

Leistungsbeurteilung unter Berücksichtigung ihrer formativen Funktion

Elfriede Schmidinger, Franz Hofmann & Thomas Stern

Leistungsbeurteilung war bereits Thema eines Beitrags im Band 2 des Nationalen Bildungsberichts 2009 (Specht, 2009; Beitrag von Eder, Neuweg & Thonhauser, 2009). Die dort vollzogene Analyse zeigte u. a. auch die unterschiedlichen gesetzlichen Intentionen und Funktionen der in der Leistungsbeurteilungsverordnung (LBVO) vorgesehenen Instrumente zur Leistungsbeurteilung auf. Den nicht zu benotenden Informationsfeststellungen (§ 1 Abs. 2 LBVO) kommt demnach ausschließlich eine pädagogische Funktion zu („Setzen von Bedeutsamkeitssignalen, Feedback für Lehrpersonen und Schüler/innen, Motivierung“; Eder et al., 2009, S. 247), indem sie einerseits der Lehrperson die individuellen Lernstände ihrer Schüler/innen zeigen, sodass diese die weiteren unterrichtlichen Maßnahmen entsprechend anpassen kann und indem sie andererseits Grundlage für ein hilfreiches weiterführendes Feedback für die Schüler/innen sind. In dem Beitrag wird auch festgehalten, dass der pädagogischen Funktion am besten mit qualitativen Rückmeldungen, wie sie für die Informationsfeststellungen vorgesehen sind, entsprochen wird (Eder et al., 2009, S. 248). Dagegen erfüllen die zu benotenden Leistungsfeststellungen (§ 3 LBVO) neben der pädagogischen und der Berichtsfunktion primär gesellschaftliche Funktionen, wie die der Zuweisung zu verschiedenen Laufbahnen innerhalb und außerhalb der Schule (Allokationsfunktion; Eder et al., 2009, S. 247 f.) und die der Vergabe von Berechtigungen (Selektionsfunktion; Eder et al., 2009, S. 247), da die Noten aufgrund ihrer für die Betroffenen und Abnehmer/innen rasch überschaubaren Information zum zentralen Instrument im Selektionsgeschehen geworden sind (Eder et al., 2009, S. 248).

Zweckentsprechend sollten mit den beiden Instrumenten auch unterschiedliche Daten und Informationen in unterschiedlichen Phasen des Unterrichts erhoben werden. Die Informationsfeststellungen zielen auf die Prozesse *während* des Lernens im Unterricht ab, um diese positiv beeinflussen zu können, hingegen werden vor allem mit den punktuellen Leistungsfeststellungen bilanzierend die erreichten Lernergebnisse *nach* den für deren Erreichung aufgewendeten Lern- beziehungsweise Unterrichtseinheiten erfasst. Die Leistungsbeurteilung hat gemäß der LBVO daher sowohl eine lernförderliche formative als auch eine ergebnisorientierte summative Funktion. Der Lehrperson werden damit zwei verschiedene, durchaus antinomische Rollen zugeschrieben, die oft schwierig zu vereinen sind: einerseits die eines Sachverständigen, der in der summativen Funktion sein Fachurteil abgibt und andererseits die des lernunterstützenden Coachs in der formativen Funktion der Leistungsbeurteilung.

Im Beitrag von Eder et al. (2009) werden die Mängel bei der praktischen Umsetzung der rechtlichen Bestimmungen zu den Leistungsfeststellungen und -beurteilungen sowie Optionen zu deren Weiterentwicklung aufgezeigt. Von diesen Optionen wurde bis jetzt nur ein Bereich der dort empfohlenen „Revision der rechtlichen Grundlagen“ (Eder et al., S. 267) in Angriff genommen. Seit 2011 arbeitet eine Arbeitsgruppe des Bundesministeriums für Unterricht, Kunst und Kultur (BMUKK), jetzt Bundesministerium für Bildung und Frauen (BMBF), an der Erneuerung der LBVO (siehe Abschnitt 3.2.3). Den Forderungen nach „Externalisierung und Objektivierung“ (Eder et al., S. 265) wurde im Bereich der „Leistungsfeststellungen, mit denen Berechtigungen verbunden sind“ (Eder et al., S. 267) nur so weit entsprochen, dass die bereits 2009 laufenden Arbeiten an der Einführung einer *teilstandardisierten kompetenzorientierten Reife- und Diplomprüfung* im Schuljahr 2015/16 abgeschlossen

Formative und summative Funktion der Leistungsbeurteilung

werden. In diesem Schuljahr werden auch zum ersten Mal an allen allgemeinbildenden und berufsbildenden höheren Schulen (AHS und BHS) die Diplom- und Reifeprüfungen in dieser Form durchgeführt.

Anknüpfung an
den Nationalen
Bildungsbericht 2009

An den Nationalen Bildungsbericht 2009 anknüpfend nimmt der vorliegende Beitrag die formative Funktion vor allem der Informationsfeststellung im Rahmen der Leistungsbeurteilung und darüber hinaus als didaktisches Prinzip in den Fokus.

1 Formative Assessment

1.1 Wurzeln des Formative Assessment

Im englischsprachigen Raum gibt es seit den 1960er Jahren einen wissenschaftlichen und bildungspolitischen Diskurs um die formative Leistungsbewertung (FLB).¹

Über 50 Jahre formatives
Assessment

Den Anstoß gab Scriven bereits 1967. Er beschrieb *Formative Evaluation* als ein Instrument zur Curriculumsentwicklung durch ständige Begleituntersuchungen ihrer Auswirkungen und darauf folgende Nachbesserungen (Scriven, 1967, S. 41). Bloom (1969, S. 48) verwendete denselben Begriff in seiner Theorie des zielreichenden Lernens (*Mastery Learning*) erstmals im Zusammenhang mit Lerndiagnosen zur Steuerung des Unterrichtsprozesses (*Classroom Assessment*). Formative Evaluationen waren für ihn kurze Tests, um alle Schüler/innen regelmäßig mit dem für das Erreichen der Lernziele notwendigen Feedback zu versorgen, und zwar – im Unterschied zu den „summativen“ Leistungsfeststellungen – ohne Benotung.

In Großbritannien setzte sich in der Folge anstelle von *Formative Evaluation* der Begriff *Formative Assessment* durch. Als verantwortlich für deren effektive Nutzung sah man unter dem Einfluss der konstruktivistischen Lerntheorie nicht allein die Lehrperson, sondern auch die Lernenden selbst. Diese müssen aktiv in der Lage sein, die Qualität ihres Lernens anhand der Rückmeldungen zu reflektieren und ihren Lernstand selbst einzuschätzen, um ihr weiteres Lernen dementsprechend regulieren zu können. Eine Voraussetzung dafür sind klare Zielvorstellungen der Lehrpersonen und der Lernenden.

1.2 Definitionen des Formative Assessment

Ihrer bahnbrechenden Metastudie „Inside the Black Box“, die die positiven Auswirkungen des Formative Assessment auf das Lernen aller, vor allem aber der leistungsschwächeren Schüler/innen zeigte, legten Black und Wiliam (1998) folgende Definition zugrunde:

„We use the general term assessment to refer to all those activities undertaken by teachers – and by their students in assessing themselves – that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs“ (Black & Wiliam, 1998, S. 140).

Diese Definition betont, dass sich die formativen Bewertungsmaßnahmen insbesondere durch die Erfüllung ihres Zwecks, der Anpassung der unterrichtlichen Maßnahmen an die aktuellen Bedürfnisse der Lernenden, von den summativen Beurteilungen unterscheiden, jedoch nicht unbedingt in ihrer Form. Es können vor dem Hintergrund dieser Definition daher auch summative Tests formativ genutzt werden und umgekehrt. Diese Definition von Black und Wiliam (1998) lässt, wie man an diesem Beispiel sieht, vieles offen, nichtsdestoweniger wurde sie zur Grundlage vieler späterer Arbeiten zum Formative Assessment

1 *Formative Assessment* wird in diesem Beitrag mit *formativer Leistungsbewertung* (FLB) übersetzt.

(z. B. Cowie & Bell, 1999; Shepard et al., 2005), die zu einer vielfältigen Praxis führten, in der jedoch oft unklar blieb, was man unter Formative Assessment verstand.

Die folgende Definition möchte klarstellen, dass FLB als Unterrichtsprozesse zu sehen sind und nicht als Tests: „Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcomes“ (Heritage, 2010, S. 9).

Broadfoot et al. (2002) plädierten in dieser Situation für den Begriff *Assessment for Learning* (Leistungsbewertung, die dem Lernen dient), der heute synonym für das Formative Assessment verwendet wird, im Gegensatz zum *Assessment of Learning* (summativ Leistungsbewertung des Lernertrags). „Assessment for Learning as the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there“ (Broadfoot et al., 2002, S. 2 f.).

Assessment for Learning

In ihren kritischen Reviews dieser und anderer Arbeiten argumentieren Dunn und Mulvenon (2009) sowie Bennett (2011), dass nicht die Leistungsfeststellung per se formativ oder summativ sein kann, sondern nur der Zweck ihrer Rückmeldung. Wird das erreichte Leistungsniveau gemessen und werden die Schüler/innen über ihren Leistungsstand informiert, hat die Leistungsfeststellung primär einen summativen Zweck, kann aber zusätzlich – sekundär – auch einzelnen Lernenden manchen Fehler aufklären und damit auch formativ wirken. Werden jedoch aus der Leistungsfeststellung Schlüsse gezogen, die der Planung der nächsten Unterrichtsaktivitäten dienen, wirkt sie primär formativ, trotzdem erfahren Lehrpersonen und Schüler/innen – sekundär – auch etwas über den Lernstand. Bennett (2011, S. 7 f.) meint daher, dass Leistungsfeststellungen, wenn sie gut gemacht sind, eine der beiden Funktionen primär gut erfüllen können sollen, die jeweils andere jedoch sekundär auch.

Zusammenschau

FLB ist eine komplexe Intervention im Unterricht. Sie beschränkt sich nicht auf die Informationsfeststellung und -bewertung sowie auf deren Rückmeldung, sondern verbindet sie mit individuellen Anregungen zum Weiterlernen und Überlegungen zur ständigen Adaptierung des Unterrichts an die Schülerbedürfnisse. Grundlage der Bewertungen sind sowohl die *Kriterial-* als auch die *Individualnorm*, indem sie für jede Schülerin und jeden Schüler die jeweilige Lernausgangslage in die Bewertung einbezieht. FLB ist dann wirksam, wenn sie so genau wie möglich erfasst, was die Lernenden schon können und was noch nicht. Sie ist grundsätzlich förderorientiert, denn für sie gilt, dass sie dem Lernen der Schüler/innen aller Leistungsniveaus dienen muss, nicht ihrer Klassifizierung (Stern, 2010). Sie erfordert einen *adaptiven Lehrstil* (Helmke, 2014).

1.3 Ordnungsversuche

Shavelson et al. (2008, S. 300 f.) kategorisieren die verschiedenen Formen FLB auf einem Kontinuum von informell bis formell:

- *On-the-Fly Formative Assessment* ist ungeplant – z. B. Beobachtungen einer Gruppenarbeit durch die Lehrperson mit einer Intervention zur Überwindung momentaner Schwierigkeiten oder die Besprechung einer fehlerhaften Aufgabenlösung mit einem Kind.
- *Planned-for-Interaction Formative Assessment* wird bereits bei der Unterrichtsplanung entwickelt – z. B. Vorbereitung einiger Fragen, die die Lehrperson mit der Klasse diskutieren möchte, um zu sehen, was die Schüler/innen bereits wissen und verstanden haben und was noch weiter bearbeitet werden muss.
- *Embedded-in-the-Curriculum Formative Assessment* ist eine von Testinstituten bereitgestellte diagnostische Aufgabensammlung mit Auswertungsanleitungen. Die Klasse führt einen solchen Test durch, sobald wesentliche Lernziele erreicht sein sollten. Die Testergebnisse zeigen, was die Schüler/innen schon können und was ihnen noch fehlt, sodass die Lehrperson ihr weiteres Vorgehen auf der Basis dieser Daten abstimmen kann.

Formen der FLB

Die beiden erstgenannten Formen sind Situationen, in denen die Lehrperson Schüleraktivitäten beobachtet und interpretiert, um Feedback zu geben. Dies soll in jeder Unterrichtsstunde (*short-cycled*) „minute-by-minute, day-by-day“ geschehen, wie es Leahy, Lyon, Thompson und Wiliam (2005) schon im Titel eines ihrer Artikel einfordern.

Die Curriculum-Embedded Formative Assessments werden dagegen bloß als standardisiertes Messinstrument eingesetzt, in den USA immer häufiger auch computerunterstützt. Diese Tests sind formativ, weil die Überprüfung nicht der Benotung, sondern dem Feedback dient, sie jedoch nur alle paar Wochen oder nur einige Male im Jahr (*long-cycled*) eingesetzt werden. Es fehlt also der unmittelbare Zusammenhang zwischen Lernprozess und Ergebnisrückmeldung zur Unterstützung des Lernens. Daher konnten für solche Settings empirisch keine Leistungsverbesserungen nachgewiesen werden (Heritage, 2010). Die Curriculum-Embedded-Tests wirken wie „mini-summative“ Tests. Auch Shavelson et al. (2008) selbst mussten in ihrer Studie feststellen, dass die Curriculum-Embedded Assessments nur schlecht in den Unterricht zu integrieren waren und keine signifikanten Auswirkungen auf die Schülerleistungen hatten. Die Autorinnen und Autoren zogen daraus die Konsequenz, die Tests auf wenige offene Verständnisaufgaben zu beschränken. Anschließend sollten die Schüler/innen ihre individuellen Lösungen erst in Kleingruppen, dann im Klassenplenum mit der Lehrperson diskutieren. Die Überprüfungen sollten sich auf Lernaktivitäten stützen, in denen das Denken der Schüler/innen explizit beobachtet werden kann, und sollten auch Raum für Selbst- und Peer-Assessments enthalten. Dieses Lernarrangement nannten sie *Reflective Lesson* (Ayala et al., 2008).

Effektivität der Reflective Lesson

Die Effektivität dieser Modifikation des Curriculum-Embedded Formative Assessments konnten Yin, Tomita und Shavelson (2015) in einem Experiment zum Thema „Sinken und Schwimmen“ mit zwei Klassen nachweisen. In einer der beiden Klassen wurde On-the-Fly Formative Assessment eingesetzt, in der anderen zusätzlich ein Curriculum-Embedded Formative Assessment auf der Grundlage eines von Curriculum- und Assessmentexperten entworfenen Lernentwicklungsplans, der auch Reflective Lessons enthielt. Die Leistungstestergebnisse der Schüler/innen der Experimentalklasse waren signifikant besser, ihnen war es auch besser gelungen, die anfänglich vorhandenen Fehlvorstellungen zu überwinden.

1.4 Merkmale einer effektiven formativen Leistungsbewertung

Leahy et al. (2005) empfehlen, bei der Planung effektiver Lehr-Lern-Prozesse von folgenden drei Leitfragen auszugehen: Wohin geht die/der Lernende, was ist das Lernziel? Wo befindet sich die/der Lernende auf ihrem/seinem Weg zum Ziel? Wie kommt sie/er dorthin? Effektive FLB erfordert die Umsetzung nachstehender fünf Schlüsselstrategien durch die drei Akteure Lehrperson, Peers und Lernende, die sich auf diese Leitfragen beziehen (siehe Abbildung 2.1):

Fünf Schlüsselstrategien

1. *Lernziele und Erfolgskriterien mit den Lernenden klären:* Die Lehrperson entwickelt mit ihren Schülerinnen und Schülern am Beginn einer größeren Unterrichtseinheit ein gemeinsames Verständnis von den Lernzielen und deren Erfolgskriterien.
2. *Effektive Diskussionen, Aktivitäten und Aufgaben arrangieren, die zu beobachtbaren Lernergebnissen führen:* Die Lernaktivitäten liefern Daten, die sichtbar machen, wie die Schüler/innen denken. Sie zeigen, wie sie eine Aufgabe gelöst haben und warum sie so vorgegangen sind.
3. *Feedback geben, das das Lernen voranbringt:* Wiliam (2013) stellt klar, dass kein Feedback allen Lernenden hilft. Was den einen weiterbringt, wirft einen anderen zurück. Effektives Feedback setzt voraus, dass die Lehrperson ihre Schüler/innen aufmerksam beobachtet, deren Bedürfnisse gut kennt und ihr Vertrauen besitzt.
4. *Schülerinnen und Schülern ermöglichen, einander als Ressourcen zu nutzen:* Kooperative Lernformen, Selbsteinschätzung und Peer-Feedback sind wesentliche Elemente von FLB. Die Selbsteinschätzung ist für die Selbstregulation des Lernens und die Übernahme der Verantwortung für das eigene Lernen eine notwendige Kompetenz. Sie wird durch Anre-

gungen zur Selbstreflexion des Lernens und durch Feedbacks der Lehrperson sowie der Peers unterstützt. Das hilft den Lernenden, eine aktive Rolle beim Planen, Überwachen und Evaluieren ihrer Lernfortschritte einzunehmen. Dabei erwerben sie metakognitive Strategien, mit denen sie ihr Lernen verbessern können.

