2000 an einem Mathematik-Test teil und wurden zu Schulbiographie, Selbsteinschätzung der Lernstrategien, Motivation und Merkmalen des Unterrichts befragt. Insgesamt nahmen mehr als 37000 Jugendliche teil.

### *2) Form der Ergebnismeldung*

Nach Auswertung wurden die Ergebnisse des Leistungstests automatisch an die beteiligten Lehrpersonen und Schulleitungen zurückgemeldet. Die Ergebnisse der Schülerbefragung konnten extra angefordert werden. Zurückgemeldet (Helmke et al. 2000) wurde erstens das Kontextprofil der Klasse in Form eines Balkendiagramms, zweitens erhielten die Lehrpersonen die Klassener-

gebnisse des Leistungstests in Mathematik, aufgeschlüsselt nach den inhaltlichen Teilbereichen des Tests und im Vergleich zu bereinigten Werten, die das Kontextprofil der Klasse berücksichtigten.

### 3) *Rezeptionsstudie*

Alle an MARKUS beteiligten Lehrpersonen<sup>8</sup> wurden angeschrieben und gebeten an der Untersuchung WALZER teilzunehmen. Bedauerlicherweise haben nur 81 der 1876 Personen die entsprechenden Fragebögen zurückgesandt. Lehrpersonen, die länger als 20 Jahre im Dienst sind, waren in der Stichprobe deutlich überrepräsentiert. Erfragt wurde in der Studie die Einschätzung des schulischen Qualitätsmanagements, die Wichtigkeit von Fortbildungsmaßnahmen, die Unterrichtsgestaltung, Schulklima, die Selbsteinschätzung der Selbstwirksamkeitserwartung sowie die Einstellung gegenüber externer Evaluation. Es wird davon ausgegangen, dass die kleine Stichprobe die aktiven, innovationsfreudigen oder aber pflichtbewussten Lehrpersonen abbildet. Zentral für die vorliegende Arbeit sind folgende Ergebnisse:

Erstens zeigte sich (Helmke 2004, 165), dass

*„je positiver Lehrkräfte externer Evaluation gegenüber stehen und je höher ihre Selbstwirksamkeit ist, umso größer ist ihre Bereitschaft, den eigenen Unterricht zu verändern.“*

Zweitens hat sich gezeigt, dass die Mehrheit berichtet, keine Maßnahmen auf Grundlage der zurückgemeldeten Informationen ergriffen zu haben, obwohl 40% der Befragten einen hohen Fortbildungsbedarf angegeben haben.

Drittens zeigte sich, dass je mehr die Lehrpersonen Unterstützung durch die Schulleitung erfahren, desto höher ist ihre Bereitschaft, individuelle Veränderungen bei der Unterrichtsgestaltung zu erwägen.

## **2.3 Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM)**

### *1) Studie, in der die zurückgemeldeten Daten erhoben wurden*

In der im Mai 1999 in Brandenburg durchgeführten QuaSUM (Peek 2004b) wurden klassenbezogene Rückmeldungen an die Schulen gegeben. Auftraggeber der Studie war das Brandenburger Ministerium für Bildung, Jugend und Sport. Erhoben wurde der Lernstand im Fach Mathematik am Ende der 5. und 9. Klassenstufe bei Schülerinnen und Schülern – in Form einer geschichteten Zufallsstichprobe – in Brandenburg. Insgesamt nahmen ca. 12.500 Schülerinnen und Schüler aus Grundschulen, Gesamtschulen und Gymnasien an der

---

<sup>8</sup> Für die Befragung der Schulleitungen siehe Schrader & Helmke 2003.

querschnittlichen Untersuchung teil. Neben den Leistungsdaten wurden der sozioökonomische Hintergrund, Unterrichtssituation, Klassenklima, psychosoziale Rahmendaten sowie kognitive Voraussetzungen der Schülerschaft erfasst. Die Teilnahme war für die Schulen freiwillig. Ziel der Studie war in erster Linie die Standortbestimmung im Verantwortungsbereich des Bildungsministeriums, um aus den Ergebnissen Konsequenzen für die Bildungsplanung ableiten zu können.

