

7. Studie 3: Drei Schweizer Instrumente zur Leistungsmessung und Ergebnisrückmeldung im Vergleich mit BeLesen

7.1 Erläuterung der Studie

Neben den oben beschriebenen Studien 1 und 2, die im Rahmen von BeLesen stattgefunden haben, habe ich mich während meiner Promotionsphase auch in anderer Hinsicht mit dem Thema der Ergebnisrückmeldungen beschäftigt. Im Herbst 2005 war ich für zehn Wochen als Gutachterin in der Bildungsdirektion des Kantons Zürich tätig und habe drei Instrumente, die in der Schweiz für die schulinterne Messung von Schülerleistungen verwendet werden, begutachtet und mit meinen Erfahrungen im Projekt BeLesen in Beziehung gesetzt: Check 5, Klassencockpit und Orientierungsarbeiten Zentralschweiz.

Da diese inhaltliche Auseinandersetzung die empirischen Erfahrungen, wie in Studie 1 und 2 beschrieben, abrunden, habe ich mich in Absprache mit der Bildungsdirektion dazu entschlossen, das Gutachten in die Dissertationsschrift mit einzubeziehen.

Hierfür folgt nun auf den nächsten Seiten die Erläuterung des Vorgehens, eine kurze Beschreibungen der Instrumente und Angaben, wo detailliertere Informationen zu den Instrumenten eingeholt werden können. Im Anschluss daran werden die drei Schweizer Instrumente in Hinblick auf die Nutzungsmöglichkeiten, Bezugsnormen und Gütekriterien mit den Ergebnisrückmeldungen von BeLesen verglichen. Als Ausblick werden in der abschließenden Diskussion Gestaltungsempfehlungen für Ergebnisrückmeldungen erläutert, die sich auch als Folge aus den in Studie 1 und Studie 2 dargelegten Erkenntnissen ergeben.

7.2 Methode

Die zu begutachtenden Instrumente Check 5, Klassencockpit und Orientierungsarbeiten Zentralschweiz wurden dokumentenanalytisch gesichtet. Die Methode der Dokumentenanalyse ist ein klassisches Verfahren der qualitativ-interpretativen Analyse und wird als non-reaktives Messverfahren verstanden (Mayring 2002). Dokumente werden im Rahmen dieses Verfahrens als Zeugnis bzw. Objektivierung menschlichen Denkens, Fühlens und Handelns betrachtet (Atteslander 1971, Ballstaedt 1987). Durch die Analyse der Dokumente können Aussagen über die Intentionen der Menschen gemacht werden, die an der Erstellung von Texten, Urkunden, Berichten etc. beteiligt gewesen

sind. Vorteil der Methode ist, dass die Daten – die Dokumente – bereits vorhanden sind. Nachteil ist jedoch, dass die subjektive Auswahl der Datenquellen gewisse Verzerrungen mit sich bringen kann (Mayring 2002). Aus den Dokumenten können Informationen über Fakten, Ereignisse (z.B. bei historischen Dokumenten) und normative Regeln generiert werden. Nachdem eine Definition des Ausgangsmaterials in Bezug auf eine Fragestellung vorgenommen wird – hier alle Dokumente, die den Lehrpersonen für die Leistungstests und deren Auswertung zur Verfügung stehen – wird der Aussagewert des Dokumentes eingeschätzt und der Informationsgehalt des Dokuments interpretiert (Mayring 2002).

In der vorliegenden Arbeit ist das Ziel der Dokumentenanalyse, empirisch gestützte Aussagen in Hinblick auf zwei Fragen machen zu können: Welche Informationen stehen den Lehrpersonen bei Check 5, Klassencockpit und den Orientierungsarbeiten im Vergleich zu BeLesen zur Verfügung?

Welche Ziele und Nutzungsmöglichkeiten der Informationen werden von den Autorinnen und Autoren der Tests und Rückmeldungen kommuniziert?

Folgende Schriften, Broschüren, Materialien sowie bereits bestehende Evaluationen der Instrumente sind in der dokumentenanalytischen Begutachtung berücksichtigt worden. Sowohl die in Papierform als auch die in elektronischer Form online vorliegenden Informationen zu den Instrumenten wurden einbezogen.

Orientierungsarbeiten

Die Orientierungsarbeiten der Klassen 5 und 7 in den Fächern Mathematik und Deutsch.

Moser, F. & Bucher, M. (2003). *Regionale Erarbeitung von Orientierungsarbeiten in der Volksschule*. Zwischenbericht. Luzern: Bildungsplanung Zentralschweiz.

Mahler, W., Armand, C., Gautschi, T., Naef, P. (2004). *Vergleichstests bzw. Orientierungsarbeiten zur Selbstevaluation auf der Gymnasialstufe*. Bericht der Spurgruppe.

Senn, W. & Moser, F. (2003). Orientierungsarbeiten. *Mittendrin*. Luzern: Bildungsplanung Zentralschweiz. 1/2003 S.3-4

Vögeli-Mantovani, U. (2003). Klassencockpit oder Orientierungsarbeiten. *Mittendrin*. Luzern: Bildungsplanung Zentralschweiz. 1/2003 S.6-7

Klassencockpit

Broschüre Klassencockpit – ein Instrument zur Ermittlung des Leistungsstandes von Schülerinnen und Schülern. Kanton Zürich.

Broschüre Klassencockpit – Informationen für Lehrpersonen. Kanton Zürich.

www.klassencockpit.ch Stand 1.10.2005

Landert, Farago Davatz & Partner (2002). *Klassenscockpit. Evaluation*. Im Auftrag der Projektleitung Klassenscockpit. Zürich.

Moser, U. (2003). *Klassenscockpit im Kanton Zürich. Ergebnisse einer Befragung von Lehrerinnen und Lehrern der 6. Klassen über ihre Erfahrungen im Rahmen der Erprobung von Klassenscockpit im Kanton Zürich im Schuljahr 2002/03*. Zürich: Bildungsdirektion des Kantons Zürich.

Zoppi, S. (2003). *Klassenscockpit – ein Instrument zur Qualitätssicherung und –entwicklung?* Diplomarbeit. St. Gallen

Check 5

Testhefte Deutsch, Mathematik und Problemlösen.

Rückmeldebroschüre in der Musterversion Durchgang Schuljahr 2004/2005.