5. *Schüler/innen als Verantwortliche ihres eigenen Lernens anerkennen*: Die Lernenden müssen die Erfahrung machen können, dass sich Anstrengung lohnt, sie kreativ sein können, Fehlermachen Teil des Lernens ist und sie Verantwortung für ihr eigenes Lernen übernehmen sollen.

Abb. 2.1: Fünf Schlüsselstrategien der formativen Leistungsbewertung

	Lernziel (Wohin die/der Lernende geht)	Lernstand (Wo die/der Lernende steht)	Lernweg (Wie sie/er dorthin kommt)
Lehrperson	Lernziele und Erfolgskriterien mit den Lernenden klären	Effektive Diskussionen, Aktivitäten und Aufgaben arrangieren, die zu beobachtbaren Lernergebnissen führen	Feedback geben, das das Lernen voranbringt
Peer		Schülerinnen/Schülern ermöglichen, einander als Ressourcen zu nutzen	
Lernende		Schüler/innen als Verantwortliche ihres eigenen Lernens anerkennen	

Anmerkung: Übers. v. Verf.

Quelle: William (2014, S. 8).

An der University of California, Los Angeles (UCLA) wird seit 2006 in einem nationalen Großprojekt *Formative Assessment for Students and Teachers* (FAST SCASS) als wesentlicher Stimulus für die Verbesserung des Lehrens und Lernens propagiert. FAST SCASS (2008) publizierte im Rahmen dieses Projekts fünf Merkmale für eine effektive FLB, die größtenteils mit den fünf Schlüsselstrategien übereinstimmen:

1. Lernentwicklungspläne (*Learning Progressions*)
2. Lernziele und Erfolgskriterien (*Learning Goals and Criteria for Success*)
3. Deskriptives Feedback (*Descriptive Feedback*)
4. Selbst- und Peereinschätzung (*Self- and Peer-Assessment*)
5. Kooperation (*Collaboration*)

Heritage (2008) sieht in der Ausrichtung der FLB auf fachbezogene Lernentwicklungspläne ein zentrales Qualitätsmerkmal. Anders als die horizontal an den Wissensdomänen orientierten Curricula sind die Lernentwicklungspläne (*Learning Progressions*) vertikal strukturiert, d. h. sie beschreiben für die zu erwerbenden Kompetenzen Lernsequenzen, die zu zunehmend anspruchsvolleren Denkweisen in komplexer werdenden, aufeinander aufbauenden Thematiken führen (Heritage, 2008; Wilson & Bertenthal, 2005). Sie erfordern einen kohärenten Lehrplan, der auf langfristige Bildungsziele ausgerichtet ist und in dem Wissen, Fähigkeiten und Fertigkeiten miteinander verknüpft sind. Eine solche Entwicklungsperspektive regt die Lehrpersonen an, Lernen als einen Prozess mit individuell unterschiedlichen Lerngeschwindigkeiten zu verstehen, der zu einem zunehmend differenzierteren Verständnis der Sache führt und nicht bloß zur Bewältigung der für die Schulstufe vorgesehenen Stoffmenge. Aufgrund der klaren Reihenfolge seiner Teilziele zeigt der Lernentwicklungsplan der Lehrperson für jede Lernende/jeden Lernenden den nächsten Schritt in der *Zone der nächstmöglichen Entwicklung* (Vygotskij, 2002). Solche Lernentwicklungspläne zu konzipieren ist jetzt Aufgabe der Lehrperson. Hilfreich wären dazu exemplarische Vorlagen – wie sie aktuell in Österreich im Zusammenhang mit der Entwicklung von Kompetenzstrukturmodellen und

Lernentwicklungsplan

Kompetenzrastern als Grundlage für Leistungsfeststellungen geplant sind (siehe Abschnitt 3.2.3) –, die von Expertengruppen aus dem Bereich der Fachdidaktik, der Lerntheorie und der Curriculumentwicklung zur Verfügung gestellt werden (Heritage, 2008).

1.5 Feedback als zentrales Element der formativen Leistungsbewertung

Aufgrund der empirischen Evidenz wird heute die FLB als eine Unterrichtsintervention und weniger als ein Messinstrument aufgefasst, in der die Lehrperson nach Möglichkeit gemeinsam mit ihren Schülerinnen und Schülern Ziele setzt, diese nach einer erfolgten Instruktion beim Lernen beobachtet und ihnen mit ihrem Feedback hilft, die Ziele zu erreichen. Ein zentrales Element bildet darin das Feedback, zu dem sehr viele Studien, jedoch mit inkonsistenten Ergebnissen vorliegen. Feedback wird aufgrund der aktuellen Forschungslage als ein mehrdimensionales und -funktionales Konzept verstanden, in dem verschiedene Feedbacktypen und ihre Rolle im Lernprozess identifiziert werden können, und das auch erkennen lässt, unter welchen Bedingungen sich Feedback auf die Lernergebnisse positiv auswirkt (Hattie & Wollenschläger, 2014; siehe auch Abschnitt 2.1.2). Damit können die unterschiedlichen Wirkungen eines Feedbacks in verschiedenen Settings und bei verschiedenen Personen erklärt werden.

Aus konstruktivistischer Sicht ist das Feedback ein Hilfsangebot an die Lernenden, um sie bei der Konstruktion von Wissen im Umgang mit komplexen Problemen zu unterstützen (vgl. Zierer, Busse, Otterspeer & Wernke, 2015). Auch ein Feedback kann summativ (z. B. eine verbale Beurteilung im Zeugnis) und formativ sein. Formatives Feedback wird von manchen Autorinnen und Autoren, wie auch von Hattie (2013), mit der FLB gleichgesetzt.

Feedback ist einer der
wirkmächtigsten Faktoren
im Unterricht

Hattie (2013) hat in seiner Studie „Lernen sichtbar machen“ u. a. auch eine Zusammenschau des Forschungsstands zum Thema Feedback in Form von Metaanalysen über 1.287 Studien vorgelegt. Obwohl Metaanalysen schwierig zu interpretieren sind, da die zugrundeliegenden Studien sehr unterschiedliche Kontexte aufweisen (Unterschiede in Schulstufen, Unterrichtsgegenstände, Schulformen, Staaten u. a.), ist die Effektstärke von $d = 0,73$, die Hattie für das Geben und Erhalten von Feedback errechnet hat, ein nicht zu ignorierender Hinweis auf die Wirkmächtigkeit dieses Faktors, der jedoch aufgrund der oben angeführten Inkonsistenz der Einzelergebnisse eine große Streubreite aufweist. Häufige, kurze Rückmeldungen (*Rapid Formative Assessments*), die die Lehrperson zu ihren unterrichtlichen Maßnahmen einholt, sind laut Hattie (2012, S. 142 ff.) besonders wirkungsvoll. Dies bestätigt die Ergebnisse der qualitativen Literatursynthese zum Formative Assessment von Black und Wiliam (1998), in der sie Werte zwischen $d = 0,4$ und $d = 0,7$ fanden. Klare und für alle transparente Ziele sind für Hattie eine wesentliche Voraussetzung für ein gelingendes Feedback. Die Effektstärken von $d = 0,75$ für „Klarheit der Lehrperson“ und $d = 0,56$ für „Klarheit der Ziele“, die Hattie gefunden hat, sind damit auch eine Bestätigung der Wirksamkeit der Schlüsselstrategie (Leahy et al., 2005) und des Qualitätsmerkmals „Lernziele und Erfolgskriterien gemeinsam klären“ (FAST SCASS, 2008).

Hattie verweist auch auf den dialogischen Charakter wechselseitigen Feedbacks. Das Erfassen von Lernfortschritten ist nicht nur die Grundlage für lernförderndes Feedback an die Schüler/innen, sondern ermöglicht auch der Lehrperson die Einschätzung der Wirksamkeit ihres Unterrichts, ist also auch ein Feedback der Lernenden an sie. Um den Lernerfolg zu steigern, empfiehlt Hattie (2013) den Lehrpersonen weiters, ihren Unterricht *mit den Augen der Lernenden* – also *lernseits* (Schratz & Westfall-Greiter, 2010) – zu betrachten. Dieses Einfühlen in die Schüler/innen macht deren Lernprozesse „sichtbar“ und ermöglicht den Lehrpersonen, die entsprechenden Schlussfolgerungen aus den Evidenzen zu ziehen.

Feedback auf der
Aufgaben-, Prozess- und
Selbstregulationsebene

Hauptzweck des Feedbacks sollte die Verringerung der „Diskrepanzen zwischen dem aktuellen Verständnis und der Leistung auf der einen Seite und einer Lernintention oder einem Ziel auf der anderen Seite“ (Hattie, 2013, S. 208 f.) sein. Um dies zu erreichen, muss das

Feedback Informationen zu den drei Leitfragen zum Lernziel, Lernstand und Lernweg bieten (siehe Abschnitt 1.4, Abbildung 2.1). Hattie (2013, S. 210) unterscheidet vier Ebenen, auf denen das Feedback zu diesen Leitfragen operieren kann: die Ebene der Aufgaben, der Verstehensprozesse, der Selbstregulation und der Person. Während Feedback auf der Aufgabenebene eher Oberflächen-Lernen unterstützt, hilft Feedback auf der Prozessebene, die Lernprozesse zu optimieren. Am effektivsten erwies sich das Feedback auf der Ebene der Selbstregulation mit Informationen zur Steuerung der Lernprozesse, da damit die metakognitiven Fähigkeiten der Lernenden gefördert werden. Feedback auf diesen drei Ebenen ist besonders lernförderlich, wenn es mit Anregungen und Tipps für die nächsten Lernschritte verbunden ist.

Auf der persönlichen Ebene Lob zu spenden erwies sich hingegen als ineffektiv. Selbst wenn zusätzliche inhaltliche Informationen zum Lob gegeben werden, schwächt ein Lob diese Information ab. Als Feedbackformen mit geringer Lernwirksamkeit erwiesen sich neben dem Lob auch die Bestrafung, kontrollierendes Feedback und der programmierte Unterricht. Ebenso waren extrinsische Belohnungen auf lange Sicht kontraproduktiv, weil sie „die Übernahme eigener Verantwortung für die Selbstmotivierung und -regulierung verringern“ (Hattie, 2013, S. 208).

Feedback auf der Ebene des Selbst

Um die Lernenden zu aktivieren und das Lernen zu verbessern, muss das Feedback klar, zweckgerichtet, sinnvoll und verständlich sein. Auch wenn anspruchsvolle Ziele angestrebt werden, sollten Schwierigkeitsgrad und Komplexität der Aufgaben nicht zu hoch sein. Niemand darf in seinem Selbstwertgefühl bedroht werden. Eine besonders für die FLB relevante Zusammenstellung der Merkmale eines effektiven Feedbacks bietet Chappuis et al. (2012, zitiert nach Winter, 2015, S. 109):

Effektives Feedback

Effektives Feedback

- richtet die Aufmerksamkeit auf das beabsichtigte Lernen, wobei die Stärken herausgehoben werden und spezifische Information angeboten wird, die eine Verbesserung anleiten kann,
- erfolgt während des Lernens, solange noch Zeit bleibt, daran zu arbeiten,
- richtet sich auf teilweise Verstandenes,
- nimmt den Schülerinnen und Schülern nicht das Denken ab,
- begrenzt die korrektive Information so, dass die Schülerin oder der Schüler damit noch umgehen kann.

1.6 Theoriezugänge zur formativen Leistungsbewertung

William (2014) selbst bringt die FLB in Beziehung zum *selbstregulierten Lernen*, um es besser theoretisch fundieren zu können. „Selbstreguliertes Lernen ist eine Form des Lernens, bei der die Person in Abhängigkeit von der Art ihrer Lernmotivation selbstbestimmt eine oder mehrere Steuerungsmaßnahmen (kognitiver, metakognitiver, volitionaler oder verhaltensmäßiger Art) ergreift und den Fortgang des Lernprozesses selbst überwacht. Aus dieser Definition folgt, dass Theorien der Selbststeuerung des Lernens sowohl die kognitive Seite des Lernprozesses als auch die motivationale und volitionale Seite thematisieren müssen“ (Schiefele & Pekrun, 1996, S. 258). Die wesentlichen Komponenten eines jeden Lernvorgangs sehen Schiefele und Pekrun (1996, S. 250) in

Theorie des selbstregulierten Lernens

1. der Zielstellung (Elaboration der Ziele, die mit einer Lernepisode erreicht werden),
2. den Lernoperationen (Lern- und Regulationsaktivitäten) und
3. den zielorientierten Kontrollprozessen (Selbstbewertung der Ergebnisse).

Selbststeuerung bedeutet die aktive Beeinflussung dieser Lernprozesskomponenten durch die Lernenden selbst. In Korrelationsstudien zeigten sich positive Zusammenhänge mit den Lernergebnissen.

Die fünf Schlüsselstrategien einer effektiven FLB (Leahy et al., 2005; siehe auch Abschnitt 1.4) haben eine solche Selbststeuerung zum Ziel. Die gemeinsame Klärung der Lernziele und der Erfolgskriterien sowie die Anregung zur Selbsteinschätzung der Lernfortschritte als komplementäre Ergänzung der Lehrer- und Peer-Feedbacks im Rahmen der FLB (siehe Abschnitt 1.5) ermöglichen und unterstützen die Selbststeuerung des Lernens der Schüler/innen. Sie fördern auch ein bewusstes Kompetenz- und Autonomieerleben sowie deren Reflexion. Dies bewirkt nach der Selbstbestimmungstheorie von Deci und Ryan (1993) intrinsische Motivation, aber auch ein positives Selbstkonzept und in der Folge verbesserte Leistungen. Eine weitere Erklärung findet Chan (2006, zitiert nach Zierer, Busse, Otterspeer & Wernke, 2015) in seiner Studie zu Feedback in Versagenssituationen: Formatives Feedback erhöht die Selbstwirksamkeitsüberzeugungen und damit indirekt die intrinsische Motivation.

Weitere theoretische Zugänge

Aus *lerntheoretischer Perspektive* ermöglicht FLB ein kontinuierliches Arbeiten in der Zone der nächstmöglichen Entwicklung. Vygotskij (2002) unterscheidet die tatsächlich erfolgte Entwicklung, die man an der eigenständigen Leistung eines Kinds erkennt, von der potenziellen Entwicklung, bei der sie nur mit fremder Hilfe gelingt. FLB hilft durch das Beobachten, Sammeln und Interpretieren der Lernevidenzen der Schüler/innen zu erkennen, wie Lernhilfen allmählich reduziert werden können, sodass die Leistung selbstständig erbracht werden kann (*Scaffolding*).

Aus *konstruktivistischer und soziokultureller Perspektive* wird die FLB als reziproke Aktivität in einem dynamischen Prozess zwischen Lehrperson und den Schülerinnen sowie Schülern betrachtet, wobei der Kommunikation eine Schlüsselfunktion zukommt. Die Lehrpersonen nützen die Evidenzen des Lernens zum Scaffolding und Feedback, die Lernenden zur Selbstregulierung ihres Lernens. Alle Beteiligten dieser Lerngemeinschaft sollen dabei Ziele, Praktiken und Normen übernehmen, die das Lernen unterstützen, indem sie in einer Atmosphäre des Vertrauens, des Respekts und der Zusammenarbeit Lernergebnisse erfassen und interpretieren sowie Feedback geben (Heritage, 2010).

Auch wenn die Prinzipien der FLB für alle Fächer gelten, müssen sie in der Praxis fachspezifisch umgesetzt werden, was vielfach zu wenig beachtet wird (Bennett, 2011). Es sind daher auch fachdidaktische Theorien zur Begründung der FLB heranzuziehen.

2 Formative Leistungsbewertung im deutschsprachigen Raum

Formative Assessment wurde in der deutschsprachigen wissenschaftlichen Literatur unterschiedlich übersetzt: *ganzheitliches Beurteilen* (Roos, 2001), *förderorientierte Beurteilung* (Roos, 2001; Smit, 2009), *förderliche Leistungsbewertung* (Stern, 2010), *formative Beurteilung* (Smit, 2009), *informatives tutorielles Feedback* (Narciss, 2006; zitiert nach Narciss, 2014), *Lernverlaufsdiagnostik* (Klauer, 2014), *formative Leistungsdiagnostik* (Maier, 2014) u. a. m. Hier wird dafür weiterhin FLB verwendet.

2.1 Studien auf der Basis des englischsprachigen Forschungsstands

2.1.1 Formative Leistungsbewertung mithilfe von Tests

Curriculum-Based Measurement

2006 griff Klauer in einem Artikel über „Curriculum-Based Measurement“ (CBM) die US-amerikanischen Forschungen auf (siehe Klauer, 2014). Er beschreibt darin, wie Lehrpersonen fachlehrplanbezogene Arbeits- und Testblätter, die von Testunternehmen angeboten werden, im Unterricht auch computergestützt einsetzen (Voß & Hartke, 2014), um für die gezielte Förderung das Ausmaß der bewältigten Lernziele für jedes Kind festzustellen. Dies führte in Deutschland zur Entwicklung von formativen, auf probabilistischen Modellen basierenden Tests für Lesen (Diehl & Hartke, 2012; Walter, 2010, 2013), Rechtschreiben (Kuhlmann & Hartke, 2011; Strathmann, Klauer & Greisbach, 2010) und Mathematik für

die Sonderpädagogik und die Grundschule (Strathmann & Klauer, 2012). Aus den Aufgabenpools können die Lehrpersonen in relativ kurzen Abständen (etwa alle zwei bis vier Wochen) formative Tests zusammenstellen, deren Ergebnisse die individuellen Lernverläufe der Schüler/innen abbilden und damit ein Feedback über die verwendeten Lehrmethoden und -strategien bieten, die Anpassungen des Unterrichts und individuelle Fördermaßnahmen ermöglichen. Diese Angebote sollen den Lehrpersonen die Arbeit erleichtern, bergen aber auch die Gefahr der Deprofessionalisierung der diagnostischen Kompetenzen des Lehrpersonals in sich.