### *2) Form der Ergebnismeldung*

Die Daten wurden von dem Team der Humboldt-Universität zu Berlin, das die Untersuchung durchgeführt hat, eingegeben und ausgewertet und von einer externen Agentur (Research & Value Berlin) für die Rückmeldung aufbereitet (Peek 2004b). Die Schulleitung und die Lehrpersonen erhielten den wissenschaftlichen Zwischenbericht sowie die eigenen Ergebnisse in schriftlicher Form und zwar ein knappes Jahr nach der Erhebung der Daten. Ziel der Rückmeldung war die Anregung der schulinternen Diskussion über Möglichkeiten der Leistungssteigerung. Die Resultate wurden klassenbezogen im Vergleich mit den landesweiten Ergebnissen der Schülerinnen und Schülern der gleichen Schulform mitgeteilt, gegliedert in inhaltliche Teilbereiche (Geometrie, Zahlenbereiche, Schulzufriedenheit etc.). Die Informationen wurden als Balkendiagramme dargestellt und mit Erläuterungen ergänzt.

### *3) Rezeptionsstudie*

Im Juni 2000 wurden die Schulen, die an QuaSUM 1 teilgenommen haben, befragt (Peek 2004b). Insgesamt nahmen 128 Schulleitungen 266 Klassenlehrerinnen und -lehrer sowie 336 Mathematiklehrkräfte an der Befragung QuaSUM 2 teil<sup>9</sup>, das entspricht einer Beteiligung von 74,4 % der Schulen. Die Befragten wurden gebeten, zum einen Auskunft über ihre Einstellung gegenüber zentralen Testuntersuchungen und zum Umgang mit Ergebnismeldungen zu geben, zum anderen wurde untersucht, wie die Befragten das Rückmeldeformat bewertet haben und wie sie selbst mit den Informationen umgegangen sind.

Eine kurze Zusammenstellung der Ergebnisse zeigt, dass zwar grundsätzlich eine positive Einstellung gegenüber Leistungsstudien in der Stichprobe erkennbar ist, die mit dem Unterrichtsgeschehen befassten Personen, also die Lehrkräfte, den Nutzen von Schulleistungsstudien im allgemeinen und im speziellen Fall von QuaSUM jedoch skeptischer betrachten als die Schulleitungen. Ein weiteres wichtiges Ergebnis der Rezeptionsstudie QuaSUM 2 ist,

---

<sup>9</sup> Auch Fachkonferenzen wurden befragt. Die Ergebnisse dieser Befragungen fließen jedoch in die vorliegende Arbeit nicht ein.

dass die befragten Lehrpersonen in Bezug auf die Veröffentlichung und Transparenz der Resultate skeptischer sind als die Schulleitungen. Die Bereitschaft zur Bekanntgabe von Ergebnissen nimmt ab, je mehr befürchtet wird, dass die Ergebnisse einen in eine Verteidigungsposition bringen. Ähnliche Resultate hat auch die Rezeptionsstudie von Kohler (2005) gezeigt, in der die Rezeption von TIMMS untersucht wurde. Die Rückmeldung von QuaSUM, das Rückmelde-Verfahren, die zur Verfügung gestellten Informationsmaterialien und die Unterstützung der Schulen durch Multiplikatoren wurden von der Mehrzahl aller befragten Personenkreise mit den Schulnoten gut oder befriedigend bewertet und die eingeschlagene Strategie insgesamt als positiv bewertet. Die Befragten zeigen ein großes Interesse an den zurückgemeldeten Informationen sowie eine hohe Akzeptanz der normorientierten, vergleichenden Evaluation. Selten haben die Lehrpersonen jedoch ausgesagt, dass aus den zurückgemeldeten Resultaten Konsequenzen für den eigenen Unterricht abgeleitet wurden. Peek (2004b, 61) fasst zusammen:

*„Offensichtlich schätzen die Lehrkräfte und die Schulleitungen die Rückmeldung als Instrument der externen Spiegelung und zur Selbstvergewisserung, halten daraus ableitbare Folgerungen allerdings für begrenzt möglich.“*

Die Ergebnisrückmeldungen wurden zur Kenntnis genommen, jedoch nicht als Impuls für den Austausch zwischen Kollegen in Hinblick auf die pädagogische Arbeit genutzt. Sofern Austausch stattfand, so war dieser fachbezogen, fand also zwischen den Mathematik-Lehrkräften statt.