Moser, U. & Tresch, S. (2004). *Erste Ergebnisse Check 5*. Zwischenbericht 2004 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich.

Moser, U. & Tresch, S. (2004). *Muster Rückmeldung der Ergebnisse Check 5*. Kanton Aargau. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich.

Moser, U. & Tresch, S. (2005). *Leistungen messen und beurteilen. Handreichung zum Umgang mit Ergebnissen von Leistungstests*. Herausgegeben vom Departement Bildung, Kultur und Sport des Kantons Aargau (BKS). Aarau: BKS.

Tresch, S. & Moser, U. (2005). *Check 5*. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich.

Kritisch ist zu beachten, dass nur die Interpretation der Texte durch die Autorin vorliegt. Zwar wurden in der Schweiz verschiedene formelle Gespräche (Interviews) und informelle Gespräche über die Informationen, die die jeweiligen Tests und Rückmeldungen liefern, geführt. Die Analyse der Dokumente wurde so durch die Interviews mit Schweizer Lehrpersonen unterfüttert.

Da die in den Interviews generierten Informationen für die Bildungsdirektion allein bestimmt waren, können die Aussagen der Befragten nicht spezifiziert werden.²⁷ Das so entstandene Wissen ist jedoch in die Bewertung der Schweizer Instrumente eingegangen.

Im Rahmen einer intensiven persönlichen Auseinandersetzung wurden die Dokumente aus dem Blickwinkel der angesprochenen Lehrpersonen interpretiert und gelesen, immer unter folgenden Fragestellungen:

- Sind die Informationen verständlich?

²⁷ vgl. Schneewind, Julia (2006b). *Gutachten zu Form und Einsatz von Ergebnisrückmeldungen für die Unterrichtsentwicklung. Die Instrumente Klassenscockpit, Orientierungsarbeiten, Check 5*. Im Auftrag der Bildungsdirektion des Kanton Zürich, Abteilung Bildungsplanung.

- Sind die Inhalte leicht zugänglich?
- Wie könnten diese Informationen auf meine Schülerinnen und Schüler bezogen werden?
- Welche Informationen werden gewonnen, wenn die Tests mit der anschließenden Auswertung durchgeführt werden?
- Ergeben sich Konsequenzen für das Unterrichtshandeln durch die neuen Informationen?

Kritisch betrachtet beinhaltet die Dokumentenanalyse einen großen subjektiven Spielraum: Jede Lehrperson kann die Daten unterschiedlich interpretieren, verstehen und so kann auch die von mir vorgenommene Begutachtung in Frage gestellt werden. Dennoch bietet die im Folgenden dargelegte vergleichende Analyse Ansatzpunkte, um die Diskussion über die Nutzbarkeit von Ergebnisrückmeldungen zu unterfüttern. Die in Studie 1 und 2 erläuterten Ergebnisse können unter Berücksichtigung von den Erfahrungen aus der Schweiz weitergehender diskutiert werden.

7.3 Beschreibung der Instrumente in Kurzform

Im Folgenden werden die begutachteten Instrumente und die Studie BeLesen kurz beschrieben. Die Auswahl der Instrumente erfolgte durch die Bildungsdirektion des Kantons Zürich. Im Anhang finden sich Beispiele der Tests und der jeweiligen Rückmeldung.

7.3.1 Orientierungsarbeiten

Das Instrument und die Rückmeldung

Die Orientierungsarbeiten sind ein Projekt in der Zentralschweiz. Hier wurden 30 Orientierungsarbeiten für die Schulstufen 1 bis 9 in den Fächern Deutsch, Mathematik, Mensch & Umwelt, Naturlehre, Musik, Geografie, Geschichte, Bildnerisches Gestalten, Technisches Gestalten, Hauswirtschaft und Lebenskunde erarbeitet. Die Orientierungsarbeiten sind eine Sammlung von vielfältigen Aufgabenstellungen jeweils eines Fachs und einer Stufe, die in einem gebundenen Heft als Kopiervorlagen für die Lehrperson zusammengestellt sind. Diese sollen zur systematischen Prüfung der Lernzielerreichung dienen. Die Orientierungsarbeiten, sprich jedes einzelne Arbeitsblatt, können zu jedem beliebigen Zeitpunkt eingesetzt werden: Als Lernstandmessung zu Beginn einer Unterrichtseinheit, als Zwischenevaluation, als Abschlusstest am Ende einer Unterrichtseinheit. Die Inhalte stützen sich auf die Lernziele des entsprechenden Fachlehrplans und wurden von einem Team aus Lehrpersonen erstellt.

Die Aufgaben weisen einen je unterschiedlichen Schwierigkeitsgrad und unterschiedliche Komplexität auf, so dass die Lehrperson entsprechend ihren selbst gesetzten Zielen Aufgaben auswählen kann. Die Aufgaben haben alle ein einheitliches Auswertungsraster für eine lernziel- und förderorientierte Beurteilung: Es werden für den Schüler sichtbar Lernziele formuliert, eine Aufgabe gestellt und ein Antwortformat vorgegeben (Beispiel siehe Anhang: Schreibe einen Brief an eine Versicherung). Die Lehrperson kann dann, nachdem das Aufgabenblatt bearbeitet wurde, anhand von vorgegebenen Zielen entscheiden, ob der Schüler die Aufgabe erfüllt bzw. nicht erfüllt hat oder die Anforderungen übertroffen wurden. Die Orientierungsarbeiten sind nicht als Übungen in Form von Arbeitsblattsammlungen für den regulären Unterricht gedacht, sondern als Element zur Prüfung des Lernstandes.

Bei den Orientierungsarbeiten wird keine externe Rückmeldung gegeben, sondern es gibt nur die von Lehrpersonen generierte Information über die Leistung der Schülerinnen und Schüler. Die Schülerinnen und Schüler bearbeiten die jeweiligen Arbeiten, die Lehrperson korrigiert diese anhand der vorgegebenen Kriterien und erhält so Informationen über den Leistungsstand

der Schülerschaft. Die Information ist objektiviert durch die vorgegebenen Kriterien, so dass auch zwischen Klassen verglichen werden kann.