Das CBM ist auch eine wesentliche Komponente im Rahmen des *Response-to-Intervention* (RTI)-Ansatzes, der im *Rügener Inklusionsmodell* (RIM) mit monatlichen CBM und zwei bis drei Screenings im Schuljahr erprobt wird (Mahlau et al., 2014). Mit den Daten aus diesen Testungen wird die Wirksamkeit der binnendifferenzierten Maßnahmen im Klassenunterricht evaluiert (Förderebene I). Wenn darin der Lernverlauf eines Kinds zu stagnieren scheint, wird dies durch die Lehrperson überprüft und bei Bedarf eine intensivere Förderung in einer Kleingruppe geboten, deren Effektivität mit wöchentlichen CBM evaluiert wird (Förderebene II). Bei Lernerfolg wird diese zusätzliche Förderung beendet, andernfalls gibt es weiteren Kleingruppen- oder Einzelunterricht durch eine Sonderpädagogin oder einen -pädagogen (Förderebene III). Die Universität Rostock begleitet das RIM-Projekt wissenschaftlich mit Konzeptentwicklung, Lehrerfortbildung und Projektevaluation. Seit 2013/14 arbeiten alle Klassen der Grundschule auf Rügen auf diese Art, wodurch nur 1,5 % der Schüler/innen des Einschulungsjahrgangs im Vergleich zu 11 % der Kontrollgruppe nicht in regulären Grundschulklassen unterrichtet werden. Dass die Leistungssteigerungen bisher bescheidener waren, als amerikanische Studien erwarten lassen, wird einerseits auf den außerordentlich hohen Anteil von Kindern mit sonderpädagogischem Förderbedarf zurückgeführt, andererseits auf die geringe Erfahrung der Lehrpersonen mit der Nutzung der diagnostischen Daten für ein korrekatives Feedback und eine tutorielle Unterstützung.

Diese Tests werden zwar als Methoden beziehungsweise Instrumente einer *formativen Leistungsdiagnostik* (Klauer, 2014; Maier, 2014) angeboten. Sie eignen sich jedoch nur zur Evaluation der Effektivität des eigenen Unterrichts, weniger für formative Rückmeldungen an die Schüler/innen, die vor allem von einer pädagogischen, förderorientierten Haltung der Lehrperson abhängen und nicht auf Testformate reduziert werden können.

2.1.2 Experimentelle Studien

Souvignier, Förster und Schulte (2014) untersuchten die Wirksamkeit der regelmäßigen formativen Erfassung der Lesekompetenz auf der 4. Schulstufe (basale Lesefertigkeiten und sinnerfassendes Lesen) mit einem internetbasierten Testprogramm, das sofort automatisiert das Testergebnis für jedes Kind sowie die Klassenstatistik liefert. Laut Studie führte bereits die Rückmeldung diagnostischer Informationen zu signifikanten Lernzuwächsen beim Lesen, obwohl nur zwei Drittel der Lehrpersonen angaben, diese für pädagogische Entscheidungen genutzt zu haben.

Im groß angelegten, mehrjährigen deutschen Forschungsprojekt „Conditions and Consequences of Classroom Assessment“ (Co²CA) einer Forschergruppe um Eckhard Klieme wurden in einer Vorstudie zu den Experimenten sowohl Lehrpersonen als auch Schüler/innen der 9. Schulstufe über verschiedene Aspekte der Leistungsüberprüfung im Mathematikunterricht befragt. Das Ergebnis des Surveys zeigt, dass in der Regel summativ beurteilt wird: Sehr häufig geben die Lehrpersonen verbale Rückmeldungen, so in zwei Drittel der Klassen immer bei der Tafelarbeit. Weniger häufig werden Noten vergeben oder (Haus-)Aufgaben explizit beurteilt. Partizipative FLB (z. B. Selbst- und Peereinschätzungen durch die Schüler/innen) kommen nur sehr selten vor. Nur ca. 10 % der Lehrpersonen gaben vor Prüfungen und Schularbeiten Kriterien der Leistungsbeurteilung bekannt. Aus Sicht der Schüler/innen geben vorwiegend jene Lehrpersonen individuelle Rückmeldungen, die auch partizipative Leistungsbeurteilungsformen anwenden und die Schüler/innen durch anspruchsvollere offene

Rügener Inklusionsmodell

2

Survey zur Praxis der Leistungsbeurteilung

Mathematikaufgaben kognitiv aktivieren. Diese verfügen über eine hohe fachdidaktische und diagnostische Kompetenz. Eine Mehrebenenanalyse der Querschnittsdaten, die keine Kausalzuschreibung zulässt, ergab, dass die notenzentrierte summative Leistungsbeurteilung sowohl mit der Testmotivation als auch mit der Testleistung der Schüler/innen negativ verbunden ist, FLB und eine individuelle Bezugsnorm hingegen positiv. Die kriteriale Rückmeldung der Testleistung auf der Basis von Kompetenzstufenmodellen hatte ebenso signifikant bessere Effekte auf die Motivation und Attribution („Begabung“ statt „Glück“) als das sozialnormierte Feedback (Klieme et al., 2010).

In Laborexperimenten untersuchten Harks und andere (Harks, Rakoczy, Hattie, Besser & Klieme, 2014; Harks, Rakoczy, Klieme, Hattie & Besser, 2014) im Rahmen des Co²CA-Projekts mithilfe eines Pfadmodells die Wirkungen von individuellem schriftlichem kompetenz- und lösungsprozessbezogenem Feedback auf Motivation und Leistungen via wahrgenommener Nützlichkeit moderiert durch Interesse und Selbstwirksamkeit im Vergleich zum Notenfeedback. Das lösungsprozessbezogene Feedback verwendet eine individuelle und kriteriale Bezugsnorm, bezieht sich auf konkrete Aufgaben und Lösungsprozesse, informiert die Lernenden über ihr Stärken-Schwächen-Profil, bietet Strategien zum Erreichen längerfristiger Lernziele an und legt instabile Attributionen im Falle eines Misserfolgs nahe. Dieses Feedback bewirkte stärkere Leistungs- und Interessensteigerungen als das Notenfeedback (Harks, Rakoczy, Hattie, Besser & Klieme, 2014).² In einer weiteren Studie zur Wirkung des weniger aufwändig zu erstellenden kompetenzorientierten Feedbacks bauten sie im Co²CA-Projekt auf einer experimentellen Vorstudie und auf den Ergebnissen eines Experiments zur Wirkung kompetenzorientierten Feedbacks von Wollenschläger (2015) auf. Im Vergleich zu Kontrollgruppen ohne und mit einem sozialnormierten Notenfeedback führte das kompetenzorientierte Feedback zu positiven Effekten auf die Genauigkeit der Schülerelbsteinschätzungen ihrer fachlichen Leistungen, die wiederum die Leistung und Motivation erhöhte.

Wirkung des kompetenz- und lösungsprozessbezogenen Feedbacks

Harks, Rakoczy, Klieme, Hattie und Besser (2014) untersuchten daraufhin ebenfalls in einem experimentellen Design die Wirkungsweise der kompetenzbezogenen im Vergleich zur lösungsprozessbezogenen Rückmeldung und der Noten. Die Ergebnisse zeigen die positive, über die wahrgenommene Nützlichkeit der Rückmeldung vermittelte motivations- und leistungsfördernde Wirkung sowohl der kompetenz- als auch der lösungsprozessbezogenen Rückmeldung im Vergleich zu Noten. Die Moderatorenanalysen ergaben, dass die wahrgenommene Nützlichkeit der Rückmeldung sich bei Schülerinnen und Schülern, die eine niedrige Selbstwirksamkeitsüberzeugung und hohes Interesse aufweisen, stärker auswirkt als bei jenen, die eine hohe Selbstwirksamkeitsüberzeugung und niedriges Interesse zeigen (Harks, Rakoczy, Klieme, Hattie & Besser, 2014, S. 185 f.). Die Ergebnisse zu den differenziellen Effekten sind wesentlich für einen adaptiven Unterricht.

Die letzte Studie im Co²CA-Projekt ist ein Feldversuch, in dem die Erkenntnisse aus den Experimentalklassen auf den Normalunterricht übertragen wurden (Bürgermeister, Klieme, Rakoczy, Harks & Blum, 2014). Die Lehrpersonen wurden geschult, mit einer dafür entwickelten Diagnosehilfe ihren Schülerinnen und Schülern lösungsprozessbezogenes Feedback zu geben. Die Schüler/innen sollten im Verlauf von 13 Unterrichtsstunden dreimal jeweils ein bis zwei diagnostische Aufgaben zum Satz des Pythagoras lösen. Eine Kontrollgruppe von Lehrpersonen unterrichtete dasselbe ohne diese diagnostischen Aufgaben. Die ersten Ergebnisse dieses Feldversuchs zeigen, dass die Rückmeldungen die Lernleistungen besonders dann steigern, wenn

- auf verständliche Formulierung Wert gelegt wird,
- sie auf konkrete Aufgaben, aber nicht auf die Persönlichkeit bezogen sind,
- sie ein fachbezogenes Stärken-Schwächen-Profil enthalten,
- sie nicht sozial-, sondern kriterial- oder individualbezogen normiert sind.

² Das lösungsprozessbezogene Feedback wurde signifikant nützlicher wahrgenommen als das Notenfeedback ($\beta = .60$, 95 % Konfidenzintervall). Die wahrgenommene Nützlichkeit hatte einen signifikanten direkten positiven Effekt auf Leistung ($\beta = .20$, 95 % Konfidenzintervall) und Interesse ($\beta = .21$, 95 % Konfidenzintervall).

Im Vergleich zu den Lernenden aus der Kontrollgruppe erlebten die Schüler/innen in den Experimentalklassen diesen Unterricht eher als adaptiv und kompetenzunterstützend, was ihre Motivation und Leistungen steigerte. Das Konzept zur Lehrerschulung bildete die Basis für die Entwicklung eines Designs für Lehrerfortbildungen zur FLB in Mathematik (Besser, Tropper & Leiss, 2013).

Wirkungen der FLB im Feld

2.2 Formative Leistungsbewertung und alternative Beurteilungsformen

Spätestens seit der fundamentalen Kritik an den Ziffernnoten aus den 1970er Jahren (u. a. Ingenkamp, 1971) sind alternative Formen der summativen Leistungsbeurteilung ein Schlüsselthema für die Schulreform. Einige deutsche Bundesländer und Schweizer Kantone führten in ihren Grundschulen flächendeckend Verbalzeugnisse anstelle der Notenzugnisse ein. In Österreich vermehrten sich im Bereich der Grundschule stetig die Standorte mit einem solchen Schulversuch. Dass in Deutschland nur geringe Auswirkungen der detaillierteren Rückmeldungen in den Zeugnissen auf Einstellungen, Schulleistungen und Persönlichkeitsentwicklung der Kinder nachweisbar waren (Valtin & Wagner, 2002), ist ein Hinweis darauf, dass sie vor allem summativ genutzt wurden.

2.2.1 Alternative Beurteilungsformen als Instrumente der formativen Leistungsbewertung in einer neuen Lernkultur

Das Lernen ist produktiv und effizient, wenn es ein aktiver und konstruktiver Prozess ist (das Gegenteil von extern vermittelter, passiv aufgenommener und mechanisch verarbeiteter Information), der kontextuiert und situiert, intrinsisch motiviert, selbstorganisiert und selbstkontrolliert ist (Weinert, 1996). Dazu müssen die Schüler/innen Gelegenheit haben, das eigene Lernen zu reflektieren und einzuschätzen – in einem Unterricht, der Freiräume für selbstbestimmte Lernaktivitäten und innere Differenzierung gibt und eine Gleichgewichtung von Lernprozess und -produkt ermöglicht (Hofmann, 2000).

Die individuellen Lernwege können mit den alternativen Formen der Leistungsbeurteilung, z. B. in einem Lernjournal oder Portfolio, dokumentiert werden. Sie eignen sich damit nicht nur als Grundlage für die summativ Beurteilung, sondern auch für die FLB, da sie Einblicke in die jeweiligen Lernprozesse der Schüler/innen bieten, ihren jeweiligen Lernstand zeigen und für dessen Weiterentwicklung genutzt werden können (Stern, 2010).

Dokumentation des Lernens mit alternativen Formen der Leistungsbeurteilung

Um als effektives Instrument der FLB dienen zu können, sollten die FLB nachstehende Anforderungen erfüllen (Schmidinger, 2012, S. 37 f.), wodurch auch die Umsetzung der fünf Schlüsselstrategien (Leahy et al., 2005) ermöglicht wird:

- Sie sollen die *Lernkompetenzen* erfassen, also nicht nur die *Sachkompetenzen*, sondern auch die in der Lernkompetenz mit ihnen verbundenen *Methoden-, Sozial- und Selbstkompetenzen* (Czerwanski et al., 2002, S. 31). Da die Lernkompetenz die Fähigkeit ist, erfolgreich zu lernen und das Gelernte beim weiteren Lernen selbstorganisiert anzuwenden, unterstützen sie das mit der FLB angestrebte selbstregulierte Lernen (siehe Abschnitt 1.6).
- Gegenstand der Leistungsbewertung sollen nicht nur die *Lernergebnisse* als Output des Unterrichts sein, sondern auch die *Qualität der Lernprozesse*. Die Leistungsbewertungen sind damit sowohl produkt- als auch prozessorientiert (Winter, 2004, S. 72).
- Ebenfalls zur Förderung des in der FLB angestrebten selbstregulierten Lernens sollen *Selbstbewertungen* der Lernergebnisse und -prozesse im Rahmen der *Selbstreflexion* des eigenen Lernens erfolgen.
- Zur Förderung der intrinsischen Motivation sollen bei allen Bewertungen der Grad der Zielerreichung, aber auch die Lernfortschritte reflektiert werden. *Bezugsnormen* sollten daher die *Kriterial- und die Individualnorm* sein (Bohl, 2004, S. 63 ff.; Jürgens & Sacher, 2008, S. 101 f.). Voraussetzung ist, dass die Lernziele und die Erfolgskriterien den Lehrpersonen und Schülerinnen und Schülern klar sind.

Anforderungen an die alternativen Formen der Leistungsbeurteilung als Instrumente der FLB

- Eine angemessene Selbstbewertung kann durch vielfältige *Rückmeldungen* von Peers, den Lehrpersonen und außerschulischen Personen unterstützt werden, die auch eine soziale Verständigung über die Beurteilungskriterien ermöglichen. Durch diese „*soziale Einbettung*“ in einer entwickelten Feedbackkultur wird das Lernen „*dialogisch-reflexiv*“ (Winter, 2004, S. 73).
- Die Leistungsbewertung soll anhand der *Lernaufgaben als integraler Teil des aktiven Lernhandelns im Unterricht* erfolgen und nicht in einer abgegrenzten Prüfungssituation.
- Die Produkte, die beim Lernen entstehen – seien es bearbeitete Aufgaben, Präsentationen des eigenen Wissens und Könnens, Selbstreflexionen des Lernens in Lerntagebüchern usw. –, sind Belege des Lernens und dienen als „*direkte Leistungsvorlagen*“ (Vierlinger, 2012). Sie sind Gegenstand der Selbstbewertungen und Rückmeldungen.
- Die Bewertung soll inhaltlich-verbal mit *qualitativen Aussagen* erfolgen, wie „Ziel erreicht“ oder der Beschreibung des Lernstands (Bohl, 2004; Winter 2004, S. 73).

Nicht alle alternativen Formen der Leistungsbeurteilung erfüllen diese Anforderungen per se. Die Arbeit mit Portfolios, Lernjournals und die *lernzielorientierte Leistungsbeurteilung* (LOB) entsprechen am ehesten diesen Anforderungen. Doch ließen sich die meisten der praktizierten Formen so weiterentwickeln, dass sie diesen Anforderungen entsprechen.

2.2.2 Forschungs- und Entwicklungsprojekte

Inkompatibilität von
Noten und FLB

In den 90er Jahren erprobte man in der Schweiz die Umstellung auf eine lernförderliche Beurteilungskultur in den Entwicklungsprojekten „Ganzheitliches Beurteilen und Fördern“ an den Grundschulen im Kanton Luzern und „Beurteilen und Fördern“ an den Sekundarschulen im Kanton Zug. Die Evaluationsergebnisse zeigen, dass die Akzeptanz aller Beteiligten in den ersten Schulstufen der notenfreien Grundschule am höchsten war, später aber mit dem Näherrücken des Übertritts zu den differenzierten Sekundarschultypen abnahm. In den Sekundarschulen sollten weiterhin sozialnormierte Noten die Grundlage für Selektionsentscheidungen sein, zusätzlich sollten verstärkt lernförderliche Rückmeldungen gegeben werden (nach der Kriterial- und Individualnorm). Die meisten Lehrpersonen gaben an, durch dieses Dilemma belastet zu sein, am geringsten noch in den schulformengemischten (integrativen) Klassen (Roos, 2001; Roos, Huber, Sandmeier, Sempert & Schuler, 2007). Der Versuch einer nachhaltigen Einführung einer FLB in einem längerfristig angelegten Entwicklungsprojekt mit begleitenden Praxisberatungen, Fortbildungen und Interventionen scheiterte an der übermächtigen Notenorientierung sowohl der Lehrpersonen als auch der Schüler/innen und Eltern. Die Studienautorinnen und -autoren schließen daraus, dass lernförderlicher Unterricht nicht so sehr neue Lernarrangements erfordert, sondern vor allem eine Haltungsänderung seitens der Lehrpersonen.