## **2.4 Thüringer Kompetenztests**

### *1) Studie, in der die zurückgemeldeten Daten erhoben wurden*

In Thüringen werden seit 2002 landesweit die Schülerinnen und Schüler der 3. und 6. Klassenstufen (aller Schularten) in den Fächern Deutsch und Mathematik untersucht. Ziel der Untersuchung Thüringer Kompetenztests (Nachtigall et al. 2005, Schmidt et al. 2005) ist eine Standortbestimmung der Leistungen. Die Resultate sollen den Lehrpersonen behilflich sein, um Schul- und Unterrichtsentwicklung zu betreiben. Es werden jährlich ca. 30.000 Schülerinnen und Schüler untersucht. Die Tests werden von den Lehrpersonen korrigiert, die Ergebnisse internetbasiert erfasst, vom Team Kompetenztests ausgewertet und elektronisch zurückgemeldet.

### *2) Form der Ergebnisrückmeldung*

Die beteiligten Lehrpersonen können die Resultate in einem, mit Passwort geschützten, Schulportal im Internet herunterladen. Als erstes erhalten sie sofort nach der Dateneingabe eine tabellarische Auswertung der Klassenergeb-

nisse (Nachtigall et al. 2005). Als zweites können nach einer kurzen Auswertungsphase so genannte Sofortberichte herunter geladen werden, in denen die Klassenergebnisse für die Besprechung mit Eltern und Schülern aufbereitet wurden. Drittens erhalten die Lehrpersonen – auch über das Internet – einen Ergebnisbericht, in dem die Resultate ausführlich analysiert, Fehlerschwerpunkte aufgezeigt und Vergleichswerte auf Landesebene mitgeteilt werden. Im Wesentlichen enthalten die Rückmeldungen konkrete Ergebnisdaten. Es besteht darüber hinaus die Möglichkeit, detaillierte Analysen der Resultate zu erhalten.

### *3) Rezeptionsstudie*

Die Rezeption der Ergebnisrückmeldungen aus dem Schuljahr 2005 wurde in zweierlei Hinsicht evaluiert (Nachtigall et al. 2005). Zum einen kann durch die rein elektronische Weitergabe der Ergebnisse an die Schulen genau analysiert werden, wie viele Berichte, in welchem Zeitraum von welcher Schule herunter geladen wurden. So konnte ein Nutzerprofil erstellt werden und es zeigte sich, dass drei Monate nach der Veröffentlichung der Endberichte über 90% der Schulen ihren jeweiligen Bericht abgerufen haben.

Zum anderen diente die weitere Evaluierung der Rezeption der Ergebnisrückmeldungen durch Befragungen nicht nur der Klärung der Frage, wie die Lehrpersonen mit den Rückmeldungen umgehen, sondern auch der Verbesserung des Instrumentes selbst. 380 Lehrpersonen und Schulleiter nahmen an der Befragung teil<sup>10</sup>. Die Befragten äußerten sich positiv zu der organisatorischen und technischen Umsetzung sowie zur Art der Information, zu Materialien und Hotline Betreuung. Kritisiert wurde der Zeitpunkt der Rückmeldung, der so spät war, dass die Ergebnisse erst zur Verfügung standen, wenn die Schülerinnen und Schüler schon in der nächsten Klasse, sprich gegebenenfalls bei einer anderen Lehrperson sind. Am häufigsten wurden die Ergebnisberichte der eigenen Klasse genutzt. Es zeigt sich in der Befragung der Lehrpersonen, die an den Thüringer Kompetenztests beteiligt waren, dass die Ableitung von Folgemaßnahmen für die eigene Klassen und den konkreten Unterricht den Schwerpunkt bei der Verwendung der Ergebnisse bilden. Darüber hinaus wurden die Resultate als Ausgangspunkt für Diskussionen zur Qualitätssicherung von Schule und Unterricht genutzt. Die Informationen, die mit den Kompetenztests generiert werden, scheinen im Empfinden der Lehrerinnen Impulse für die Unterrichtsentwicklung geben zu können und als Instrument

---

<sup>10</sup> Darüber hinaus wurden auch Eltern interviewt, auf die Darstellung dieser Untersuchung wird jedoch verzichtet. Vgl. hierfür Nachtigall et al. (2005) Landesbericht Thüringer Kompetenztests 2005 <http://www.kompetenztest.de/downloads1.html>

für eine objektivierbare Leistungsdiagnostik verwendbar zu sein. Nachtigall et al. (2005, 72) weisen darauf hin, dass

*„den Schulen eine Vielzahl von Unterstützungsmöglichkeiten zur Schulentwicklung zur Verfügung stehen, deren Potentiale aber zu einem erheblichen Teil noch nicht ausreichend genutzt werden.“*

Die Verwendung ist demnach nach wie vor ausbaufähig.