Ziele der Orientierungsarbeiten

Die Orientierungsarbeiten dienen dazu, das zielorientierte Lernen zu üben, die Ziele der Aufgabenstellung werden den Lernenden erläutert. Außerdem können die Orientierungsarbeiten die Umsetzung der Lehrpläne in den Unterricht erleichtern und die in den Lehrplänen verfolgten Zielorientierungen unterstützen. Die kontinuierliche Arbeit mit diesen Aufgaben soll die Kompetenz der Lehrkräfte in Hinblick auf förderorientiertes Beurteilen stärken und als Instrument der förderorientierten Planungshilfe fungieren. Der individuelle Lernstand der Schüler wird festgestellt, Leistungen werden differenziert beschrieben, die Lehrperson ist angehalten, den Schülerinnen und Schülern jeweils eine förderorientierte Rückmeldung zu geben, um das Leistungsprofil der Person zu veranschaulichen. Als Folge der Orientierungsarbeiten sollte der Unterricht auf Stärken und Schwächen der jeweiligen Schülerinnen und Schüler ausgerichtet werden. Von den Autoren der Orientierungsarbeiten wird ausdrücklich darauf hingewiesen, dass dies kein Controlling- oder Selektionssystem ist und die Leistungsmessung oder der Vergleich von Leistungsniveaus verschiedener Klassen nicht das Ziel sind.

7.3.2 Klassencockpit

Das Instrument und die Rückmeldung²⁸

Dieses Instrument ist ein Testsystem zur Qualitätssicherung im Volksschulbereich. Die Testreihe ist einsetzbar von der 3. bis 9. Klasse und dient der Überprüfung der Lernziele in Mathematik und Deutsch. Zwei bis drei Module können jeweils für beide Fächer zu verschiedenen Zeitpunkten im Laufe des Jahres durchgeführt werden, meist wird Klassencockpit im November, März und Juni eingesetzt. Die Tests sind geeicht und standardisiert.

Die Lehrperson erhält eine Rückmeldung, in der die Klassenwerte mit dem Wert einer geeichten Stichprobe von 450 bis 600 Schülerinnen und Schülern verglichen werden. Die Rückmeldung erfolgt online und enthält im Wesentlichen ein Diagramm, das die Ergebnisse der Klasse mit der Eichstichprobe in Beziehung setzt. Weitere Informationen sind aus den Rückmeldungen ablesbar: Es werden Vergleichswerte mitgeteilt, da sowohl die Ergebnisse der Eichstichprobe als auch die der eigenen Klasse vorliegen. Die durchschnittliche Anzahl richtig gelöster Aufgaben, die Klassenmittelwerte und Punktzah-

²⁸ www.klassencockpit.ch

len pro Schüler werden in der Rückmeldung erläutert: Wie viele Aufgaben wurden richtig gelöst?

Außerdem wird der Prozentrang angegeben, also die relative Stellung der Schülerinnen und Schüler in Bezug zur Eichstichprobe (soziale Bezugsnorm). In der Rückmeldung werden Notenvorschläge gemacht, die sich an den Noten und Ergebnissen der an der Eichung beteiligten Schülerinnen und Schüler orientieren. Außerdem wird die prozentuale Verteilung der einzelnen Notenstufen ausgewiesen. Die Lehrperson kann Informationen über die eigene Benotungspraxis ablesen: Wird eher zu streng oder eher zu mild bewertet? Diese Information kann auch für den Austausch zwischen Lehrpersonen einer Schule nützlich sein.

Die Leistung der Schülerinnen und Schüler werden in Anforderungsniveaus eingeteilt: Schüler, die den leistungsstärksten 25 % der Eichstichprobe entsprechen sind in der Gruppe Niveau 1. Schüler, die den leistungsschwächsten 25 % der Eichstichprobe entsprechen sind in der Gruppe Niveau 3. Schüler die dazwischen liegen sind in der Gruppe Niveau 2.

Ziele von Klassencockpit

Klassencockpit dient – anders als die Orientierungsarbeiten – der kontinuierlichen und in erster Linie streng standardisierten Qualitätskontrolle. Der Unterschied zwischen Klassencockpit und den Orientierungsarbeiten besteht darin, dass bei Erstgenannten standardisierte Tests verwendet werden, die Antworten nur in den Kategorien richtig oder falsch bewertet werden können, die Lehrpersonen keine eigene Bewertung der Leistung vornehmen, sondern die Auswertung extern erfolgt. Der extern vorgenommene Vergleich mit anderen soll von den Lehrpersonen zur Standortbestimmung genutzt werden, um davon ausgehend den Unterricht zu reflektieren und zu verbessern. Klassencockpit kann prüfen, ob die Lernziele erreicht wurden, sofern die Inhalte der vergangenen Lerneinheiten mit den in den Tests vorgesehenen Themen kompatibel sind. Die Qualität des Unterrichts und die Leistungsfähigkeit der Schülerschaft sollen durch Selbstevaluation gesichert werden. In den einzelnen Kantonen, die Klassencockpit verwenden, wird sehr unterschiedlich gehandhabt, wem die Ergebnisse von Klassencockpit zugänglich sind. Teilweise erhalten die Eltern eine Rückmeldung, teilweise werden die Ergebnisse im Lehrerkollegium ausgetauscht, teilweise möchte auch die Schulaufsicht Einsicht in die Daten erhalten.

7.3.3 Check 5

Das Instrument und die Rückmeldung²⁹

Check 5 ist eine vom Kompetenzzentrum für Bildungsevaluation und Leistungsmessung (KBL) an der Universität Zürich entwickelte Schulleistungsmessung mit anschließender Ergebnisrückmeldung. Bisher wurde das Instrument nur im Kanton Aargau eingesetzt. Zu Beginn der 5. Klasse werden mit den Schülerinnen und Schülern Tests in den Lernbereichen Deutsch, Mathematik und Problemlösekompetenz durchgeführt. Von den Lehrpersonen wird entschieden, ob sie mit ihren Klassen an der freiwilligen Leistungsmessung teilnehmen möchten.