2.3 Fazit des Forschungsstands

Die britische Metastudie „Inside the Black Box“ (Black & Wiliam, 1998), die neuseeländische Metastudie „Lernen sichtbar machen“ (Hattie, 2013) und die deutschen Forschungsergebnisse im Co²CA-Projekt (Harks, Rakoczy, Hattie, Besser & Klieme, 2014; Harks, Rakoczy, Klieme, Hattie & Besser, 2014; Klieme et al., 2010) wiesen die Wirksamkeit der FLB zur Steigerung der Leistung, der intrinsischen Motivation und der Anstrengungsbereitschaft nach. Dass im deutschsprachigen Raum immer noch die Notenrückmeldung vorherrscht, führen Roos et al. (2007) in den Ergebnissen des Projekts „Beurteilen und Fördern“ einerseits auf die Macht der Gewohnheit sowie auf die Illusion der Schüler/innen und Eltern von der Klarheit und Gerechtigkeit der Noten, andererseits auf das subjektive Unvereinbarkeitserleben der Lehrpersonen von der summativ-selektiven und formativ-lernfördernden Leistungsbeurteilungsfunktion zurück.

FLB, eine anspruchsvolle
Unterrichtsintervention

Die FLB ist eine anspruchsvolle Intervention im Unterricht, bei der die Lehrperson den momentanen Lernstand der einzelnen Schüler/innen in Unterrichtsroutinen laufend beobachtet

und festgestellte (Verständnis-)Lücken spontan mit Hinweisen für die Weiterarbeit rückmeldet (On-the-Fly Formative Assessment). Die Unterrichtsroutrinen beinhalten vielfältige Instrumente, wie kleine Lernkontrollen mit diagnostisch aufschlussreichen Aufgaben, Beurteilungsraster, Präsentationen, Qualitätsdiskussionen und den Lernprozess antizipierende, begleitende oder rückblickende schriftliche Reflexionen (vgl. Winter, 2015). Anhand der gemeinsam mit der Lehrperson erstellten Erfolgskriterien reflektieren die Schüler/innen selbst, aber auch mit Peers und der Lehrperson deren Bearbeitung, wobei sie, wenn nötig, Hinweise für Verbesserungen erhalten. Die kriteriale Bewertung kann durch eine Bewertung auf der Basis der Individualnorm ergänzt oder ersetzt werden, wenn vor allem die Lernfortschritte bewusst gemacht werden sollen. Die FLB kann auch mithilfe alternativer Leistungsbeurteilungsformen erfolgen, wie mit Portfolios oder Lernjournals im *Dialogischen Lernen* (Ruf & Gallin, 2011). Diese eignen sich für eine geplante FLB (*Planned-for-Interaction Formative Assessment*), wenn sie den Qualitätsanforderungen (siehe Abschnitt 2.2.1) entsprechen. Die Leistung der punktuellen formativen Tests im Rahmen der FLB ist eher in der Evaluation ihrer Effektivität zu sehen.

Die Erfahrungen aus den vorgestellten Entwicklungsprojekten zeigen, dass für die Einführung von FLB längerfristige Prozessbegleitung und Unterstützung erforderlich ist, insbesondere durch systematische Reflexion der eigenen Praxis. Die gleichzeitige Erfüllung der Selektions- und Förderfunktion mithilfe der Leistungsbeurteilung erscheint schwierig.

3 Analyse der Leistungsbeurteilungssituation in Österreich

Zur Praxis der FLB in Österreich gibt es zwar keine vergleichbar großen Studien, es ist jedoch zu vermuten, dass die im Abschnitt 2.1.2 vorgestellten Befragungsergebnisse zur Leistungsbeurteilung aus dem Co²CA-Projekt (Klieme et al., 2010) auch für Österreich zutreffend sind. Zusätzlich kann die Studie von Kobarg, Thoma, Dalehefte, Seidel und Prenzel (2011) zu den „lernwirksamen Bedingungen im alltäglichen Unterricht“ angeführt werden. Den beiden Studien zufolge kommen FLB oder einzelne ihrer Elemente, wie „Klarheit der Ziele“ und „Rückmeldungen“ nur selten vor. Sie spielten auch in der Unterrichtsplanung keine besondere Rolle.

Einen gewissen Einblick in konkrete österreichische Initiativen im Klassenzimmer und in die vorhandene Lehrerexpertise zum Thema der Leistungsbeurteilung geben die Projektberichte von *Innovationen Machen Schulen Top* (IMST; früher *Innovations in Mathematics, Science and Technology Teaching*) auf der Lernplattform IMST-Wiki.³ Diese von Lehrpersonen publizierten Beschreibungen und Reflexionen zu selbst erprobten Unterrichtsinnovationen sind zwar nicht repräsentativ für die didaktischen Reformbestrebungen an österreichischen Schulen, zeigen aber in ihrer Summe, wo von den Akteurinnen und Akteuren Reformbedarf im Klassenzimmer geortet wird. Im Rahmen von IMST führen derzeit jährlich etwa 100 Lehrerteams an ihren Schulen Innovationsprojekte durch und schreiben darüber Berichte, von denen über 1.000 im IMST-Wiki als Ideenpool veröffentlicht wurden. Sie wären generell eine Ressource für qualitative Forschungen, die von IMST soweit wie möglich auch genutzt wird.

So sind im IMST-Wiki nicht weniger als 439 Berichte, also etwa 40 %, unter dem Suchbegriff „Leistung“ kategorisiert. 67 dieser Berichte befassen sich im engeren Sinn explizit mit Innovationen im Bereich der Leistungsüberprüfung. Im Themenprogramm „Prüfungskultur“ begleitete und evaluierte das IMST 2010 bis 2012 32 Projekte (Schuster, 2013). Neben den Problemanalysen zur primär summativen Leistungsbeurteilung (Transparenz, Objektivität, Gerechtigkeit etc.) und Entwürfen von Alternativen zur Benotung finden sich auch Projekte, die nach praktikablen Formen der FLB suchen. Die Studien (siehe auch Altrichter & Nagy,

Wenige Daten zur FLB
in Österreich

³ IMST-Wiki, Lernplattform für Unterricht und Schule. Berichte über Unterrichtsinnovationen. Verfügbar unter <https://www.imst.ac.at/imst-wiki/index.php/Kategorie:Leistung> [zuletzt geprüft am 01.12.2015].

2010) zeigen, dass für die Lehrpersonen – trotz zahlreicher Seminare, Workshops und schulbezogener Fortbildungsveranstaltungen – nach wie vor summative Aspekte der Leistungsbewertung im Vordergrund stehen und Ansätze für eine systematische Verwendung formativer Leistungsdiagnosen erst im Ansatz zu erkennen sind; siehe dazu auch die Handreichungen des IMST mit Anregungen für eine innovative Leistungsbewertung (Amrhein-Kreml et al., 2008; Stern, 2008).

3.1 Strukturelle Hemmnisse für die formative Leistungsbewertung

So wie Roos in seinen beiden Evaluationsstudien (siehe Abschnitt 2.2.2) für die Schweiz feststellte, dominieren auch in Österreich in der Wahrnehmung sowohl der Lehrpersonen als auch der Schüler/innen sowie deren Eltern die Leistungsfeststellungen aufgrund ihrer Allokations- und Selektionsfunktion. Wenn auch die Allokationen, wie die Zuweisung in eine Vorschulstufe, das Wiederholen einer Schulstufe, die Feststellung eines sonderpädagogischen Förderbedarfs, die Übertrittsbestimmungen in weiterführende Schulen nach der Grundschule und der Hauptschule (HS) beziehungsweise nach der Neuen Mittelschule (NMS) als formative Maßnahme zur Lernförderung gedacht sind, werden sie von den Betroffenen meist als für ihren Lebensweg bedeutsame Selektionsmaßnahme empfunden.

Die *Noten* werden dadurch zu einem starken *Motivationsfaktor*, der die anfänglich intrinsische Motivation für das schulische Lernen, ein Lernen aus Freude und Interesse, zur weniger lerndienlichen extrinsischen Motivation werden lässt, zu einem Lernen um der guten Noten willen. Die Leistungsfeststellungen erhalten auch durch die größere Regelungsdichte mehr Aufmerksamkeit der Lehrpersonen sowie der Schüler/innen als die Informationsfeststellungen, für die es nur die Regelung gibt, dass sie nicht benotet werden dürfen (Eder et al., 2009; Schmidinger, 2012, 2013). Da die einzelnen erfassten Leistungen bei der (Leistungs-) Feststellung der Mitarbeit auch nicht gesondert benotet werden dürfen, sind die Informationsfeststellungen für die Schüler/innen nur schwer als solche erkennbar.

Fehlende Trennung
von Lern- und
Beurteilungsphasen

Der Eindruck des *Fehlens beurteilungsfreier Phasen* verstärkt sich für die Schüler/innen in den häufigen Unterrichtsgesprächen, in denen die Lehrpersonen die Schülerbeiträge in der Regel verbal oder nonverbal, oft auch im Hinblick auf die Mitarbeitsnote bewerten. Fehler werden daher von den Schülerinnen und Schülern im gesamten Unterricht möglichst vermieden und können so nicht lerndienlich aufgeklärt werden (Amrhein-Kreml et al., 2008).

Die Selbstbeurteilung und -reflexion des eigenen Lernens im Rahmen der FLB werden durch die regelmäßigen Leistungsfeststellungen, die ausschließlich Lehreraufgabe sind, konterkariert. Das Schulunterrichtsgesetz (SchUG) und die LBVO verlangen zwar eine Beurteilung auf der Grundlage der kriterialen Bezugsnorm, die auch für die lernfördernden Effekte eine Voraussetzung ist. Bei der stärkeren Gewichtung des zuletzt erfassten Leistungsstands bei der Bestimmung der (Halb-)Jahresnote (§ 20 LBVO) sollte auch die individuelle Bezugsnorm Berücksichtigung finden. Mangels konkreter kompetenzbezogener Kriterien wird jedoch in der Regel nach der sozialen Bezugsnorm beurteilt, die negative Effekte auf die Leistungsmotivation und das Selbstkonzept leistungsschwächerer Schüler/innen nach sich zieht. Aufgrund dieser Bedingungen sind die Noten für die Schüler/innen die wichtigsten Rückmeldungen über den Erfolg ihres Lernens, denen sie jedoch aufgrund ihrer Abstraktheit (Eder et al., 2009) nicht die Informationen entnehmen können, die sie für ihr weiteres Lernen verwerten könnten.

Zu abstrakte Rückmeldung
durch Noten

Das Fehlen ausreichender FLB kann auch an den mangelnden Diagnosekompetenzen der Lehrpersonen liegen. Im Band 2 des Nationalen Bildungsberichts 2012 (Herzog-Punzenberger, 2012) wird z. B. die große Unsicherheit bei der Identifizierung der Kinder mit größeren Lesedefiziten als Problembereich bei der Leseförderung nachgewiesen (Schabmann, Landerl, Bruneforth & Schmidt, 2012).

3.2 Günstige Voraussetzungen für die formative Leistungsbewertung

Der Analyse der Hemmnisse für die FLB in Abschnitt 3.1 sind einige positive Entwicklungsansätze zur Förderung der FLB während der letzten Jahre gegenüberzustellen.

3.2.1 Legalisierungsbemühungen um die alternativen Formen der Leistungsbeurteilung in der Grundschule

Zu den alternativen Formen der Leistungsbeurteilung werden gemäß dem Bericht des Rechnungshofs (RH) schon seit 1966/67 auf Initiative der Schulen und der Landesschulräte sowie des Stadtschulrats von Wien Schulversuche an Österreichs Volksschulen geführt. Mittlerweile ist dies der Schulversuch mit der größten Anzahl an Klassen (37 % aller durchgeführten Schulversuche im Schuljahr 2012/13 (Rechnungshof, 2015).

Dies wurde durch den am 01.09.1998 in Kraft getretenen § 78a SchUG ermöglicht, der die Anzahl der möglichen Schulversuchsklassen auf 25 % der Anzahl aller Klassen an öffentlichen Schulen im Bundesgebiet erhöhte, da kein politischer Konsens zur Änderung der gesetzlichen Bestimmungen erreicht wurde. Dies kam einer Legalisierung nahe, da dadurch in allen Klassen der ersten drei Schulstufen ein solcher Schulversuch möglich wäre. Es gibt zwar keine empirisch belegbare Qualitätsaussage zu den Wirkungen der Schulversuche, da sie nie bundesweit evaluiert wurden. Trotzdem sind nun die Regierungspartner der Empfehlung des RH gefolgt und haben die Übernahme der „alternativen Leistungsbeschreibung“ in den schulautonomen Bereich bis einschließlich der 3. Schulstufe in ihr Arbeitsprogramm für 2013 bis 2018 (Republik Österreich, 2013, S. 41) aufgenommen. Damit würden die Ressourcen, die durch die Administration der Schulversuche entstehen, eingespart – Mittel, die zur Unterstützung einer wirksamen Umsetzung vor allem der formativen Funktion der alternativen Leistungsbeurteilung gut gebraucht werden könnten.

Schulversuch mit den meisten Klassen

3.2.2 Die Bildungsstandards als kriteriale Bezugsnorm

Die gesetzliche Verankerung (2008) von *Bildungsstandards* (BIST) für die 4. und 8. Schulstufe im § 17 Abs. 1a des SchUG und den beiden Verordnungen zu den „Bildungsstandards im Schulwesen“ (BGBl. II Nr. 1/2009 und BGBl. II Nr. 282/2011) haben eine stärkere Ergebnisorientierung im Unterricht, einen nachhaltigen Kompetenzaufbau und eine gezielte individuelle Förderung der Schüler/innen zum Ziel. Die BIST sind als Regelstandards konzipiert, die die Kompetenzen beschreiben, die von den Schülerinnen und den Schülern der 4. und 8. Schulstufe im Durchschnitt erreicht werden sollen. Die Verordnung BGBl. II Nr. 1/2009 enthält auch den Auftrag zu periodischen Standardüberprüfungen, deren Ergebnisse sowohl auf Grundlage der kriterialen als auch der sozialen Bezugsnorm ausgewertet und rückgemeldet werden. Die BIST-Maßnahmen bieten damit u. a. entsprechende Rahmenbedingungen für einen stärker kompetenzorientierten Unterricht und Beispiele, wie Leistungen auf eine Kriterialnorm bezogen beurteilt werden können. Dies könnte eine Umorientierung von der in Österreich vorherrschenden sozialen zur kriterialen Bezugsnorm einleiten, die neben der Individualnorm auch eine wesentliche Bedingung für eine effektive FLB ist.

BIST als kriteriale Bezugsnorm anstelle der sozialen

Auch die berufsbildenden Schulen entwickeln seit 2004 eigene BIST, die auf die Abschlussqualifikationen der 11. Schulstufe der berufsbildenden mittleren Schulen (BMS) und der 13. Schulstufe der BHS abzielen. Sie sind auch als Regelstandards konzipiert, werden in den neuen Lehrplänen in den Bildungs- und Lehraufgaben enthalten sein und bilden so, anders als in der AHS, die Grundlage der Leistungsbeurteilung in den berufsbildenden mittleren und höheren Schulen (BMHS). Sie beinhalten ein Kompetenzmodell, Deskriptoren und Unterrichtsbeispiele, einige auch Kompetenzraster (in Entwicklung). Sie sollen damit – wie in den allgemeinbildenden Schulen – den kompetenzorientierten Unterricht und eine transparente, auf die kriteriale Norm bezogene Leistungsbeurteilung und FLB fördern.

Die Begleitforschungen über die „Rezeption und Nutzung von Ergebnissen der Bildungsstandardüberprüfungen in Mathematik auf der 8. Schulstufe unter Berücksichtigung der Rückmeldemoderation“ (Rieß & Zuber, 2014) und der 4. Schulstufe (Dinges & Egger, 2015), die vom Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens (BIFIE) durchgeführt wurden, zeigen, dass zumindest alle Schulleitungen angeben, sich mit den Rückmeldungen der Standardtestungen auseinanderzusetzen. In der Folge leiten sie oft auch Schulentwicklungsmaßnahmen ein, die jedoch nicht immer in einem sachlichen Zusammenhang mit den rückgemeldeten Ergebnissen stehen. Die Schülerergebnisse werden für die Schulleitungen im Schulbericht auf der Basis der sozialen und der kriterialen Bezugsnorm dargestellt, wobei die sozial normierten Daten von den Schulleitungen sehr viel häufiger als die kriterial normierten rezipiert werden. Dies ist ein Hinweis einerseits, wie bedeutsam die Sozialnorm für die Schulleitungen ist, und andererseits, wie wenig Bedeutung und Wert die Kriterialnorm heute noch für die Lehrerschaft hat, und dass auch das Vorliegen verbindlicher kriterialer Normen nicht selbstverständlich deren Anwendung zur Folge hat.