## **2.5 Erfahrungen aus der Schweiz: Befragung von Primarschullehrerinnen und -lehrern des Kantons Zürich**

### *1) Studie, in der die zurückgemeldeten Daten erhoben wurden*

Zum Ende des Schuljahrs 2000 wurde die Leistung der Schülerinnen und Schüler der 3. Primarschulklassen des Kantons Zürich flächendeckend evaluiert, und zwar in den Fächern Deutsch und Mathematik. Außerdem wurde die Leistungsbereitschaft und Selbsteinschätzung der Schulzufriedenheit erhoben (für nähere Informationen zu Tests und Ergebnissen vgl. Moser, Keller, Tresch 2002). Diese Untersuchung wurde vom Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich (KBL) durchgeführt, ebenso wie die Ergebnisrückmeldung und die daran anschließende Rezeptionsstudie.

### *2) Form der Ergebnisrückmeldung*

Die mit ihren Klassen an der Untersuchung beteiligten Lehrpersonen, erhielten nach den Sommerferien eine umfassende Rückmeldung über die Ergebnisse der Klasse. Bereit gestellt wurden die Ergebnisse der Klasse insgesamt im Vergleich mit dem Mittelwert aller beteiligten Klassen, sowohl tabellarisch als auch grafisch aufbereitet. Sie erhielten Angaben zur Leistung jedes einzelnen Schülers in jedem einzelnen Test in Prozent sowie die Einschätzung der Leistungsbereitschaft und Schulzufriedenheit auf Klassenebene. Die Ergebnisse wurden durch kurze Informationen erläutert.

### *3) Rezeptionsstudie*

Ein halbes Jahr nach dem die Lehrerinnen die Rückmeldung erhalten haben, wurden sie gebeten bei einer Befragung teilzunehmen (Moser & Keller 2002). Es wurde die allgemeine Beurteilung der Rückmeldung erfasst, die Beurteilung der Funktion der Rückmeldung, die Nutzung der Information sowie die Schwierigkeiten in der Interpretation der Informationen. 66 Lehrerinnen und Lehrer nahmen an der Befragung teil. Über 90% der befragten Personen fanden die Rückmeldung informativ und verständlich, 12% fanden die Information überflüssig und unbrauchbar, 82% empfanden die Rückmeldung als nützlich. Kritisiert wurde, dass die Rückmeldung erst erfolgte als die Lehrerinnen

die untersuchte Klasse bereits abgegeben hatten. 93% der Befragten gaben an, dass die eigene Einschätzung der Klasse durch die zurückgemeldeten Ergebnisse genau oder eher bestätigt wurde. 94% der Lehrpersonen bestätigten, dass die Rückmeldung für eine Standortbestimmung der eigenen Klasse genutzt werden kann. Knapp 70% der Befragten attestieren der Rückmeldung, dass sie zur Prüfung der Lehrplanziele beiträgt, die Optimierung des Unterrichts unterstützt, Informationen zu Leistungsstand liefert und bei der Objektivierung der Bewertung hilfreich ist. Nur 50% finden die Rückmeldung für die Anpassung der Bewertung nützlich. Kritisiert wurde jedoch, dass die einmalige externe Evaluation nicht so hilfreich ist wie kontinuierliche zusätzliche Informationen. 96% der befragten Personen gaben an, dass sie die Rückmeldung tatsächlich für eine Standortbestimmung genutzt haben, 46% haben den Unterricht auf Grundlage der Information optimiert, 28% haben die Beurteilung angepasst und 14% haben die Beurteilung sogar revidiert als Folge der Rückmeldung. Am Schwierigsten gestaltete sich für die Lehrpersonen der Übertrag der Informationen auf die Klasse und die Umsetzung in Rückschlüsse für die Optimierung des Unterrichts. Aus diesen Ergebnissen hat das Team des KBL abgeleitet, dass die Informationen früher versendet, übersichtlicher gestaltet und mit Interpretationshinweisen versehen werden sollten sowie für die Lehrpersonen Weiterbildungsmöglichkeiten bestehen sollten, um die Interpretation und den Umgang mit solchen Informationen besser bewältigen zu können.