Die Rückmeldung erfolgt nach einer externen Auswertung durch das KBL ca. 8 Wochen nach dem Test. Die Ergebnisrückmeldung enthält Informationen über die eigenen Schüler sowie anonymisierte Klassenwerte der anderen Teilnehmerinnen und Teilnehmer. Die Individualergebnisse der Schülerinnen und Schüler werden in Prozent dargestellt, und zwar aufgeschlüsselt nach Teilbereichen und im Vergleich zum erreichten mittleren Prozentwert der Klasse und der Stichprobe. Ablesbar sind also die individuellen Leistungen, in Beziehung zur Referenzgruppe Klasse und Gesamtstichprobe. Außerdem wird die Verteilung der Leistungen innerhalb der Klasse (Mathematik und Deutsch) in Hinblick auf die Mindestziele dargestellt sowie der Grad der Heterogenität innerhalb der Klasse durch ein Schaubild illustriert. Hier kann die Lehrperson ablesen, welche Leistung die einzelnen Schülerinnen und Schüler in beiden Fächern erzielt haben, so dass ein vollständigeres Bild gegeben wird, als wenn die Fächer unabhängig voneinander dargestellt werden.

Die Leistungen der Klasse in beiden Fächern, wieder Angaben in Prozentwerten, werden im Vergleich zu den Leistungen der anderen Klassen und im Verhältnis zum Durchschnitt in einem weiteren Schaubild dargestellt. Hier ist der Standort der eigenen Klasse bestimmbar. Die Umsetzung von Check 5 wird regelmäßig durch das KBL evaluiert.

Ziele von Check 5

Check 5 möchte eine unabhängige Standortbestimmung der Leistungen der Schülerinnen und Schüler zu Beginn der 5. Klasse ermöglichen, um so die Förderung der Schülerinnen- und Schülerleistung gezielter vornehmen zu können. Das Instrument und die damit verbundene Ergebnisrückmeldung möchten die Lehrpersonen in der Weiterentwicklung eines wirkungsvollen, leistungsorientierten Unterrichts unterstützen. Check 5 wird nicht als Instrument zur Selektion verstanden. Wichtig ist, dass Check 5 sowohl Weiterbil-

²⁹ www.ag.ch/leistungstest

dungsangebote für Lehrpersonen zur Thematik der Leistungsmessung beinhaltet als auch die Lehrkräfte explizit dazu auffordert, aus den Ergebnissen Maßnahmen abzuleiten.

7.3.4 BeLesen im Vergleich zu den anderen drei Instrumenten

BeLesen beinhaltet im Gegensatz zu den anderen drei Formaten kein Einzelinstrument, sondern eine Zusammenstellung verschiedener Tests. In dieser Schulleistungsstudie wird mit bestehenden Instrumenten gearbeitet: Hamburger Schreibprobe (HSP), Würzburger Leise Lese Probe (WLLP), Test zum Leseverständnis (ELFE), Diagnostischer Rechtschreibtest (DRT), DEMAT Mathematiktest. Hiermit werden Grundschul Kinder im Rahmen einer Längsschnitt-Studie von der ersten bis zur vierten Klasse untersucht, um ihre Lernentwicklung zu dokumentieren. Die Lehrpersonen erhalten acht Wochen nach der Durchführung der Leistungstests eine Ergebnismeldung in schriftlicher Form. Ziel ist es, die Lehrpersonen über den Leistungsstand der Schülerinnen und Schüler zu informieren und ihnen so eine Datenbasis zur Verfügung zu stellen, aufgrund der sie ihr Unterrichtshandeln reflektieren können. Für die nähere Erläuterung der Studie und des Rückmeldeformats vergleiche Kapitel A_6.

Vergleicht man BeLesen und die drei anderen Instrumente, so wird ein grundlegender und entscheidender Unterschied deutlich: BeLesen ist eine Schulleistungsstudie, die Ergebnismeldungen als Nebenprodukt generiert. Die Rückmeldungen werden zwar mit der Intention an die Lehrerinnen gegeben, dass sie daraus für sich hilfreiche Informationen ableiten, jedoch war das Ziel der Studie ein anderes, nämlich *best practice* im Anfangsunterricht für Kinder mit Migrationshintergrund zu identifizieren. Klassencockpit, Orientierungsarbeiten und Check 5 hingegen sind Instrumente, die explizit dafür entwickelt wurden, den Leistungsstand einer Klasse bzw. eines Jahrgangs zu erfassen, und gleichzeitig Informationen über das Leistungsprofil der Schülerinnen und Schüler zu liefern. Die drei Schweizer Instrumente haben das Ziel, eine Qualitätssteigerung von Leistungsbeurteilung durch die Einführung von Verfahrensstandards zu erreichen, da die Leistungsbeurteilungen von Klasse zu Klasse bisher nicht vergleichbar sind (vgl. Moser & Bucher 2003).

Check 5 nutzt differenziertere Aufgaben und Unterbereiche und ist stärker als Klassencockpit mit einer Nutzungsverpflichtung verbunden. Zum Beispiel ist der Informationsabend verpflichtend. Die Lehrpersonen sind verpflichtet, die Daten zu analysieren, Maßnahmen zu entwickeln, diese umzusetzen und später im Schuljahr die Effektivität der Maßnahmen selbst zu überprüfen. Klassencockpit ist unverbindlicher.

Die Orientierungsarbeiten haben als Fokus zielerreichendes Lernen, individuelle Lerndiagnose, Unterrichtsentwicklung. Die Rückmeldung ist differenzier-

ter und stärker auf das einzelne Kind bezogen. Bei Klassencockpit steht die klassenübergreifende Vergleichbarkeit der Leistungsbeurteilung mit hoher Testqualität im Vordergrund. Die Abweichungen und Resultate müssen von den Lehrpersonen bei Klassencockpit selbstständig analysiert werden. Aus den so generierten Daten können nur bedingt Schlüsse für die Unterrichtsentwicklung gezogen werden, da beispielsweise keine Fehlerprofile mitgeteilt werden, sondern explizit die Standortbestimmung der Klasse das Ziel von Klassencockpit ist.

Die Orientierungsarbeiten und Klassencockpit setzen an unterschiedlichen Punkten an. Beurteilungen in Orientierungsarbeiten liefern die Antworten, wo Ergebnisse von Vergleichstests erst Fragen aufwerfen (Vögeli-Mantovani 2003). Die Anwendung eines Formates wie z.B. die Orientierungsarbeiten könnte also nach Erhalt der Rückmeldungen von BeLesen hilfreich sein, um die Resultate mit weiteren Informationen zu ergänzen.