Unterstützungsmaßnahmen seitens des BIFIE

Das BIFIE unterstützt die Umsetzung der BIST im Unterricht auch durch die Publikation von Praxishandbüchern, deren Inhalte an Beispielen in zusätzlichen Themenheften fachdidaktisch vertieft werden. So findet sich im „Themenheft für die Kompetenzbereiche ‚Einsicht in Sprache durch Sprachbetrachtung‘ und ‚Rechtschreiben‘“ (BIFIE, 2014) eine Kompetenzentwicklungsplanung („Kompetenzerwerbsschemata“) für den Kompetenzaufbau über die vier Schulstufen der Grundschule, die auch eine FLB erleichtern kann (siehe Abschnitt 1.4). Darüber hinaus werden Aufgabenbeispiele, die die verschiedenen Kompetenzstufen veranschaulichen, zur Verwendung im Unterricht veröffentlicht, die ebenfalls für die FLB genützt werden könnten.

Zusätzlich werden den Lehrpersonen Diagnoseinstrumente zur *Informellen Kompetenzmessung* (IKM) für Deutsch (Lesen/Einsicht in die Sprache) und Mathematik der Grundschule sowie für Deutsch (Lesen/Sprachbewusstsein), Mathematik und Englisch (Listening/Reading) der Sekundarstufe I angeboten. Sie können schon im Schuljahr vor der Standardüberprüfung⁴ – in der Sekundarstufe I zusätzlich auch in der 6. Schulstufe – zur Ermittlung des Förderbedarfs in Bezug auf die Standardtestung eingesetzt werden, um im verbleibenden Unterricht noch entsprechende Schwerpunktsetzungen oder individuelle Förderungen vornehmen zu können.

Bereits im Sommersemester 2013 wurden die IKM von den Grundschulen gut genützt. Die Bestellraten variierten in den Bundesländern zwischen 43 % und 82 % der Schulen (BIFIE, 2014, S. 145). Die jüngsten Daten vom Frühjahr 2015 (BIFIE, 2015) zeigen ein ähnliches Bild: 61 % der Volksschulen und knapp 60 % aller Klassen der 3. Schulstufe nahmen an der IKM teil, wobei für fast alle Klassen alle drei Aufgabenpakete bestellt wurden (Deutsch/Lesen: 53.556 Hefte, Deutsch/Sprachbetrachtung: 53.510 Hefte, Mathematik: 55.442 Hefte). Damit nahmen maximal 71 % der Schüler/innen der 3. Schulstufe an der IKM teil, es könnten aber auch weniger gewesen sein, da man nicht weiß, ob alle bestellten Hefte tatsächlich eingesetzt wurden.

Die Schulen der Sekundarstufe I nahmen das Online-Angebot⁵ der IKM nicht in diesem Ausmaß an. Nur 47 % dieser Schulen nutzten es, vor allem die NMS (57 %), gefolgt von der HS (27 %), am wenigsten häufig die AHS (16 %). Jeweils rund ein Drittel dieser Schulen haben die IKM in einem Fach, in zwei Fächern oder in allen drei angebotenen Fächern durchgeführt. Deutsch/Lesen für die 7. Schulstufe wurde von allen Angeboten am häufigsten eingesetzt. Es wurden damit aber nur 18 % aller Schüler/innen dieser Schulstufe erreicht. Dagegen werden die Orientierungs-Checks auf der 9. Schulstufe im BMHS-Bereich schon fast flächendeckend verwendet (siehe Abschnitt 3.2.5).

⁴ Das ist in der Grundschule die 3. Schulstufe, sie kann aber auch noch zu Beginn der 4. Schulstufe eingesetzt werden und in der Sekundarstufe I die 7. Schulstufe.

⁵ Für den Onlinetest mussten die Lehrpersonen auf der Plattform Sessions anlegen. Im Folgenden werden Kennzahlen zu den erfolgreich abgeschlossenen Sessions angegeben.

Für die 9. Schulstufe werden im Bereich BHS „Orientierungs-Checks“ für die Unterrichtsgegenstände Deutsch, Englisch (Reading/Listening) und Angewandte Mathematik für Lernstandserhebungen in der Schuleingangsphase angeboten, deren Ergebnisse in Form individueller Stärken-Schwächen-Analysen (ohne Noten) dargestellt werden. Sie haben somit ausschließlich diagnostischen Charakter und werden den Lehrpersonen zum Einsatz während der ersten Schulwochen empfohlen. Sie sollen ein integrativer Bestandteil der Feedbackkultur werden (Dorninger, Niemeier, Pachatz, Pregesbauer & Winkler-Rigler, 2013, S. 35). Die Lehrpersonen können sie als Grundlage für individualisierte Unterstützungsmaßnahmen und für an die Eingangsvoraussetzungen besser angepasste Unterrichtsplanungen verwenden. Selbstevaluierungstools (Lernausgangsmessungen, Orientierungsaufgaben) sind seitens des BIFIE auch für Englisch und Mathematik der 5. Schulstufe geplant, weitere für Mathematik auf der 4. sowie für Mathematik und Englisch (Reading/Listening) auf der 8. Schulstufe (persönl. Mitteilung im Rahmen des BMBF-Hintergrundgesprächs, 29.04.2015).

Mit den Selbstevaluierungsinstrumenten IKM und Orientierungsaufgaben kann die Effektivität des eigenen Unterrichts eingeschätzt werden. Dies hat zwar im Rahmen einer FLB eine bedeutsame Funktion, aufgrund der limitierten Einsatzmöglichkeiten der bis jetzt angebotenen Instrumente können sie diese Funktion jedoch nur punktuell leisten.

IKM und Orientierungs-Checks zur Evaluierung der Effektivität des eigenen Unterrichts

3.2.3 Weiterentwicklung der Leistungsbeurteilung durch das BMBF auf der Grundlage der bestehenden gesetzlichen Bestimmungen

Im Abschnitt 3.1 wurde gezeigt, dass in der Praxis die Leistungsfeststellungen und -beurteilungen im Vergleich zu den FLB bei Weitem dominieren. Dazu kommt, dass auch das Prozedere der Notenfindung (Punkteverrechnung beziehungsweise Prozentsätze der gelösten Aufgaben u. a. m.) in der Praxis weder den Regelungen der geltenden LBVO entspricht noch sachgemäß ist, was am Abstraktheitsgrad der LBVO-Regelungen liegen mag (Neuweg, 2014). Die durch die Einführung der BIST auf der 4. und der 8. Schulstufe und durch die begonnene Aufnahme von Bildungsstandards in die Bildungs- und Lehraufgaben der BMHS-Lehrpläne (siehe Abschnitt 3.2.2) eingeleitete stärkere Kompetenzorientierung des Unterrichts erfordert auch Anpassungen in der LBVO.

So begann 2011 eine Arbeitsgruppe im BMUKK (jetzt BMBF), wissenschaftlich begleitet, mit Arbeiten an einer Weiterentwicklung der LBVO – eine Entwicklungsoption, die schon im Beitrag von Eder et al. (2009) im Nationalen Bildungsbericht 2009 angedacht war. Da die LBVO auch weiterhin für alle Schulformen gelten soll, beteiligten sich seit 2012 beide pädagogischen Sektionen (allgemeinbildendes und berufsbildendes Schulwesen) sowie die Rechtssektion mit zunehmender Intensität an den Tätigkeiten der Arbeitsgruppe. Da eine Erneuerung der LBVO im aktuellen Regierungsprogramm nicht enthalten ist, können in dieser Regierungsperiode jedoch nur die notwendigen ministeriumsinternen Vorbereitungsarbeiten für eine Aufnahme dieses Projekts im nächsten Regierungsprogramm erfolgen. Um eine gewisse Außenwirksamkeit zu erzielen, wurden zwei Tagungen durchgeführt:

2013:

„Pädagogische Expertinnen- und Expertentagung zur kompetenzorientierten Leistungsbeurteilung“ am 02.05.2013 an der Pädagogischen Hochschule Oberösterreich (PH OÖ).

2014:

„Enquete zur Leistungsbeurteilung“ am 25.04.2014 im BMBF: Darstellung und Kommentierung des bisherigen Diskussionsstandes auf der Grundlage eines ministeriumsinternen Entwurfs zur LBVO (2014) durch Prof. Eder (Universität Salzburg), Dr. Westfall-Greiter (Universität Innsbruck), Dr. Schmidinger (Landesschulrat für Oberösterreich) und Prof. Moser (Pädagogische Hochschule Salzburg).

Arbeiten an der Weiterentwicklung der Leistungsbeurteilung

Zum Zeitpunkt der Erstellung dieses Beitrags ist für Ende 2015 die Einrichtung von sechs Arbeitsgruppen mit Praktikerinnen und Praktikern geplant, die in den nächsten zwei Jahren für die Sekundarstufen I und II Kompetenzraster für die Unterrichtsgegenstände Deutsch, Englisch, Mathematik, Geschichte und Sozialkunde/Politische Bildung, Bewegung und Sport sowie Informations- und Kommunikationstechnologie entwickeln und in ihrem Unterricht erproben sollen. Die Arbeit dieser Gruppen soll von der PH OÖ sowie von *Sounding Boards* unterstützt und auch wissenschaftlich begleitet werden. Diese Vorgehensweise ist darauf ausgerichtet, der Ressortleitung bis zum Ende der Legislaturperiode 2018 eine entsprechende Empfehlung für eine neue LBVO und für die Aufnahme dieses Vorhabens in das nächste Regierungsprogramm geben zu können.

Ziele dieser Arbeiten sind eine konsequente Kompetenzorientierung der Leistungsfeststellungen und -beurteilungen, die Anpassung der Leistungsfeststellungsinstrumente an die zu messenden Kompetenzen und eine Stärkung ihrer formativen Wirksamkeit. Dies erfordert die Auflistung der im jeweiligen Beurteilungszeitraum (Semester, Schuljahr) zu erwerbenden Kompetenzen in Kompetenzrastern, wobei zwischen Grundanforderungen und darüber hinausgehenden Anforderungen unterschieden werden soll. Die erworbenen und demonstrierten Kompetenzen der Schüler/innen sollen in eigenen kompetenzspezifischen Beurteilungsrastern erfasst werden, in denen ebenfalls zwischen Grundanforderungen und darüber hinausgehenden Anforderungen unterschieden wird, für die jeweils Indikatoren für zwei Stufen (überwiegend und vollständig erfüllt) enthalten sein sollen.

Nach dem ministeriumsinternen Entwurf der LBVO ist die Leistungsbeurteilung eine Gesamtbeurteilung des am Ende des Beurteilungszeitraums vorliegenden Kompetenzprofils der Lernenden (§ 13 Abs. 1 LBVO). Sie erfolgt weiterhin mit der fünfstufigen Notenskala, jedoch nach definierten Konversionsregeln, die den Kriterien der Beurteilung im SchUG (§ 18 Abs. 3) und den Definitionen der einzelnen Notenstufen in der geltenden LBVO (§ 14) entsprechen.

Diese Überlegungen des BMBF zur Weiterentwicklung der Leistungsbeurteilung sind für die formative Wirksamkeit der Leistungsfeststellungen insofern von Relevanz, als

- die Beurteilungsraster den Lernenden die erreichte Kompetenzstufe zeigen und gleichzeitig im Indikator der nächsthöheren Stufe das Ziel eines möglichen nächsten Lernschritts;
- das Ergebnis einer Leistungsfeststellung von den Schülerinnen und Schülern individuell durch ein besseres Ergebnis einer späteren Leistungsfeststellung im selben Beurteilungszeitraum ergänzt beziehungsweise ersetzt werden können soll.

Formative Wirkung der Leistungsfeststellungen

Ein solches Vorgehen würde längere notenfreie Perioden und Transparenz schaffen, das selbstständige und selbstverantwortete Lernen fördern und auch eine sowohl auf die Sachnorm als auch auf die Individualnorm bezogene Selbst- und Fremdbeurteilung unterstützen. Die Leistungsfeststellungen, die nicht mehr im aktuellen Ausmaß normiert werden sollen und auch von den Lernenden selbst initiiert (z. B. Portfolieinlagen) und entwickelt werden können, hätten damit auch lernfördernde Eigenschaften. Sie würden sich damit für ein auch die FLB unterstützendes, laufendes Monitoring des Leistungsniveaus durch die Lehrpersonen und durch die Schüler/innen selbst eignen.

Stärkere Kompetenzorientierung

Da jedoch einerseits nur so viele mündliche und schriftliche Leistungsfeststellungen vorzusehen sind, wie für eine sichere Leistungsbeurteilung unbedingt notwendig (LBVO § 3 Abs. 1), und andererseits für die formative Wirksamkeit vom Beginn eines Lernprozesses an hochfrequent (short-cycled, siehe Abschnitt 1.3.) formative Maßnahmen gesetzt werden müssen, stellt sich die Frage, ob die vorgesehenen Leistungsfeststellungen für eine optimale formative Wirksamkeit ausreichend sein werden.

Wenn die Leistungsfeststellungen die alleinige Grundlage der Noten in den Zeugnissen bleiben, an die die verschiedenen selektionswirksamen Berechtigungen geknüpft sind, kann erst die Praxis zeigen, in welchem Ausmaß die jetzt stark ausgeprägte, auf die Noten ausgerichtete extrinsische Motivation der Schüler/innen zu einer lerndienlicheren intrinsischen wird.

3.2.4 Neue Formen der Leistungsbeurteilung in der Neuen Mittelschule

Als die NMS als Modellversuch 2008 begonnen wurde, wollte man in der Versuchszeit eine gemeinsame Schulform für alle 10- bis 14-jährigen Schüler/innen entwickeln, die durch ihr didaktisches Konzept in der Lage sein sollte, „die Bildungserträge zu steigern, Bildungsbarrieren abzubauen, den Einfluss der sozialen Herkunft zu kompensieren und damit mehr Chancengleichheit und Bildungsgerechtigkeit zu schaffen“ (Eder, Altrichter, Bacher, Hofmann & Weber, 2015, S. 444).

Dieses didaktische Konzept baut auf der Förderung eigenverantwortlichen und selbstgesteuerten Lernens auf. Die Leistungsbeurteilung soll diesem didaktischen Konzept entsprechen und es vor allem durch ihre formative Funktion unterstützen. Konsequenterweise bemühte sich das *Zentrum für lernende Schulen – NMS-Entwicklungsbegleitung* (ZLS) seit seiner Gründung im Schuljahr 2012/13 mit seinen Fortbildungsmaßnahmen die Leistungsbeurteilung an den NMS auf die vorgesehene kriteriale Bezugsnorm auszurichten und ein Bewusstsein für die lernfördernde Funktion der FLB zu schaffen (z. B. Westfall-Greiter, 2012). Für diesen Zweck sollen die erworbenen Kompetenzen während des Schuljahrs mit alternativen Formen wie Portfolios und gestuften Kompetenzrastern dokumentiert und bewertet werden, die den Lernenden zeigen, wo sie sich auf dem Weg zum jeweiligen Ziel befinden (NMS-Lehrplan, S. 12). Damit sollen auch Selbstkompetenz und Selbsteinschätzung der Schüler/innen gefördert werden. Das Jahreszeugnis wird mit einer *verbalen differenzierenden Leistungsbeschreibung* (EDL) ergänzt, in der die besonderen Stärken ausgewiesen werden sollen, vor allem aber auch, um überfachliche Kompetenzen sinnvoll bewerten zu können. Gemäß § 19 Abs. 1a SchUG müssen überdies die individuellen Leistungsstärken und der jeweilige Lernstand regelmäßig in einem *Kind-Eltern-Lehrperson-Gespräch* (KEL) rückgemeldet werden, wobei die Lernenden ihre Leistungen selbst darstellen sollen. Westfall-Greiter (2012) empfiehlt ein solches Gespräch in ihrer „Orientierungshilfe Leistungsbeurteilung“ einmal pro Schüler/in in jedem Semester.

Didaktisches Konzept
enthält FLB

Die Evaluation der 2012 zur Regelform gewordenen und damit die HS ersetzende NMS (Eder et al., 2015) zeigt, dass das anspruchsvolle didaktische Konzept an den verschiedenen Standorten mit unterschiedlicher Qualität umgesetzt wird. Insgesamt berichten zwar die Schüler/innen eine Zunahme von förderorientierten Unterrichtsmaßnahmen und einen Rückgang überfordernder Verhaltensweisen der Lehrpersonen sowie mehr individuelle Wahlmöglichkeiten. Dies führte jedoch noch nicht zu den erwarteten Leistungssteigerungen und auch zu keinen Verbesserungen persönlichkeitsbezogener Merkmale wie dem Selbstkonzept. Vergleicht man hingegen die Klassen, die das didaktische Konzept am besten umsetzen, mit den restlichen, so werden in den „modellschulartigen“ Klassen mehr Qualitätsindikatoren im Unterricht umgesetzt und auch bessere Leistungen gemessen.