## **2.6 Erfahrungen aus Großbritannien: Studien der National Foundation for Educational Research (nfer)**

In Großbritannien<sup>11</sup> wurden von der National Foundation for Educational Research (nfer) zwei Studien durchgeführt, die jeweils untersuchen, wie mit nationalen Daten von Seiten der Lehrpersonen umgegangen wird. Im Folgenden werden diese Studien zur Rezeption – Use of Data – dargestellt.

### *1a) Studie, in der die zurückgemeldeten Daten erhoben wurden*

Die Datengrundlage der ersten Studie (Ashby & Sainsbury 2001), in der die Rezeption von Rückmeldungen untersucht wurde, ist eine regelmäßig durchgeführte Leistungserhebung: Das *National Curriculum* wird in Form von *National Tests* untersucht. Erhoben wurden hier die Fähigkeiten im Lesen, Schreiben, Buchstabieren, Mathematik, Englisch und Sachkunde, jeweils zu

---

<sup>11</sup> Exemplarisch werden hier zwei Studien aus Großbritannien erläutert. Für den interessierten Leser sind in Hinblick auf den internationalen Vergleich auch empfehlenswert die Vorträge der ACER 2005: <http://www.acer.edu.au/workshops/conferences.html>, hieraus besonders Rowe (vergleiche auch Kapitel B\_8 dieser Arbeit).

zwei Zeitpunkten (Key Stage 1, Siebenjährige; Key Stage 2, Elfjährige), Schulformübergreifend.

### *2a) Form der Ergebnisrückmeldung*

Die Ergebnisse werden in Anlehnung an die National Curriculum Levels als numerisches Ergebnis ausgegeben. Die Tests werden entweder in der Schule selbst ausgewertet oder ausgewertet an die Schulen zurückgegeben.

### *3a) Rezeptionsstudie*

Im September 2000 wurden knapp 800 Schulen angeschrieben und gebeten an der Untersuchung teilzunehmen; 415 haben die Fragebögen ausgefüllt zurückgesandt. In dem Fragebogen wurden drei Teilbereiche abgefragt: Erstens wurde gefragt, wie die grundsätzliche Einstellung zum Nutzen der Ergebnisse aus dem National Curriculum Test ist, zweitens wurde nach der tatsächlichen Nutzung der Ergebnisse gefragt und drittens danach, wie die Resultate in die Lehrplanung einbezogen werden. Es zeigte sich, dass die zurückgemeldeten Ergebnisse zum Beispiel genutzt wurden

- für die individuellen Zielsetzungen einzelner Schülerinnen und Schüler und um individuelle Förderpläne zu entwickeln,
- für die Analyse von inhaltlichen Missverständnissen in der Klasse oder bei einzelnen Schülern,
- für die Zuordnung von Schülerinnen und Schüler zu Klassen oder Levels und
- für die Analyse von Schwächen in der Schule, die durch zusätzliche Ressourcen in Hinblick auf Zeit, Personal oder Know-How ausgeglichen werden können.

In einigen Schulen wurde eine sorgfältig geplante Strategie deutlich, wie die zurückgemeldeten Daten effizient genutzt werden können, zum Beispiel um schulhausspezifische, klassenspezifische und personenspezifische Schwächen herauszuarbeiten. Nicht alle, aber der Großteil der Schulen nutzt die Information, um die Schulleistung zu verbessern. Ashby & Sainsbury (2001, 23) stellen heraus:

*“The results of this survey indicate that schools have shown themselves able to respond strongly to recent standards-raising initiatives. Future policy development must ensure that tests and their results continue to be used appropriately.”*

### *1b) Studie, in der die zurückgemeldeten Daten erhoben wurden*

In Großbritannien gibt es neben den oben beschriebenen National Tests eine Vielzahl von statistischen Daten über das Schulsystem, Hintergrundvariablen der Schülerschaft und der Schulen, die hauptsächlich vom Office for Stan-

dards in Education (Ofsted) publiziert werden. Kirkup et al. (2005) untersuchen die Rezeption und die Verwendung dieser Information.

### *2b) Form der Ergebnisrückmeldung*

Über die Form der Rückmeldung der Daten an die Schulen wird von Kirkup et al. (2005) keine Aussage gemacht. Da es hier um die allgemeine Verwendung der Daten geht und die Studie der National Foundation for Educational Research (nfer) grundsätzliche Aussagen dazu machen kann, kann auf weiterführende Darstellungen<sup>12</sup> an dieser Stelle verzichtet werden.