Alle Instrumente bedeuten eine Zusatzarbeit für Lehrpersonen. Der Mehrwert des Aufwandes liegt nach Aussage der Lehrpersonen darin, dass alle Formate für die Standortbestimmung und in Elterngesprächen oder für Argumentationen gegenüber der Schulleitung nützlich sind. Die drei Schweizer Instrumente, ebenso wie BeLesen, liefern Ergebnisse, die die Selbstkontrolle und professionelle Reflexion unterstützen. Bei allen vieren ist die Verwendung freiwillig. Die an BeLesen teilnehmenden Klassen wurden zwar von den Lehrerinnen auch freiwillig dazu angemeldet, jedoch ist diese Verpflichtung durch die vierjährige Laufzeit der Studie wesentlich länger als bei einem ein- oder zweimalig einzusetzenden Instrument.

Es folgen nun Detailvergleiche in Hinblick auf Bezugsnormen, Nutzungsmöglichkeiten und Gütekriterien.

7.4 Vergleichende Analysen

Die drei Instrumente und BeLesen werden nun in Hinblick auf vier Aspekte miteinander in Beziehung gesetzt. Diese vier Aspekte stehen für die im Zyklusmodell von Helmke und Hosenfeld (2005) abgebildeten Schritte Rezeption, Reflektion und Aktion. Auch der Vergleich der Instrumente lehnt sich an dieses Modell an, um den Verlauf von der Evaluation zur Innovation abzubilden.

- Gestaltung
- Bezugsnormen
- Grad der Differenziertheit
- Nutzungsmöglichkeiten

Die **Rezeption** wird mit der Betrachtung der Gestaltung der Rückmeldung und den hierbei umgesetzten Aspekten abgebildet. **Reflektion** wird über die Differenziertheit der Formate sowie durch die Bezugsnormen dargestellt, die Nutzungsmöglichkeiten widmen sich dem Aspekt **Aktion**.

7.4.1 Rezeption: Regeln der Gestaltung

Im Folgenden werden die drei zu begutachtenden Instrumente anhand der im Kapitel A_5.2 erläuterten Kriterien begutachtet. Die Begutachtung der Gestaltung beruht auf den theoretisch aus der oben dargelegten Literatur abgeleiteten Aspekten, die bei der Gestaltung berücksichtigt werden sollten. Diese wurden im Kapitel A_5.2 ausführlich erläutert. Die Dokumente wurden in Hinblick auf die Frage analysiert, ob die Formate die einzelnen, zusammengetragenen Aspekte bei der Gestaltung umsetzen. Dazu gehört beispielsweise die Aufbereitung der Information, der Zeitpunkt der Rückmeldung, die Verständlichkeit und Lesbarkeit der Abbildungen und Tabellen. In der folgenden Tabelle werden Abkürzungen verwendet, die hier kurz erläutert werden:

OA	Orientierungsarbeiten
KC	Klassenscockpit
C 5	Check 5
BL	BeLesen
+	Dieser Aspekt wurde umgesetzt.
±	Dieser Aspekt wurde berücksichtigt, ist aber in der Umsetzung verbesserungswürdig.
–	Dieser Aspekt wurde nicht umgesetzt.
?	Dieser Aspekt kann nach jetzigem Kenntnisstand nicht beurteilt werden.
/	Dieser Aspekt kommt in dem Format aufgrund der grundsätzlichen Anlage nicht vor.

Tabelle 71: Gestaltung der Ergebnisrückmeldung, Vergleich verschiedener Kriterien

<i>Kriterien, die betrachtet werden</i>	<i>Bewertung der Kriterien</i>			
	KC	C 5	OA	BL
1) Der Feedbackgebende Part: Die Institution, die die Rückmeldung gibt, ist kompetent, vertrauenswürdig	+	+	+	+
2) Ziel- und Zweckgebundenheit:				
<i>a. Definition und Kommunikation des Ziels des Instruments sind gegeben.</i>	±	±	±	±
<i>b. Die Grenzen der Information werden thematisiert</i>	±	±	±	±
3) Übermittlung der Information: Rückmeldung ist einfach zu erhalten (online oder schriftlich).	+	+	+	+
4) Ansprechpartner: Für organisatorische, inhaltliche, fachliche und statistische Fragen steht jemand zur Verfügung.	±	±	±	±
5) Aktualität und Zeit. Die Rückmeldung erfolgt zeitnah.	+	±	+	±
6) Curriculare Validität: Berücksichtigung des Lehrplans.	?	?	+	?
7) Zielgruppenorientierung: Orientierung an der Handlungsebene der Zielgruppe (Klassen- und Schülerwerte).	+	+	+	+
8) Handlungsbezug: Interpretationshinweise sind vorhanden.	-	±	?	-
9) Verständlichkeit und Darstellungen:				
<i>a. Verständlich, in der Fachsprache der Zielgruppe, Glossar, Übersetzung statistischer Kennwerten</i>	±	±	/	±
<i>b. Tabellen, Diagramme sind vollständig und nachvollziehbar beschriftet, lesbar und übersichtlich.</i>	-	-	/	±
10) Informationsaufbereitung:				
<i>a. Fehleranalysen werden angeboten</i>	-	-	-	-
<i>b. Stärken/Schwächenprofile der Schüler/der Klasse</i>	±	±	+	±
<i>c. Vergleich mit vergleichbaren Referenzgruppe ist möglich*</i>	-	-	-	-
<i>d. Hinweise zur Förderdiagnostik</i>	-	-	-	-
<i>e. Ausführliche Wegleitung für die Verwendung</i>	-	±	-	-
<i>f. Minimalstandards, die erreicht werden müssen, sind angegeben</i>	-	+	+	-

*In keinem Format werden Kontextvariablen erhoben, daher gibt es auch in keinem der hier begutachteten Instrumente adjustierte Werte.

Insgesamt zeigt sich, dass alle Schweizerischen Instrumente sowie die Ergebnismeldung von BeLesen verbesserungswürdig sind, vor allem in Hinblick auf die Informationsaufbereitung. Meiner Begutachtung nach bietet keines der Instrumente die Möglichkeit, die Daten mit *vergleichbaren* Klassen in Beziehung zu setzen. Alle drei Instrumente bieten nach meiner Einschätzung in den Ergebnismeldungen keine Hinweise für eine gezielte Förderung an, wenn in einzelnen Aufgabenbereichen die Ergebnisse ungenügend waren. Die Informationen sind nicht vielfältig aufbereitet, sondern nur in Form von Prozentwerten (Check 5, Klassencockpit), Mittelwerten und Leistungsgruppen (BeLesen) oder als Text (Orientierungsarbeiten). Die Kombination verschiedener Informationen, die sich ergänzen und auf verschiedene Ebenen übertragbar sind (Klasse und Schülerschaft) bieten die Formate ansatzweise an, es fehlt jedoch immer der Hinweis, wie mit den Informationen umzugehen ist.