Das pädagogische Potenzial der EDL und der KEL wird von den Lehrpersonen noch nicht ausreichend wahrgenommen und in den Fallstudien als zusätzliche Belastung gesehen (Altrichter, Nagy & Pocrnja, 2015, S. 436). Die Befragung der Lehrpersonen der repräsentativen Evaluationsstichprobe ergibt zwar, dass rund 80 % das KEL durchführten, deren formative Nutzung aber noch zu untersuchen wäre. Rund 70 % der Lehrpersonen ließen ein Portfolio führen, Lernjournale jedoch nur rund 7 % (Eder et al., 2015, S. 52 f.). Irritierend an diesen Ergebnissen ist, dass das KEL verpflichtend, also von 100 % durchzuführen ist, und dass unklar bleibt, was hier unter Portfolio verstanden wurde.

Evaluationsergebnisse zur
Leistungsbeurteilung und
FLB

Die Zusammenschau der Evaluationsergebnisse zeigt, dass das anspruchsvolle didaktische Konzept der NMS in der Fläche innerhalb der vier Schuljahre des Evaluationszeitraums selbst bei einem relativ hohen Ressourceneinsatz (noch) nicht effektiv umgesetzt werden konnte. Wie weit die zu vermutende mangelhafte Umsetzung der FLB als wesentliches Konzeptelement eine Ursache für die noch nicht zufriedenstellenden Evaluationsergebnisse ist, kann diesen nicht entnommen werden.

3.2.5 Schulqualität Allgemeinbildung und Qualitätsinitiative Berufsbildung

Schulqualität Allgemeinbildung (SQA) und *Qualitätsinitiative Berufsbildung* (QIBB) sind pädagogische Qualitätsinitiativen des BMBF auf der Grundlage des § 18 Bundes-Schulaufsichtsgesetz und § 56 SchUG, die die Schulen bei ihrem Qualitätsmanagement unterstützen sollen.

Für das allgemeinbildende Schulwesen lautete die Rahmenvorgabe für die Perioden 2012/13 bis 2015/16 „Weiterentwicklung des Lernens und Lehrens [...] in Richtung Individualisierung und Kompetenzorientierung“ (BMUKK, 2013, S. 2), die im jüngsten diesbezüglichen Rundschreiben für das Schuljahr 2015/16 mit „in inklusiven Settings“ ergänzt wurde (BMBF, 2015, S. 2). Im Rahmen der „Erläuterung zur Bearbeitung der Ressortschwerpunkte“ findet sich im Rundschreiben Nr. 14/2013 (BMUKK, 2013, S. 4) für die Schulen der Sekundarstufe I zu einer Leitfrage für die Entwicklungsplanung auch ein Hinweis auf Maßnahmen der FLB: „Wie können [...] pädagogische Diagnoseinstrumente, förderliche Leistungsbeurteilung, innere Differenzierung und Begabungsförderung im Unterricht eingesetzt werden?“ (BMUKK, 2013, S. 4). Zum oben angeführten Rahmenziel müssen alle allgemeinbildenden Schulen jährlich ihrer Schulaufsicht einen für das Schuljahr adaptierten Entwicklungsplan vorlegen und mit ihnen entsprechende Vereinbarungen im *Bilanz- und Zielvereinbarungsgespräch* treffen. Diese sind die Basis für die Entwicklungspläne, die von der Schulaufsicht für ihre Region und von der zuständigen Sektion des BMBF für das Bundesgebiet jährlich zu erstellen sind.

Qualitätskriterien für Teilaspekte der FLB in SQA

Zur Unterstützung der Entwicklungs- und Qualitätsmanagementarbeit wurde die Plattform Schulqualität Allgemeinbildung (www.sqa.at) eingerichtet, die auf die 1999 begonnene Initiative *Qualität in Schulen* (Q. I. S.) aufbaute. Die Schulen finden dort u. a. Kriterien zu sechs Bereichen der Unterrichts- und Schulqualität und einen Schülerfragebogen für ein freiwilliges Individualfeedback an die Lehrperson. Sowohl der zweite Qualitätsbereich „Lernen und Lehren“ als auch der Bereich „Unterstützung beim Lernen“ im Individualfeedbackfragebogen enthalten Items, die für die FLB relevante Qualitätskriterien erfassen. Da die Entwicklungspläne aller Ebenen im Rahmen von SQA nicht öffentlich zugänglich sind, kann leider nicht festgestellt werden, in welcher Weise und Intensität die FLB Inhalt der Schulentwicklung geworden ist.

QIBB hat im Gegensatz zu SQA ihre 2004 ebenfalls auf der Basis von Q. I. S. begonnene Initiative nahtlos weitergeführt, sodass es seit Beginn dieser Initiative eine kontinuierliche Arbeit an bundesweiten Qualitätsschwerpunkten gibt, die in den alle drei Jahre veröffentlichten Landes- und Bundesqualitätsberichten dokumentiert sind. Es werden auf ihrer Plattform (www.qibb.at) ebenfalls Systemfeedback-Fragebögen für alle Ebenen (Schüler/innen, Lehrpersonen, Führungskräfte) zur Selbstevaluation der Bereiche „Leistungsbeurteilung“ und „Förderung & Individualisierung“ sowie ein Screeningfragebogen für die Schüler/innen angeboten, der u. a. auch Fragen zu diesen Bereichen enthält.

Der erste in allen Schularten bearbeitete bundesweit vorgegebene Qualitätsschwerpunkt war im Schuljahr 2005/06 „Transparente Leistungsbeurteilung“, dessen Ziele im Schuljahr 2010/11 im nachfolgenden Schwerpunktthema „Individualisierung“ weitergeführt wurden. Die jüngsten Bundesqualitätsberichte liegen für das technische, gewerbliche und kunstgewerbliche Schulwesen (Dorninger et al., 2013), für die kaufmännischen Schulen (Kiss, Gintzenstorfer & Andre, 2015; Sperl, Kiss & Müllauer-Hager, 2013) sowie für die Bildungsan-

stalten für Kindergartenpädagogik (BAKIP) und Sozialpädagogik (BASOP; Zug, 2013) für den Berichtszeitraum 2010 bis 2012 und den Planungszeitraum 2012 bis 2014 vor.

Zum Schwerpunkt „Transparente Leistungsbeurteilung“ wurden in allen Schulformen Entwicklungsmaßnahmen gesetzt, die dazu dienten, die Leistungsbeurteilungskonzepte am Standort zu konkretisieren, innerhalb der und zwischen den Fachgruppen abzustimmen sowie den Schulpartnern bekannt zu geben. Auch zur FLB finden sich Maßnahmen, wie die Intensivierung des individuellen, fördernden Feedbacks durch Kompetenzportfolios, Coaching- und Tutoringsysteme, Förderung der Selbstreflexion und Selbsteinschätzung, Verringerung der Fehlerkultur zugunsten einer lernförderlichen Rückmeldekultur, verstärkter Einsatz des COOL-Konzepts (*Cooperatives Offenes Lernen*) und Entwicklung einer individuellen Stärken-Schwächen-Beratung für die Schüler/innen.

Die erreichte Qualität dieser Maßnahmen zeigt sich im Screeningfragebogen in den drei Items zur „Leistungsbeurteilung“ und den fünf Items zur „Förderung & Individualisierung“, den die Schüler/innen auf der Plattform von QIBB regelmäßig und alle drei Jahre verbindlich beantworten sollen.⁶ Am Ende des Berichtszeitraums 2010 bis 2012 werden die Items, die Leistungsbeurteilungsmaßnahmen beschreiben, von den Schülerinnen und Schülern höherer technischer Lehranstalten (HTL) im Bundesschnitt zwischen 2,5 und 2,9 und damit gut bewertet, die für die FLB relevanten, auf die Individualisierung bezogenen Maßnahmen jedoch nur zwischen 3,1 und 3,9. Diese Werte sollen den Schulen Verbesserungsbedarf anzeigen. Die schlechteste, im Rahmen von QIBB gerade noch nicht negativ interpretierte Bewertung von 3,9 erhielt das Item „Meine Lehrer/innen beraten mich über meine Stärken und Schwächen im betreffenden Unterrichtsgegenstand“ (Dorninger et al., 2013, S. 26). Der Bundesqualitätsbericht 2014 (Babel & Spitzbart, S. 34 f.) zeigt, dass die Bundesergebnisse der verschiedenen Schularten in diesem Berichtszeitraum einander sehr ähnlich sind. So bewerten die Schüler/innen aller berufsbildenden Schularten einzelne Items des Qualitätsbereichs „Individualisierung“ zwischen 3,0 und 4,0 (in den BASOP: zwei Items, in den HTL: alle fünf Items). In den kaufmännischen Schulen, deren Bundesqualitätsbericht 2015 (Kiss, Gintenstorfer & Andre, 2015) zum Zeitpunkt der Manuskripterstellung schon online war, sind die Ergebnisse ähnlich. Der Vergleich mit 2012, 2013 und 2014 zeigt also noch keine wesentlichen Fortschritte.

Qualitätsschwerpunkte
Leistungsbeurteilung und
Individualisierung im QIBB

Die Orientierungs-Checks (siehe 3.2.2) werden zur Lernstandsmessung zu Beginn des ersten Schuljahrs der BMHS in fast allen Bundesländern nahezu flächendeckend eingesetzt. Für die HTL zeigte sich darin beispielsweise bei der Auswertung der Erhebung 2012, dass im Bundesland mit dem größten Förderbedarf 28,5 % der Schüler/innen in Deutsch und 25,5 % der Schüler/innen in Mathematik in mehr als einem Kompetenzbereich Förderbedarf aufwiesen. Dem wird mit Förderkursen, verstärkter Individualisierung im Unterricht, individueller Beratung der Schüler/innen sowie ihrer Eltern, Tutorensystemen, einem Coaching bezüglich der Lernstrategien u. a. begegnet. Diese Maßnahmen führten u. a. im Bundesqualitätsschwerpunkt *Schulerfolg*, zu dem die Aktionsfelder „Förderung & Individualisierung“ und „Leistungsbeurteilung“ zählen, zu einer Steigerung des Schulerfolgs⁷ um 5 Prozentpunkte in den drei Jahren des Berichtszeitraums (Dorninger et al., 2013, S. 23 ff.). Die Qualitätsberichte belegen damit, dass in den berufsbildenden Schulen die Unterrichtssituation im für die FLB relevanten Bereich Individualisierung zwar noch verbesserungsbedürftig ist, aber auch schon Weiterentwicklungen der Individualisierungsmaßnahmen erreicht werden konnten, an denen jedoch weiterhin im Rahmen des Qualitätsmanagements QIBB gearbeitet werden muss.

Formative
Lernstandserhebungen

6 Die Schüler/innen bewerten die in den Items beschriebenen Maßnahmen (z. B. „Meine Lehrer/innen beraten mich über meine Stärken und Schwächen im betreffenden Unterrichtsgegenstand“) auf einer sechsstufigen Skala (1: trifft vollständig zu; 6: trifft überhaupt nicht zu). Werte unter 3,0 sind positive Beurteilungen, Werte über 4,0 negative Beurteilungen. 3,9 wird im Rahmen von QIBB daher als gerade noch akzeptabel interpretiert.

7 Zur Berechnung des Schulerfolgsindikators wird der Quotient (in Prozent) aus der Anzahl der Schüler/innen, die sich im Berichtsjahr in den Abschlussklassen der 5-jährigen HTL und der 4-jährigen Fachschulen befinden, und der Anzahl jener Schüler/innen, die 5 Jahre (4 Jahre) vorher in eine HTL (Fachschule) eingetreten sind, ermittelt (Datenquelle: Unterrichts-, Personal-, Informations-System [UPIS-Schülerzahlen]).

3.3 Fazit

Die rechtlichen Rahmenbedingungen und die aufgrund der Schulorganisation notwendigen häufigen Selektionsentscheidungen bedingen das unausgewogene Verhältnis der FLB zur summativen Leistungsbeurteilung. In den letzten Jahren sind einige Initiativen gestartet worden, die die FLB unterstützen könnten, damit sie häufiger praktiziert und vor allem wirksam werden können. Da dieser Begriff nur sehr selten verwendet wird, weil in Österreich dafür keine einheitliche Begrifflichkeit eingeführt wurde, ist es schwierig einzuschätzen, was mit diesen Initiativen wirklich beziehungsweise gefördert wird. Neben der uneinheitlichen Terminologie erschwert auch die Form der Qualitätsberichte in QIBB diese Einschätzung, weil in diesen meist nur beabsichtigte und/oder durchgeführte Maßnahmen genannt werden, deren Häufigkeit und Wirkung man nur vermuten kann. Die wenigen quantitativen Ergebnisse, wie die fachlichen Leistungen und der überfachliche Kompetenzerwerb in der NMS-Evaluierung oder die Screeningergebnisse in QIBB zeigen noch keine oder nur geringe Wirkungen. Daraus lässt sich schließen, dass eine wirksame Ausbalancierung der formativen und der summativen Funktion der Leistungsbeurteilung ein langfristiger Prozess ist, der mit den vorgestellten Entwicklungssträngen erst ansatzweise begonnen wurde.

4 Ausblick und Empfehlungen

4.1 Ausblick und Empfehlungen für die Forschung

Es liegen international bereits viele Studien zur FLB vor, trotzdem gibt es vor allem auf Österreich bezogen noch Forschungsbedarf. Auch in diesem Beitrag wurde die FLB vor allem aus einer lerntheoretischen und allgemeindidaktischen Perspektive behandelt. Die Umsetzung erfolgt jedoch im Fach. Eine Konzeptualisierung und Konkretisierung der FLB in den Fachdidaktiken steht vielfach noch aus. Die FLB in anspruchsvollen und komplexen Wissensgebieten (Maier, 2014) sowie die Entwicklung von schulstufenübergreifenden Lernentwicklungsplänen (siehe 1.4) in den verschiedenen Domänen sind weitere wichtige Forschungs- und Entwicklungsfelder der Fachdidaktiken.

4.1.1 Empirische Studien zur Praxis der formativen Leistungsbewertung in Österreich

Vielfältige
Forschungsprojekte
notwendig

Wie im Abschnitt 3.2.1 berichtet werden seit Jahrzehnten Schulversuche zu den alternativen Formen der Leistungsbeurteilung durchgeführt, sie wurden aber nie evaluiert. Eine „breitflächige Erprobung alternativer Formen der Leistungsbeurteilung [...] in Verbindung mit Evaluationen, in denen Forschung und Entwicklung verknüpft werden“ wurde bereits von Eder et al. (2009, S. 266) im Nationalen Bildungsbericht 2009 (Specht, 2009) empfohlen.

- Solche *Evaluationen der Schulversuche* zu den alternativen Formen der Leistungsbeurteilung könnten zeigen, ob und wie weit diese formativ genützt werden und welche Schlussfolgerungen man daraus für die Übertragung in den Regelfall ziehen kann.
- Derselben Zielstellung können auch *genauere Auswertungen der leistungsbeurteilungsbezogenen Projekte* in IMST, QIBB und SQA durch Forschungsgemeinschaften von Lehrpersonen und universitären Wissenschaftlerinnen und Wissenschaftlern dienen, die auch Aufschluss über die gängigen Bewertungspraktiken geben könnten, insbesondere über die von Lehrpersonen, die Fortbildungsimpulse aufnehmen und Neues probieren.
- Ein *wissenschaftlich konzipierter Feldversuch* kann das Wirkungspotenzial der FLB auf die Motivation, die fachlichen sowie überfachlichen Leistungen und auf das Schulklima im Zusammenspiel mit verschiedenen Formen der summativen Leistungsbeurteilung ausloten. Dies könnte wichtige Impulse für den schulpolitischen Diskurs ergeben.

- Der *Lehrplan der NMS* sowie die Broschüren des ZLS „*Kriterienorientierte Leistungsbeurteilung mit der 4.0-Skala*“ (Schlichtherle, Weiskopf-Prantner & Westfall-Greiter, 2013) und „*Orientierungshilfe Leistungsbeurteilung*“ (Westfall-Greiter, 2012) enthalten spezifische Maßnahmen für die Umsetzung sowohl der formativen als auch der summativen Funktion der Leistungsbeurteilung. Zur Feststellung des Erfolgsausmaßes dieser Weichenstellungen in Richtung Förderorientierung sind Evaluationen und Begleitforschungen über Gelingensbedingungen, Fallstudien guter Praxis sowie deren breite öffentliche Diskussion dringend nötig.

4.1.2 Forschungsdesiderata bezüglich der Integration der sachlogischen und der motivationalen Perspektive

Der Begriff *formative Leistungsbewertung* beziehungsweise *formatives Feedback* wird im Kontext dieses Forschungsausblicks im Einklang mit den bisherigen Ausführungen zur FLB wie folgt verstanden: FLB umfasst – aus der Perspektive der Schüler/innen betrachtet – Maßnahmen selbstorganisierter und/oder durch Lehrpersonen angeregter oder durch sie bereitgestellter Zwischenbilanzierungen in Lernprozessen, aus denen die Lernenden Informationen gewinnen können, in welchem Ausmaß sie welche weiteren sach- und/oder lernverhaltensbezogenen Aktivitäten zur Zielerreichung setzen können. Es wird bei dieser Definition bewusst darauf Wert gelegt, dass im Kontext solcher formativ-orientierender Bilanzierungen im Hinblick auf eine qualitätsvolle Zielerreichung nicht nur auf sachlogische Überlegungen (z. B. „Was habe ich als Schüler/in bisher richtig/falsch gemacht? Welche meiner Stärken könnte ich noch ausbauen und welche Fehler in Zukunft vermeiden?“, „Welche weiteren Informationsquellen brauche ich, um mein Wissen zu erweitern?“) fokussiert wird, sondern mit der Sachlogik gleichwertig auch die motivationale Befindlichkeit der Schülerin beziehungsweise des Schülers in den Blick genommen wird (z. B. „Wodurch fühle ich mich bei der Aufgabenbearbeitung gehemmt/gestresst beziehungsweise angezogen [positiv motiviert]?“, „Was macht es mir schwer/leicht, mich den bei der Aufgabenbearbeitung auftretenden Schwierigkeiten zu stellen?“). Viele der vorliegenden Studien zur FLB weisen fördernde Effekte auf die Motivation und Leistung aus (siehe Abschnitt 2.1.2), die durch die genaueren Selbsteinschätzungen sowie durch die wahrgenommene Nützlichkeit des Feedbacks vermittelt werden. Auch ein Zusammenhang mit den Selbstwirksamkeitsüberzeugungen und den Attributionen der Schüler/innen wurde aufgezeigt. Damit werden einerseits Lernprozesse nicht sachlogisch-kognitiv verengt, sondern unter einer ganzheitlichen Perspektive in den Blick genommen.