### *3b) Rezeptionsstudie*

Von den angeschriebenen 1820 Schulleitungen und Lehrkräften haben 529 im Frühjahr 2005 einen ausgefüllten Fragebogen zurückgesendet. Ziel der Befragung war herauszufinden, ob und wie die Ergebnisse genutzt werden, um Lernen in Schulen zu verbessern und um zum Beispiel *best practice* identifizieren zu können. Außerdem wurde angestrebt, Empfehlungen aus den Ergebnissen der Rezeptionsstudie ableiten zu können, um Kollegien darin zu unterstützen, die Lernleistung der Schülerschaft aufgrund der Datenanalysen zu verbessern. Als Ergebnis wurde nach der Befragung festgehalten, dass die Daten der landesweiten Schulstatistik genutzt werden,

- um Ressourcen und Mitarbeiter möglichst effektiv einzusetzen, gegebenenfalls auch umzustrukturieren bzw. zu versetzen,
- um *evidence-based* Auseinandersetzungen führen zu können, sowohl intern als auch extern mit der Fachöffentlichkeit oder Verantwortlichen aus Politik und Verwaltung,
- um herausfordernde Zielstellungen zu formulieren, deren Erreichung in einem vorgegebenen Zeitraum angestrebt wird – sowohl auf Schüler- als auch auf Klassenebene,
- um Entscheidungen für die Versetzung von Schülerinnen und Schülern besser treffen zu können.

Die auf Grundlage der Daten initiierten Maßnahmen, die am häufigsten berichtet wurden, sind zusätzliche Unterstützung in Form von Eins-zu-Eins-Betreuung von Schülern sowie Änderungen im Lehrplan in Richtung eines differenzierteren Unterrichts. Die Daten wurden darüber hinaus dann als effektiv und stimulierend empfunden, wenn sich daraus Fragen ergeben haben, zum Beispiel aufgrund von Diskrepanzen zwischen der Eigen- und Fremdwahrnehmung oder in Hinblick auf die Frage, wie Lernen effektiver gestaltet

---

<sup>12</sup> vgl. van Ackeren 2003

werden kann. Die Befragten machten deutlich, dass die benutzerfreundliche Übermittlung und Darstellung der Daten sowie ein Daten-Beauftragter im Schulhaus grundlegend für die Nutzbarkeit sind. Auch in dieser Studie zeigte sich, dass je stärker die Daten bezogen auf den Einzelschüler aufbereitet waren, desto ansprechender waren sie auch für die Lehrperson. Die größte Herausforderung war es, die Zeit aufzubringen, sich überhaupt mit den Daten beschäftigen zu können und sich selbst als fähig für die Datennutzung anzusehen, also im Verständnis der eigenen professionellen Rolle die Auseinandersetzung mit externen Daten als Entwicklungsmöglichkeit zu verankern.

## **2.7 Entwicklung und Implementation eines Systems zur Qualitätssicherung an Schulen unter besonderer Berücksichtigung des naturwissenschaftlichen Unterrichts (QuaSSU)**

Ergänzend zu den oben beschriebenen Rückmeldungen im Rahmen von nationalen und internationalen Leistungsstudien soll das Projekt *QuaSSU* noch Erwähnung finden sowie eine Fragebogenstudie. Beides wurde im Rahmen des Schwerpunktprogramms *Qualität von Schule und Unterricht* durchgeführt. Ditton et al. (2002a, 2002b) haben im Rahmen der verschiedenen Befragungen untersucht, wie die Einstellung von Lehrpersonen gegenüber zentralen Leistungstest ist und wie mit den Ergebnissen aus den Untersuchungen verfahren werden sollte.

### *a) Fragebogenstudie*

In einer Fragebogenstudie wurden insgesamt 1754 Lehrpersonen und Schulleitungen befragt (Ditton et al. 2002a). Es zeigte sich, dass Schulleitungen gegenüber Leistungsstudien wesentlich positiver eingestellt sind als Lehrpersonen, beide Gruppen befürworten jedoch eine Rückmeldung der Ergebnisse an die teilnehmenden Schulen bzw. Klassen. Insgesamt wird in der Studie deutlich, dass die Lehrerschaft gespalten ist: Kritiker und Befürworter von Rückmeldungen halten sich die Waage, als Ergebnis wird festgehalten (Ditton et al., 2002b, 30), dass Lehrpersonen die Ergebnisrückmeldungen nur für sinnvoll halten und nutzen, wenn die Informationen