7.4.2 Reflektion: Matrix der Differenziertheit der Instrumente

Die vier Instrumente werden verglichen in Hinblick auf die Differenziertheit der Information. Zurückgemeldete Ergebnisse können sowohl in Hinblick auf den Aspekt *Aufgaben*, als auch in Hinblick auf den Aspekt *Stichprobe* unterschiedlich differenziert ausgewertet werden.

Das heißt, die Ergebnisse können pro Aufgabe, pro Teilbereich (Geometrie, Division, Grammatik, etc.) oder pro Fach (Mathematik) aufgeschlüsselt werden: Differenziertheit des Aspektes Aufgaben.

Außerdem können die Ergebnisse in Werte der Gesamtstichprobe, Werte der Schule, Werte der Klasse und Werte des Schülers aufgeschlüsselt werden: Differenziertheit des Aspektes Stichprobe.

Abbildung 8 zeigt die vier beschriebenen Instrumente und die jeweils präsentierten Ergebnismeldungen. Auf der Matrix abgetragen wurde jeweils die differenzierteste Information, die *ohne weiterführende Analyse* durch den Leser abgelesen werden kann.

Ablesbar ist hier, dass Klassencockpit die Ergebnisse nur pro Fach auswertet, die Orientierungsarbeiten nur pro Aufgabe. Auswertungen pro Fach bedeutet, dass pro Fach die Ergebnisse des Tests dargestellt werden. Selbst wenn der Test ein spezifisches Feld abfragt, zum Beispiel Grammatik, werden die Ergebnisse bei Klassencockpit nur für diesen ganzen Test zurückgemeldet und nicht in Hinblick auf die Aufgabe a, b und c oder in Hinblick auf die Aspekte Kommasetzung oder Großschreibung. Check 5 und BeLesen bieten Auswertungen pro Teilbereich an. Keines der Instrumente bietet Informationen über

die Schule, Check 5 und Klassencockpit weisen die Ergebnisse der Gesamtstichprobe aus. Die Werte der Klasse und der einzelnen Schülerinnen und Schüler werden in allen vier Formaten präsentiert.

Matrix der Differenziertheit von Schulrückmeldungen

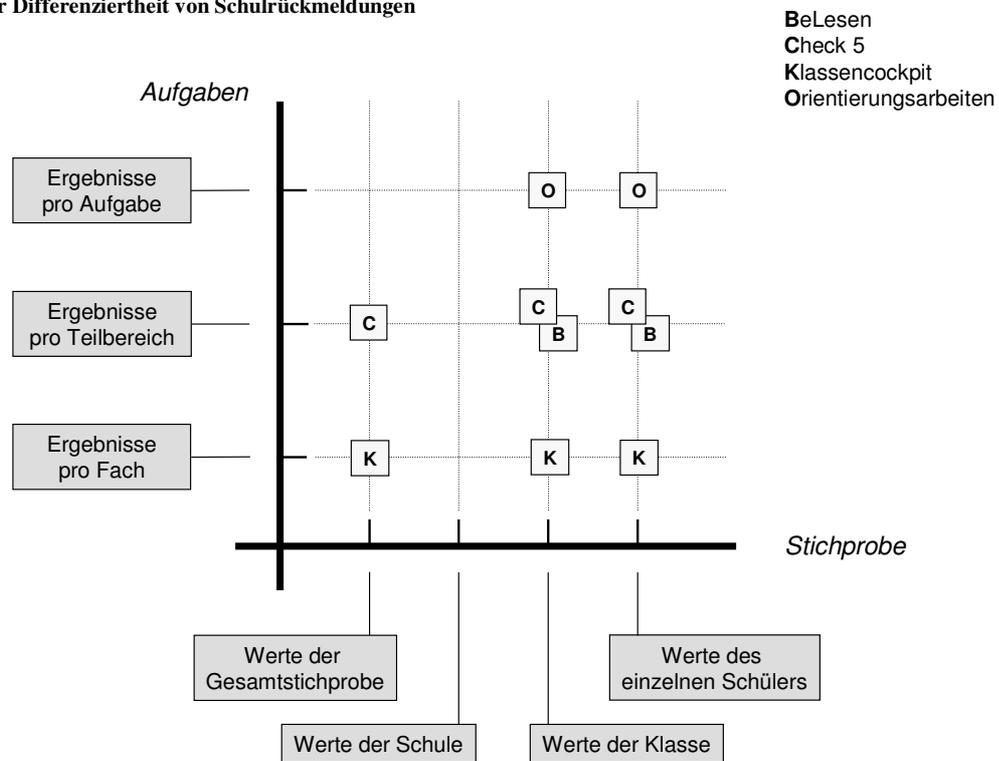


Abbildung 8: Vergleichende Darstellung der vier Formate in Hinblick auf den Grad der Differenziertheit bei Stichprobe und Aufgaben

7.4.3 Reflektion: Bezugsnormen

Die mit Klassencockpit, Check 5, den Orientierungsarbeiten der Zentralschweiz und BeLesen generierten Daten können zu der sozialen, kriterialen und individuellen Norm jeweils unterschiedlich in Beziehung gesetzt werden. Die **fettgedruckten** Inhalte in Tabelle 72 weisen immer darauf hin, dass diese Bezugsnormen den zurückgemeldeten Informationen inhärent sind. Die anderen Bezugsnormen können nur durch eine weitere Verarbeitung, die detailliertere Interpretation der Informationen oder eine erneute Datenerhebung hergestellt werden, sind aber nicht auf den ersten Blick ablesbar.