Vor dem Hintergrund dieser Überlegungen wird für weitere zukünftige empirische Studien zur FLB vorgeschlagen, differenzielle Aspekte formativen Feedbacks – egal ob durch die Schüler/innen selbst generiert oder durch die Lehrperson an die Schüler/innen herangetragen – entlang der Prüfung folgender Hypothesen zu rekonstruieren (vgl. Kuhl, 2001):

Studien zur FLB aus der Perspektive der Schüler/innen

- Schüler/innen, die *intuitiv-intrinsisch motiviert* beziehungsweise *schwierigkeitsvermeidend lernen mit der Intention, schnell herzeigbare Ergebnisse zu erzielen*, kümmern sich beispielsweise während des Lernprozesses nicht um die Frage der Qualität, in der sie diesen durchlaufen. Für diese ist formatives Feedback durch die Lehrperson wichtig, weil sie es sich nicht selbst geben. Dabei ist in weiterer Folge ein Blick auf die Frage zu werfen, *wie* Lehrpersonen solchen Schülerinnen und Schülern formatives Feedback geben sollen, damit sie leistungsförderliche Effekte erzielen.
- Schüler/innen, die *selbstgesteuert lernen*, geben sich selbst bereits während des Lernprozesses formatives Feedback und leiten daraus konstruktive Konsequenzen ab, und zwar nur, wenn sie auf Schwierigkeiten beim Lernen stoßen. Empirische Studien könnten untersuchen, wie diese Schüler/innen auf *zusätzliches* formatives Feedback reagieren, insbesondere, wenn der Zeitpunkt von der Lehrperson gewählt wird (invasives Vorgehen) und nicht von den Schülerinnen und Schülern (respondentes Lehrer/innen-Verhalten).
- Schüler/innen hingegen, die *ängstlich-zielfixiert lernen mit der Intention, nichts falsch zu machen*, geben sich aufgrund inneren Drucks, das Ziel unter allen Umständen erreichen

zu müssen, selbst kein formatives Feedback. Es herrscht eher die Mentalität „Augen zu und durch“ vor. Wenn Lehrpersonen sachbezogenes Feedback geben, wird es von solchen Schülerinnen und Schülern abgewehrt. Zu erforschen wäre, in welchem Ausmaß formatives Feedback, das weniger auf die Lernergebnisse als auf ihre Einstellungen und Selbststeuerungsstrategien beim Lernen abzielt, ihre Selbstwirksamkeit erhöht.

- Schüler/innen schließlich, die *passiv vermeidend* disponiert sind, beurteilen sich, noch bevor sie mit ihrem Lernprozess beginnen, sehr selbstkritisch daraufhin, ob sie der gestellten Aufgabe gewachsen sind und setzen im Zweifelsfall keine weiteren Aktivitäten. Bezüglich eines solchen Schülerverhaltens stellt sich die Frage, wie Lehrpersonen durch Feedback den Lernprozess in Gang bringen können. Für diesen Schülertypus ist empirisch zu untersuchen, in welchem Ausmaß formatives Feedback, das sich nicht nur auf Lernergebnisse und bisheriges Lernverhalten, sondern insbesondere auf die psychischen Prozesse bezieht, die zu einem Handlungsabbruch geführt haben, lernwirksam ist.

Betrachtet man die aktuelle Forschungslage unter einer solchen Perspektive, werden folgende Punkte deutlich:

Forschungslage

- In vielen Forschungsarbeiten wird wenig beachtet, dass *Lernprozesse zunächst aus einer individuellen Schülerperspektive* zu betrachten sind. Sehr schnell wird auf die Steuerung durch die Lehrperson fokussiert, womit die Eigenaktivitäten der Schüler/innen – auch im Hinblick auf selbstgeneriertes formatives Feedback – aus dem Blick geraten. Es wird zwar auf Begriffe des selbstgesteuerten beziehungsweise selbstregulierten Lernens rekurriert, dies aber manchmal nur auf einer oberflächlichen Ebene mit geringen Auswirkungen auf die Forschungsdesigns.
- Auch die *Differenzierung zwischen sachlogischen und an der psychischen Befindlichkeit orientierten Aspekten formativer Rückmeldungen* soll konsequenter umgesetzt werden. Neuere Curricula mit einer stärkeren Kompetenzorientierung und einer Anbindung an Bildungsstandards können gut für die sachlogische Seite formativer Rückmeldungen genutzt werden; bei einer Orientierung an dieser Struktur bleibt der für die Entwicklung der Fähigkeit und Bereitschaft zum lebenslangen Lernen wichtige Feedbackaspekt auf die psychische Befindlichkeit (vgl. z. B. die Tendenz zum Aktionismus oder zur Schwierigkeitsvermeidung; Gefangensein in passiver Vermeidung) aber eher ausgeklammert.
- In den Abschnitten 2.2.2, 2.3 und 4.2.2 wird konstatiert, dass bei den Lehrpersonen eine *Haltungsänderung* in Richtung einer stärkeren Förderorientierung Voraussetzung für eine effektive FLB wäre, wofür auch schulstrukturelle Zwänge ein Hindernis sind. In künftigen Forschungsarbeiten ist daher auch der Frage nachzugehen, welche Konsequenzen die machtmotivationale Disposition der Lehrperson auf ihr formatives Feedbackverhalten hat. Haltungsänderungen werden damit nicht auf einer ideologischen Ebene angesiedelt, sondern auf der Ebene der Umsetzungsformen des Machtmotivs (vgl. die Operationalisierung des Haltungsbegriffs bei Kuhl, Schwer & Solzbacher, 2014).

4.2 Ausblick und Empfehlungen für die Bildungspolitik

4.2.1 Formative Leistungsbewertung, ein didaktisches Prinzip in den Lehrplänen

Didaktisches Prinzip: Formative Leistungs- rückmeldung und Lernförderung

Die international vorliegenden Forschungsergebnisse weisen die FLB als außerordentlich wirkmächtige Unterrichtsintervention sowohl im Bereich der fachlichen als auch der überfachlichen Kompetenzen aus (siehe Abschnitt 2.3). Bereits im Nationalen Bildungsbericht 2009 (Specht, 2009) wird eine „stärkere Beachtung prüfungsdidaktischer Prinzipien“ gefordert, wobei einige auch für eine formative Wirksamkeit der Leistungsbeurteilung relevant sind, wie „[...] die Steigerung der Frequenz und Qualität der Rückmeldungen (keine Aufgabe ohne Rückmeldung); die Gewährung prüfungsfreier Lernphasen, die dem risikofreien Erproben dienen und in denen (auch missglückte) Leistungen als Erfahrungen interpretiert werden können, die das weitere Lernen positiv beeinflussen, wenn sie reflektiert werden; [...] Förderung der Selbstbeurteilungskompetenz“ (Eder et al., 2009, S. 266).

Es sollten daher möglichst alle Lehrpersonen diese Unterrichtsintervention routinemäßig praktizieren. Um ihnen dies vermitteln zu können, wird eine entsprechende Terminologie benötigt. Der Begriff FLB ist in Österreich bis jetzt nicht sehr gebräuchlich, sodass er – wie in der NMS-Begleitung begonnen – erst bekannt gemacht werden muss. In den vorgestellten österreichischen Entwicklungsansätzen für eine FLB (siehe Abschnitt 3.2) wird der Begriff *expressis verbis* nicht verwendet, an seiner Stelle werden nur einzelne seiner Teilaspekte angeführt, wie (lern)förderliche Rückmeldungen, kriterialnormierte Beurteilung, Stärken-Schwächen-Beratung. Die FLB ist jedoch eine komplexe Unterrichtsintervention (siehe Abschnitt 2.3), in der mehrere Schlüsselstrategien zusammenwirken (Leahy et al., 2005; siehe Abschnitt 1.4). *Formative Leistungsrückmeldung und Lernförderung* entspricht beispielsweise der Komplexität dieser Unterrichtsintervention beziehungsweise dieses Lernarrangements semantisch besser.

Die Verankerung der formativen Leistungsrückmeldung und Lernförderung als didaktisches Prinzip im SchUG und in den Lehrplänen würde zur begrifflichen Klärung beitragen und gleichzeitig zu deren Umsetzung verpflichten.

4.2.2 Schaffung struktureller Bedingungen für eine stärkere Förderorientierung

Eine effektive Umsetzung der FLB (siehe Abschnitte 2.2.2 und 2.3) bedarf einer umfassenden Förderorientierung der Lehrpersonen. Erschwert wird dies im österreichischen Bildungssystem durch die Dominanz der Notenbeurteilung, die vorrangig der Selektion und Allokation dient.

Die Berechtigungsfunktion der Noten spielt in den ersten drei Schulstufen der Grundschule im Vergleich zu den weiterführenden Schulen eine geringere Rolle. Dies sowie die vielen erfolgreichen Schulversuche zu den alternativen Formen der Leistungsbeurteilung in der Grundschule (siehe Abschnitt 3.2.1) sprechen für eine rasche Umsetzung des Vorhabens im aktuellen Regierungsabkommen, die Noten durch Leistungsbeschreibungen in den ersten drei Schulstufen zu ersetzen. Dieser Reformschritt müsste für Fortbildungen und begleitende Unterrichtsentwicklungen genützt werden, damit das Potenzial der alternativen Formen der Leistungsbeurteilung auch als FLB während des Schuljahrs wirksam wird.

Alternative Leistungs-
beschreibung in der
Grundschule

Eine konsequente Förderorientierung sowie die Inklusion von Kindern mit besonderen Bedürfnissen wirft auch die Frage auf, ob auf das Repetieren verzichtet werden kann. Bereits jetzt wird an den Grundschule im Vergleich zu allen anderen Schulformen am wenigsten selektiert. Der Anteil der nicht zum Aufsteigen berechtigten Schüler/innen betrug im Schuljahr 2013/14 nur 0,8 % (bezogen auf die gesamte Grundschule), ein Viertel dieser Kinder erhielt einen sonderpädagogischen Förderbedarf und konnte im Klassenverband verbleiben (Statistik Austria, 2015). Anstelle des Repetierens könnte die im Schuleingangsbereich bereits vorhandene Flexibilität, die Grundstufe I in ein, zwei oder drei Jahren durch Schulstufenwechsel während des Schuljahrs zu bewältigen, ohne die Klassengemeinschaft verlassen zu müssen, auf die gesamte Grundschulzeit ausgedehnt werden. Die bloße „Überleitung der alternativen Leistungsbeurteilungen für die 1. bis 3. Klasse Volksschule in den schulautonomen Gestaltungsraum“ (Bundesministerium für Wissenschaft, Forschung und Wirtschaft & BMBWF, 2015, S. 7), wie sie im Vortrag der Bildungsreformkommission an den Ministerat am 17.11.2015 im Volksschulpaket vorgesehen ist, ist jedoch zu wenig weitgreifend, da damit nur das Schulversuchsprozedere durch das Schulautonomieprozedere ersetzt würde. Anstelle dessen sollte im SchUG der § 18 Abs. 2 so geändert werden, dass in den ersten drei Schulstufen der Volksschule anstelle der Noten eine alternative Form der Leistungsbeurteilung schulautonom festzulegen und zu verwenden ist.

Auch wenn diese Reformschritte erfolgen, werden sich weiterhin die am Ende der 4. Schulstufe notwendigen Schullaufbahnentscheidungen für die HS/NMS oder AHS belastend auf die Förderorientierung der Lehrpersonen und auf die Leistungsbeurteilungssituation der

Grundschule auswirken. Viele Eltern machen Druck auf ihre Kinder und die Lehrperson, um ein Sehr gut oder Gut in Deutsch und Mathematik in der 4. Schulstufe zu erreichen – die Voraussetzung für einen AHS-Besuch. Diese Belastung kann durch eine der folgenden beiden Maßnahmen minimiert werden:

- *Die Validierung der selektionsentscheidenden Noten anhand extern erhobener Testergebnisse durch die Lehrperson:* Dies wäre ein Schritt in Richtung Externalisierung und Objektivierung der mit Berechtigung verbundenen Beurteilungen, die bereits von Eder et al. (2009, S. 265) in ihrem Beitrag im Nationalen Bildungsbericht 2009 empfohlen wurde. Gleichzeitig würde dies die Förderorientierung der Lehrpersonen stärken. Heute könnten dazu die Standardüberprüfungen herangezogen werden, wenn sie rechtzeitig vorliegen können und der § 17 Abs. 1a SchUG sowie die Verordnung zu den Bildungsstandards geändert werden.
- *Eine Verschiebung der Schullaufbahnentscheidung auf eine höhere Schulstufe:* Eine Verschiebung z. B. auf die 8. Schulstufe würde nicht nur eine stärkere Förderorientierung in der Grundschule, sondern auch in der Sekundarstufe I unterstützen. Die Konsequenz ist eine gemeinsame Schule für alle Sekundarstufenschüler/innen. Diese wird schon seit Jahrzehnten zur Verbesserung der Bildungschancengleichheit empfohlen, u. a. auch in den vorhergehenden Nationalen Bildungsberichten 2009 (im Beitrag von Eder et al., 2009) und 2012 (in den Beiträgen von Bruneforth, Weber & Bacher, 2012 sowie Herzog-Punzenberger & Schnell, 2012). Aktuell legt die Industriellenvereinigung (2014) in ihrem Programm „Beste Bildung“ ein Modell für eine solche „neue in sich differenzierte gemeinsame Schule“ vom Schuleintritt bis zur „mittleren Reifeprüfung“ (anhand eines Tests und einer Projektarbeit) für die Grundschule und Sekundarstufe I vor.

4.2.3 Weiterentwicklung der Bildungsstandards und der Informellen Kompetenzmessung

Die österreichischen Bildungsstandards sind als Regelstandards definiert, beschreiben also, welcher durchschnittliche Leistungsstand in den Fächern Deutsch, Mathematik und lebende Fremdsprache am Ende bestimmter Schulstufen erwartet wird (siehe Abschnitt 3.2.2). Die geplante Erweiterung auf die naturwissenschaftlichen Fächer, ohne sie zu testen, ist zu begrüßen. Lernschwache Schüler/innen erfahren jedoch durch die Regelstandards nur ihre Defizite, also wie weit sie unter dem Durchschnitt liegen (vgl. Klieme et al., 2007). Eine Ergänzung durch Mindeststandards ist daher wünschenswert, sodass auch lernschwächere Schüler/innen sich als erfolgreich erleben können. Damit sollte aber auch die Verpflichtung zu schulischen Fördermaßnahmen für jene Schüler/innen einhergehen, die diese noch nicht erreicht haben.

Ergänzung der Regelstandards mit Mindeststandards

- Empfohlen wird daher die Festlegung von Mindeststandards mit der Verpflichtung zur individuellen Förderung (in Kleingruppen, aber auch in zusätzlichen Förderkursen) jener Schüler/innen, die gefährdet sind, diese zu erreichen.
- Ebenso wäre eine Weiterentwicklung der den BIST zugrunde liegenden Kompetenzstrukturmodellen zu Beschreibungen von Lernsequenzen, die zu zunehmend anspruchsvolleren Denkweisen über zunehmend komplexere Thematiken führen, für die FLB hilfreich (siehe Abschnitt 1.4). Diese würden den Lehrpersonen bei der Unterrichtsplanung, der Entwicklung von Beurteilungsrastern zur Bewertung des erreichten Kompetenzniveaus und bei der Bestimmung des nächsten Schritts für einzelne Schüler/innen im Unterricht helfen.
- Damit die Lehrpersonen selbst evaluieren können, ob beziehungsweise in welchem Ausmaß FLB die Effektivität ihres Unterrichts verbessert, sollten ihnen IKM-Pakete für jede Schulstufe (1.–12.) sowie ein größeres Angebot an komplexen und inhaltlich für die Schüler/innen bedeutsamen Diagnoseaufgaben mit lösungsprozess- und/oder kompetenzbezogenen⁸ Zusatzinformationen für die Verwendung im Unterricht zur Verfügung gestellt werden (siehe Abschnitt 2.1.2 sowie Eder, Neuweg & Thonhauser, 2009).

8 Ähnlich den freigegebenen Beispieltitems aus den Standardtestungen.