*„konkret und spezifisch über das eigene Abschneiden, idealerweise in mehreren Klassen und zu mehreren Zeitpunkten ... informiert werden.“*

### *b) QuaSSU*

Hier wurden insgesamt 194 Lehrkräfte zu ihren Einstellungen in Hinblick auf Schule und Unterricht befragt, bezogen auf die Fächer Mathematik, Deutsch und Englisch (Ditton et al. 2002a, 2002b). Auf die Ergebnisse dieser Befra-

gung wird nicht näher eingegangen, da der Fokus der vorliegenden Arbeit auf dem Umgang mit der Rückmeldung liegt. Die Ergebnisse wurden für die Befragten aufbereitet und zurückgemeldet, die (im verschlossenen Umschlag versendeten) Broschüren wurden über den in der Schule bestimmten Koordinator verteilt: Jeweils ein Koordinator pro Schule war dafür verantwortlich, die aufgrund der Untersuchung zur Unterrichtsqualität erstellten Rückmeldungen an seine Kollegen weiterzugeben. Die Rückmeldungen enthielten das eigene Ergebnis als grafische Darstellung im Vergleich zu einer Referenzgruppe (Schulart und Fach). Mit der Rückmeldung wurde gleichzeitig der Fragebogen der Rezeptionsstudie versendet. Es wurden auch Interviews und Gruppendiskussionen geführt, auf die hier nicht näher eingegangen wird. Grundsätzlich äußerten sich die Befragten positiv über die Rückmeldungen, wobei Rückmeldungen in Hinblick auf die Verwertbarkeit der Ergebnisse nur bei zwei Drittel der Teilnehmer als zufrieden stellend bewertet wurden.

In der Rezeptionsstudie wurde geäußert, dass die Verarbeitung der Ergebnisrückmeldungen mit dem gesamten Kollegium im Rahmen eines „Pädagogischen Tags“ hilfreich wäre und die Unterstützung durch externe Instanzen oder durch ein Schulnetzwerk wünschenswert wäre. Außerdem wünschten sich die Befragten, dass bei Rückmeldungen neben Mittelwerten auch differenzierte Informationen über das Leistungsprofil der Klasse und/oder der Schülerinnen und Schüler mitgeteilt werden. Sie gaben an, dass die Ergebnisse mit den Schülern und im Fachkollegium besprochen wurden und der Wunsch nach einer Beratung durch eine externe und neutrale Instanz besteht. Ditton et al. (2002b, 387) bewerten die Ergebnisse zusammenfassend dahingehend, dass

*„die Verweigerung von Einheitslösungen (one size fits all) möglicherweise aber auch das ohnehin vielfach geforderte plurale Selbstverständnis der Schulen sowie der Professionalität oder das Selbstbewusstsein einer Berufsgruppe zeigt.“*

Die verschiedenen Rezeptionsstudien aus Deutschland, Großbritannien und der Schweiz zeigen ein breites Spektrum von Untersuchungen, in denen Rückmeldungen vorgenommen werden. Es zeigt sich, dass jeweils ähnliche Aspekte von Interesse waren: Wie wurde die Art der Rückmeldung bewertet? Gab es Kritik oder Wünsche in Hinblick auf die Gestaltung der Information? Wie wurden die Informationen von den Lehrerinnen und Lehrern genutzt?

Die Resultate der Befragungen der Lehrpersonen erweisen sich als relativ deckungsgleich. Zusammenfassend kann festgehalten werden, dass die Informationen als verständlich bewertet wurden, häufig jedoch zu spät zur Verfü-

gung standen. Die Ergebnisse werden akzeptiert und sie sollen möglichst so gestaltet sein, dass sie auf die einzelnen, eigenen Schülerinnen und Schüler beziehbar sind. Die Resultate werden von den Rezipienten reflektiert, aber vor allem dahingehend, dass als Konsequenz oberflächliche Maßnahmen ergriffen werden: Die in den Rezeptionsstudien berichteten Ergebnisse weisen nicht darauf hin, dass das eigene pädagogische Handeln von den Lehrpersonen auf Grundlage der Ergebnisrückmeldungen analysiert wird. Fraglich ist, ob durch die Rückmeldungen tatsächlich Schul- und Unterrichtsentwicklung in Gang gebracht wird. Im anschließenden Kapitel widme ich mich dem Prozess, der von der Evaluation hin zur Innovation durchlaufen werden kann.