Tabelle 72: Vergleichende Darstellung der vier Formate in Hinblick auf die Bezugsnormen

<i>Bezugsnormen</i>	<i>Klassencockpit</i>	<i>Check 5</i>	<i>Orientierungsarbeiten</i>	<i>BeLesen</i>
Sozial	Leistung der Klasse wird mit einer Eichstichprobe verglichen.	Vergleich des eigenen Klassenmittelwertes mit anderen Klassen und mit einem Gesamtmittelwert.	Vergleich von Parallelklassen, die die gleichen Aufgaben lösen. (Eigenleistung der Lehrkraft)	Unterschiede zwischen den Ergebnissen von Kindern deutscher und nicht-deutscher Herkunft.
Kriterial	Ergebnisse verschiedener Module in Mathe oder Deutsch werden miteinander in Beziehung gesetzt. (Eigenleistung der Lehrkraft)	Leistungen der Schüler in Hinblick auf vorher festgelegte Mindestziele.	Überprüfung der Erreichung der deklarierten Lernziele.	Erreichen der Leistungsstufen in unterschiedlichen Lernbereichen.
Individuell, Ipsativ	Ergebnisse verschiedener Module werden über die Zeit hinweg betrachtet – klassenweise oder individuell. (Eigenleistung der Lehrkraft)	Nicht möglich, nur eine Messung vorgesehen.	Bei mehrmaligem Einsatz der gleichen Aufgaben oder des gleichen Teilbereichs (z.B. Mathe) kann der Lernfortschritt festgestellt werden. (Eigenleistung der Lehrkraft)	Entwicklung einzelner Kinder über die Messzeitpunkte hinweg in den Bereichen Lesen, Schreiben und Mathematik.

BeLesen bietet aufgrund seiner Anlage als Längsschnittuntersuchung alle drei Bezugsnormen an und ist somit in Hinblick auf den Aspekt Bezugsnormen die informativste Ergebnisrückmeldung. Klassencockpit legt den Fokus auf den sozialen Vergleich, die Orientierungsarbeiten legen den Schwerpunkt auf die

kriteriale Bezugsnorm. Diese beiden Formate bieten jeweils nur eine Referenz an, zu der die zurückgemeldeten Ergebnisse in Beziehung gesetzt werden. Check 5 bietet sowohl die kriteriale als auch die soziale Bezugsnorm an und schafft so die Möglichkeit, die Informationen in zweierlei Weise zu reflektieren.

7.4.4 Aktion: Nutzungsmöglichkeiten

Ausgehend von den in Kapiteln B_7.2 und B_7.3 genannten Berichten, Broschüren und Evaluationen, die es von den drei Schweizer Instrumenten gibt, wird im Folgenden gegenübergestellt, wofür die durch die drei Instrumente (Check 5, Klassencockpit und Orientierungsarbeiten) sowie im Rahmen der Tests in der Studie BeLesen generierten Informationen genutzt werden *können*. Etwas unscharf ist aufgrund einer bisher nicht erfolgten empirischen Vergleichsuntersuchung, ob die Informationen tatsächlich für den einen oder anderen Teilbereich (vergleichen, beurteilen, kontrollieren, kommunizieren, verbessern, fördern) genutzt *werden*.

In verschiedenen Evaluationsstudien (Landert, Farago Davatz & Partner 2002, Moser & Bucher 2003, Moser & Tresch, 2004, Moser 2003, Tresch & Moser 2005, Vögeli-Mantovani 2003) wurde die jeweilige Nutzung der einzelnen Instrumente untersucht und erläutert. Hierbei wurde teilweise auch quantifiziert wie die Instrumente genutzt wurden. Skalen, Fragen und Antwortformate unterschieden sich allerdings und es gibt keine Untersuchung, die alle drei Instrumente gleichermaßen untersucht. Daher wäre eine Aussage, die sich auf eine empirische Basis stützt, unzulässig. Die Erstellung der Tabelle erfolgte aufgrund der Dokumentenanalyse sowie aufgrund von Gesprächen und Diskussionen in der Bildungsdirektion.³⁰

Tabelle 73 zeigt, für welche Nutzung sich die jeweiligen Daten – nach der von mir vorgenommenen Dokumentenanalyse – eignen, für welche Nutzung sie nicht gewünscht sind und welche Nutzung nicht möglich ist.

Dabei ist zu beachten, dass es vorkommen kann, dass Lehrpersonen die Informationen der Instrumente auch für die Funktionen heranziehen, zu denen im Einzelnen keine Angabe gemacht worden ist bzw. für die die Informationen nicht gedacht waren.

³⁰ Die Gespräche waren fünf Interviews, die im Auftrag der Bildungsdirektion mit Lehrpersonen geführt wurden, die die Instrumente einsetzen. Die vollständigen Inhalte der Interviews sowie die Information, wer die Personen waren, sind nur der Bildungsdirektion zugänglich. Daher beziehe ich zwar die Informationen, die ich aus den Interviews ableiten kann, mit ein. Ich weise jedoch ausdrücklich darauf hin, dass die Tabelle keine empirische Grundlage hat, sondern Ergebnis einer explorativen, vergleichenden Dokumentenanalyse ist.

Ich möchte betonen, dass ich nicht davon ausgehe, dass die Informationen der Instrumente gleichermaßen für die Verwendung in den Bereichen vergleichen, beurteilen, kontrollieren, kommunizieren, verbessern und fördern geeignet sind.

In weiteren Forschungen wäre es wünschenswert, dass diese Grauzone zwischen Möglichkeiten der Nutzung und der tatsächlichen Nutzung detaillierter untersucht wird. Bis weitere empirische Untersuchungen vorliegen, lässt sich nur durch klare Kommunikation über Nutzbarkeit und Zielsetzung des Instruments die intendierte Nutzung der Informationen sicherstellen. Dem sind die Entwicklerinnen und Entwickler der Instrumente weitestgehend nachgekommen. Der Wortlaut in den einzelnen Kategorien, die in der Tabelle 73 abgebildet sind, orientiert sich an den in oben benannten Berichten und Broschüren gewählten Formulierungen. Darüber hinaus wird dargestellt, welche Verwendungsmöglichkeiten in Hinblick auf die im Projekt BeLesen generierten Ergebnisrückmeldungen möglich, nicht möglich oder nicht gewünscht sind. Für diese Analyse wurden die Rückmeldebroschüre BeLesen und die Aussagen des Forschungsteams dazu zu Grunde gelegt.

Die Tabelle 73 zeigt eine Übersicht der möglichen Verwendungsformen, wobei für jede Nutzung angegeben ist, ob die Nutzung *möglich und erwünscht* ist (✓), die Nutzung *nicht gewünscht* ist, weil die Entwicklerinnen und Entwickler der Formate dies nicht als Ziel ansehen (○), die Nutzung *nicht möglich* ist, da die Instrumente sich dazu nicht eignen (✗).

Tabelle 73: Nutzungsmöglichkeiten der vier Formate, unterschieden nach möglich & gewünscht, nicht gewünscht, nicht möglich

Nutzungsmöglichkeiten	Klassenscockpit	Check 5	Orientierungsarbeiten	BeLesen
<i>Vergleichen</i>				
Standortbestimmung	✓	✓	✓ nur in Hinblick auf die eigene Klasse	✓ nur in Hinblick auf die eigene Klasse
Vergleich mit anderen Klassen	✓	✓	✗ nicht möglich	✗ nicht möglich
<i>Beurteilen</i>				
Hilfsmittel zur Selektion der Schülerinnen und Schüler (Versetzungen, Empfehlungen für weiterführende Schulen, o.ä.)	○ nicht gewünscht	○ nicht gewünscht	○ nicht gewünscht	○ nicht gewünscht
Unterstützung bei Prognosen, Einschätzungen der Leistung einzelner Kinder, Förderung der Diagnosekompetenz	✓ Nur infolge des Engagements der Lehrperson	✓	✓	✓
Objektivierung der eigenen Beurteilungspraxis an einem externen Maßstab, Ergänzung und Anpassung der eigenen Schülerbeurteilung	✓	✓	✓ Umsetzung ziel- und kriterienorientierter Beurteilung	○ nicht gewünscht
Leistungsstand der Schülerinnen und Schüler	✓ Information über die Leistung, Bestätigung der Einschätzung des Leistungsstandes	✓ Zuverlässige, differenzierte Informationen über Schülerleistungen	✓ Lernkontrolle für jeden einzelnen Schüler	✓ Informationen über den Leistungsstand der einzelnen Schülerinnen und Schüler in Form von Leistungsgruppen

Nutzungsmöglichkeiten	Klassenscockpit	Check 5	Orientierungsarbeiten	BeLesen
Kontrollieren				
Überprüfung der Zielerreichung in Hinblick auf den Lehrplan	✓ x Je nach Lehrplan	✓	✓	x nicht möglich
Rechenschaftslegung gegenüber externen Personen oder Institutionen	○ nicht gewünscht	✓	○ nicht gewünscht	○ nicht gewünscht
Kommunizieren (die Ergebnisse können, müssen aber nicht zwingend Anlass für ein Gespräch sein)				
Grundlage für Elterngespräche	✓	✓	✓	✓
Grundlage für förderorientierte Gespräche mit Schülerinnen und Schülern	x nicht möglich	x nicht möglich	✓	x nicht möglich
Grundlage für Diskurse im Kollegium	✓	✓	✓	✓
Fördern				
Differenzierte Informationen als Grundlage für die individuelle Förderung	x nicht möglich	✓	✓	✓
Verbessern				
Professionalisierung: Hilfestellung zur professionellen Ausführung der Arbeit, Abbau von Unsicherheit, Selbstevaluation der Lehrtätigkeit	✓	✓	✓	✓
Qualitätssicherung durch Information, Grundlage für Unterrichtsentwicklung	✓	✓	✓	✓

In der Übersicht zeigt sich, dass die zurückgemeldeten Informationen der Standortbestimmung der eigenen Klasse dienen sollen. Der Vergleich mit anderen Klassen ist nur bei Check 5 und Klassencockpit möglich. Alle vier Instrumente sollen nicht für die Selektion der Schülerinnen eingesetzt werden, jedoch für andere Arten der Beurteilung. Die zurückgemeldeten Informationen sollen die Lehrpersonen bei der Einschätzung der Leistung der Kinder unterstützen, sie sollen die eigenen Beurteilungspraxis objektivieren und eine Beurteilung über den Leistungsstand der Schülerinnen und Schüler liefern.

Check 5 ist für die Kontrolle, also die inhaltliche Prüfung, ob die Lehrplanziele erreicht wurden sowie für die Rechenschaftslegung gegenüber externen Personen und Institutionen geeignet, diese Funktionen sind auch erwünscht. Bei BeLesen sind diese beiden Kontrollfunktionen nicht gewünscht bzw. aufgrund der Tests, die verwendet werden, nicht möglich. Bei den Orientierungsarbeiten kann die Erreichung der Lehrplanziele kontrolliert werden, die Rechenschaftslegung ist jedoch nicht gewünscht. Bei Klassencockpit verhält es sich ähnlich, die Überprüfung der Zielerreichung in Hinblick auf den Lehrplan ist – abhängig von der Übereinstimmung zwischen Test und Lehrplan – mal möglich, mal nicht.

Als Grundlage für Gespräche mit Eltern und Kollegen sind alle vier Formate geeignet, jedoch bieten nur die Orientierungsarbeiten Informationen an, die auch für ein förderorientiertes Gespräch mit den Schülerinnen und Schülern selbst brauchbar sind. In den drei anderen Formate BeLesen, Check 5 und Klassencockpit werden Werte zurückgemeldet, aber keine Inhalte. Die Rückübersetzung der Werte in fachliche Inhalte, auf denen dann ein förderorientiertes Gespräch aufbauen kann, muss von den Lehrpersonen selbst vorgenommen werden. Zur Förderung der individuellen Leistung sind alle Formate bis auf Klassencockpit geeignet. Bereits bei der vergleichenden Analyse der Formate in Hinblick auf die Gestaltung der Rückmeldungen wurde gezeigt, dass Verbesserungen bei der Aufbereitung der Informationen wünschenswert wären, um diese besser für die Förderung einsetzen zu können.

Aus der vergleichenden Analyse ist ablesbar, dass alle vier Formate, ausgehend von den Informationsbroschüren und Ergebnisrückmeldungen, zur Verbesserung in Hinblick auf Professionalisierung der Lehrperson und Qualitätssicherung geeignet sind. Wie bereits erläutert, wird hier keine Aussage dazu gemacht, ob die Rückmeldung dazu tatsächlich von den Nutzerinnen und Nutzern verwendet wird.

Nach der vergleichenden Analyse der vier Instrumente folgt nun die Gesamtdiskussion der drei vorgelegten Studien.