4.2.4 Weiterentwicklung der Leistungsbeurteilung auf der Grundlage der bestehenden gesetzlichen Bestimmungen im Rahmen des BMBF

Im Abschnitt 3.2.3 wurden die Arbeiten zum ministeriumsinternen Entwurf zur LBVO (BMBF, 2014) vorgestellt. Eine solche Weiterentwicklung der LBVO ist zu begrüßen, da durch ihre Umsetzung längerfristig die Umorientierung von der sozialen zu einer kriterialen Bezugsnorm erreicht werden könnte. Ein besonderer Entwicklungsbedarf liegt beim Herzstück der LBVO, den Kompetenzrastern. Idealerweise sollten diese von den Fachteams an den Schulen entwickelt werden (Professionistenmodell), da sie die Bedingungen des eigenen Standorts mitberücksichtigen können, dabei auch Ownership übernehmen und eine sachgemäße Anwendung zu erwarten ist sowie gleichzeitig ihre Professionalität weiterentwickelt wird.

Eine seriöse Entwicklung der Kompetenzraster erfordert jedoch fundierte fachwissenschaftliche und -didaktische Kompetenzen und entsprechende Arbeitszeit, was nicht an allen Schulen im notwendigen Ausmaß gegeben sein wird. Die Entwicklung der Raster durch Expertinnen und Experten geht vielleicht rascher und sichert die Qualität der Arbeit (Expertenmodell). Wie diese Raster an den Schulen beziehungsweise in den Klassen aufgenommen werden, bleibt jedoch offen. Hier ist ein Weg zu suchen, der den Lehrpersonen grundlegende Arbeiten abnimmt, sie trotzdem in die Entwicklung einbindet und ihnen noch Gestaltungsfreiräume lässt (vgl. Dorninger & Schrack, 2013). Die geplante Entwicklung und Erprobung der Kompetenzraster durch Praktiker/innen in begleiteten Arbeitsgruppen kann helfen diesen Weg zu finden (siehe Abschnitt 3.2.3).

Auch wenn die im ministeriumsinternen Entwurf einer neuen LBVO vorgesehenen Leistungsfeststellungen durch die Leistungsbewertung mit Kompetenzrastern lernförderlicher werden, sollte die FLB nicht auf sie beschränkt sein, da die Leistungsfeststellungen nur in dem Maß durchzuführen sind, wie für eine sichere Kompetenzdiagnose nötig ist (siehe Abschnitt 3.2.3). Sie würden jedoch die FLB sehr erleichtern und auch ergänzen, da sich die FLB und die Leistungsfeststellungen auf dieselben Kompetenzen beziehen und daher die für beide Formen notwendige Klärung der Ziele und der Beurteilungskriterien mit den Schülerinnen und Schülern dieselbe ist. FLB erfordert aber darüber hinaus die Umsetzung der fünf Schlüsselstrategien im Unterricht (Leahy et al., 2005; siehe Abschnitt 1.4), wodurch vom Beginn jedes Lernprozesses an die Schüler/innen kontinuierlich Rückmeldungen samt Lerntipps erhalten und der Unterricht anhand der beobachteten Lernergebnisse möglichst individuell an die jeweiligen Vorkenntnisse und Bedürfnisse der Schüler/innen angepasst wird. Dies sollte daher in der LBVO auch als Verpflichtung verankert werden.

Kompetenzraster

4.2.5 Lehreraus- und Lehrerfortbildung

Um mit FLB systematisch das Lernen zu fördern, um Schülerleistungen zu verbessern, müssen ihre theoretischen Grundlagen sowie die praktischen Umsetzungsmöglichkeiten und Instrumente in der Lehreraus- und Lehrerfortbildung einen Schwerpunkt bilden. Lehrpersonen benötigen für FLB spezifisches pädagogisch-psychologisches, fachdidaktisches und unterrichtsmethodisches Wissen sowie Kenntnisse, wie man das Vorwissen der Schüler/innen erfasst und passende Lernarrangements für FLB schafft, damit ihre Schüler/innen die Fähigkeit zur Selbst- und Peereinschätzung entwickeln und die Lehrpersonen selbst die richtigen Schlüsse aus den Beobachtungen für die Adaptierung ihres Unterrichts ziehen können.

Für entsprechende Fortbildungsprogramme, die sich über längere Zeit erstrecken sollten, empfiehlt sich eine Kombination von Workshops und schulinternen Lerngemeinschaften, eine Vorgangsweise, die sich in Evaluationen von „Keep Learning on Track“⁹ erfolgreich er-

Workshops und
Lerngemeinschaften
über längere Zeit

⁹ Dieses Programm wurde von Wiliam und Leahy entwickelt und wird heute mit dem Titel „Embedding Formative Assessment“ von <http://www.ssatuk.co.uk/cpd/teaching-and-learning/embedding-formative-assessment/> [zuletzt geprüft am 01.12.2015] vertrieben.

wiesen hat. Der systematische Austausch sowie die gemeinsame Reflexion von Erfahrungen, die bei Versuchen von FLB im eigenen Unterricht gemacht wurden, und die Planung weiterer Handlungsschritte in Lerngemeinschaften schaffen günstige Voraussetzungen für die tatsächliche Veränderung des Unterrichts sowie für die Anpassung der FLB an die eigenen Erfordernisse (Wiliam, 2006).

Literatur

Altrichter, H. & Nagy, G. (2010). *Forschungsprojekt „Geschlechteraspekte in der schulischen Leistungsbewertung“* (Endbericht im Auftrag von IMST). Zugriff am 26.02.2016 unter https://www.bmbf.gv.at/schulen/unterricht/ba/vs_fb_t1_21126.pdf?4dzgm2

Altrichter, H., Nagy, G. & Pocrnja, M. (2015). Akteure und ihre Einschätzungen: Die Neue Mittelschule in der Wahrnehmung der Schulpartner und der Öffentlichkeit. In F. Eder, H. Altrichter, F. Hofmann & C. Weber (Hrsg.), *Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten. Forschungsbericht* (403–442). Graz: BIFIE. Zugriff am 02.05.2015 unter https://www.bmbf.gv.at/schulen/bw/nms/eval_forschungsbericht.pdf?4sr7p3

Amrhein-Kreml, R., Breyer, G., Dobler, K., Koenne, C., Mayr, J. & Schuster, A. (2008). *Prüfungskultur. Leistung und Bewertung in der Schule*. Klagenfurt: Institut für Unterrichts- und Schulentwicklung der Alpen-Adria-Universität.

Ayala, C. C., Shavelson, R. J., Ruiz-Primo, M. A., Brandon, P. R., Yin, Y., Furtak, E. M. et al. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education*, 21 (4), 315–334.

Babel, H. & Spitzbart, G. (2014). *Bundesqualitätsbericht des berufsbildenden Schulwesens für den Berichtszeitraum 2010–2012 und den Planungszeitraum 2012–2014. Bericht der QIBB Steuergruppe*. Zugriff am 30.10.2015 unter https://www.qibb.at/fileadmin/content/QIBB/Dokumente/Q-Berichte/BBS-Bundes-Q-Bericht_2010-12_QIBB-STG_FINAL.pdf

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles Policy and Practice*, 18 (1), 5–25.

Besser, M., Tropper, N. & Leiss, D. (2013, März). *Lehrern Lehren lehren – Entwicklung und Evaluation von Lehrerfortbildungen zu formativem Assessment*. Vortrag auf der Tagung der Gesellschaft für Didaktik der Mathematik (GDM) in Münster.

Black, P. J. & Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80 (2), 139–148.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. W. Tyler (Hrsg.), *Educational evaluation: new roles, new mean* (the 68th yearbook of the National Society for the Study of Education, Part II, S. 26–50). Chicago, IL: University of Chicago Press.

Bohl, T. (2004). *Prüfen und Bewerten im Offenen Unterricht*. Weinheim: Beltz.

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M. & Stobart, G. (2002). *Assessment for learning: 10 principles*. Cambridge, UK: University, School of Education.

Bruneforth, M., Weber, C. & Bacher, J. (2012). Chancengleichheit und garantiertes Bildungsminimum in Österreich. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 189–228). Graz: Leykam. Zugriff am 27.10.2015 unter <https://www.bifie.at/buch/1915>

Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE; Hrsg.). (2014). *Themenheft für die Kompetenzbereiche „Einsicht in Sprache durch Sprachbetrachtung“ und „Rechtschreiben“*. Deutsch, Lesen, Schreiben. Volksschule Grundstufe I + II. Graz: Leykam. Zugriff am 23.10.2015 unter <https://www.bifie.at/node/2708>

Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE). (2015). *Informelle Kompetenzmessung (IKM). Nutzung im Frühjahr 2015*. Unveröffentlichter Auswertungsbericht.

Bundesministerium für Bildung und Frauen (BMBF). (2014). *Ministeriumsinterner Entwurf einer Verordnung über die Beurteilung von Schülerinnen- und Schülerleistungen an Pflichtschulen sowie mittleren und höheren Schulen mit Vorbemerkungen von Neuweg, G. H.* (Paper zur Vorbereitung der Enquete zur Leistungsbeurteilung am 25.04.2014 im BMBF). Unveröffentlichtes Dokument, BMBF, Wien.

Bundesministerium für Bildung und Frauen (BMBF). (2015). *SQA – Schulqualität Allgemeinbildung: Richtlinien für das Schuljahr 2015/16. Rundschreiben Nr. 13/2015*. Wien: BMBF. Zugriff am 26.02.2016 unter https://www.bmbf.gv.at/ministerium/rs/2015_13.html

Bundesministerium für Unterricht, Kunst und Kultur (BMUKK). (2012). *Lehrplan der Neuen Mittelschule*. Verfügbar am 23.10.2015 unter https://www.bmbf.gv.at/schulen/unterricht/lp/lp_nms.html

Bundesministerium für Unterricht, Kunst und Kultur (BMUKK). (2013). *SQA – Schulqualität Allgemeinbildung: Richtlinien für das Schuljahr 2013/14*. Rundschreiben Nr. 14/2013. Wien: BMUKK. Zugriff am 09.08.2015 unter http://www.sqa.at/pluginfile.php/1369/course/section/546/rundschreiben_nr_14_2013.pdf

Bundesministerium für Wissenschaft, Forschung und Wirtschaft (BMWFW) & Bundesministerium für Bildung und Frauen (BMBF). (2015). *Bildungsreformkommission – Vortrag an den Ministerrat*. Zugriff am 13.01.2016 unter <https://www.bmbf.gv.at/ministerium/vp/2015/20151117.pdf>

Bürgermeister, A., Klieme, E., Rakoczy, K., Harks, B. & Blum, W. (2014). Formative Leistungsbeurteilung im Unterricht: Konzepte, Praxisberichte und ein neues Diagnoseinstrument für das Fach Mathematik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufdiagnostik* (S. 41–60). Göttingen: Hogrefe.

Cowie, B. & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles Policy and Practice*, 6 (1), 32–42.

Czerwanski, A., Solzbacher, C. & Vollstädt, W. (Hrsg.). (2002). *Förderung von Lernkompetenz in der Schule. Band 1: Recherche und Empfehlungen*. Gütersloh: Bertelsmann Stiftung.

Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, 223–238.

Diehl, K. & Hartke, B. (2012). IEL-1. *Inventar zur Erfassung der Lesekompetenzen von Erstklässlern. Ein curriculumbasiertes Verfahren zur Abbildung des Lernfortschritts*. Göttingen: Hogrefe.

Dinges, S. & Egger, M. (2015). *BIST-Begleitforschung: Rezeption und Nutzung von Ergebnissen der Bildungsstandardüberprüfung in Mathematik auf der 4. Schulstufe unter Berücksichtigung der Rückmeldemoderation im Bundesland Oberösterreich*. Salzburg: BIFIE. Zugriff am 01.08.2015 unter <https://www.bifie.at/node/3090>

Dorninger, C., Niemeyer, S., Pachatz, W., Pregesbauer, C. & Winkler-Rigler, G. (2013). *QIBB. Bundesqualitätsbericht HTL-Q-SYS für den Berichtszeitraum 2010–2012*. Zugriff am 01.08.2015 unter https://www.qibb.at/fileadmin/content/QIBB/Dokumente/Q-Berichte/II2_Bundes-Q-Bericht_HTL_2010-12_Sept._2013.pdf

Dorninger, C. & Schrack, C. (2013). Kompetenzorientierte Leistungsbeurteilung aus Sicht des berufsbildenden Schulwesens. *Erziehung und Unterricht*, 163 (S. 9–10), 795–803.

Dunn, K. E. & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, 14 (7), 1–11. Zugriff am 15.06.2015 unter <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.4370&rep=rep1&type=pdf>

Eder, F. (2009). Die Schule der 10- bis 14-Jährigen als Angelpunkt der Diskussion um Struktur und Qualität des Schulsystems. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich 2009, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (36–57). Graz: Leykam. Zugriff am 27.10.2015 unter <https://www.bifie.at/buch/936>

Eder, F., Altrichter, H., Bacher, J., Hofmann, F. & Weber, C. (2015). Executive Summary. In F. Eder, H. Altrichter, F. Hofmann & C. Weber (Hrsg.), *Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten*. Forschungsbericht. Graz: BIFIE.

Eder, F., Neuweg, G. H. & Thonhauser, J. (2009). Leistungsfeststellung und Leistungsbeurteilung. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich 2009, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 245–267). Graz: Leykam. Zugriff am 27.10.2015 unter <https://www.bifie.at/buch/1024>

FAST SCASS (Hrsg.). (2008). *Attributes of effective formative assessment*. A work product coordinated by Sarah McManus, NC Department of Public Instruction. Paper prepared for Formative Assessment for Students and Teachers (FAST), State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO). Zugriff am 11.06.2015 unter http://www.ccsso.org/Resources/Publications/Attributes_of_Effective_Formative_Assessment.html

Harks, B., Rakoczy, K., Hattie, J., Besser, M. & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: The role of feedback's perceived usefulness. *Educational Psychology*, 34 (3), 269–290.

Harks, B., Rakoczy, K., Klieme, E., Hattie, J. & Besser, M. (2014). Indirekte und moderierte Effekte von schriftlicher Rückmeldung auf Leistung und Motivation. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 163–194). Münster: Waxmann.

Hattie, J. (2012). *Visible learning for teachers. Maximizing impact of learning*. London: Routledge.

Hattie, J. (2013). *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“ besorgt von W. Beywl, K. Zierer*. Baltmannsweiler: Schneider.

Hattie, J. & Wollenschläger, M. (2014). A conceptualization of feedback. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 135–149). Münster: Waxmann.

Helmke, A. (2014). *Unterrichtsqualität und Lehrprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.

Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Paper prepared for Formative Assessment for Students and Teachers (FAST), State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO). Zugriff am 11.06.2015 unter http://www.ccsso.org/Resources/Publications/Learning_Progressions_Supporting_Instruction_and_Formative_Assessment.html

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Paper presented for the Council of Chief State School Officers. Zugriff am 25.02.2013 unter http://www.ccsso.org/Resources/Publications/Formative_Assessment_and_Next-Generation_Assessment_Systems.html

Herzog-Punzenberger, B. (Hrsg.). (2012). *Nationaler Bildungsbericht Österreich 2012, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen*. Graz: Leykam. Zugriff am 01.12.2015 unter <https://www.bifie.at/buch/1024>

Herzog-Punzenberger, B. & Schnell, P. (2012). Die Situation mehrsprachiger Schüler/innen im österreichischen Schulsystem – Problemlagen, Rahmenbedingungen und internationaler Vergleich. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 229–267). Graz: Leykam. Zugriff am 27.10.2015 unter <https://www.bifie.at/buch/1915>

Hofmann, F. (2000). *Aufbau von Lernkompetenzen fördern. Neue Wege zur Realisierung eines bedeutsamen pädagogischen Ziels*. Innsbruck: StudienVerlag.

Hofmann, F. & Katstaller, M. (2015). Innere Differenzierung. In F. Eder, H. Altrichter, F. Hofmann & C. Weber (Hrsg.), *Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten. Forschungsbericht* (S. 165–177). Graz: Leykam. Zugriff am 27.10.2015 unter https://www.bmbf.gv.at/schulen/bw/nms/eval_forschungsbericht.pdf?4sr7p3

Industriellenvereinigung (IV). (2014). *Beste Bildung für Österreichs Zukunft. Bildung neu denken. Schule besser leben*. Zugriff am 20.05.2015 unter http://www.iv-net.at/iv-all/publikationen/file_657.pdf

Ingenkamp, K. (1971). *Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz.

Jürgens, E. & Sacher, W. (2008). *Leistungserziehung und pädagogische Diagnostik in der Schule. Grundlagen und Anregungen für die Praxis*. Stuttgart: Kohlhammer.

Kiss, K., Gintinstorfer, A. & Andre, G. (2015). *QIBB. Bundesqualitätsbericht der kaufmännischen Schulen für den Berichtszeitraum 2012–2014 und den Planungszeitraum 2014–2016. Bericht der Abteilung III/3 des Bundesministeriums für Bildung und Frauen (BMBF)*. Zugriff am 15.07.2015 unter https://www.hak.cc/files/attachments/service_attachments/8.7.2015%20Bundesqualit%C3%A4tsbericht%20kfm%20Schulen%20final.pdf

Klauer, K. J. (2014). Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufdiagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufdiagnostik* (S. 1–17). Göttingen: Hogrefe.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2007). Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. In Bundesministerium für Bildung und Forschung (BMBF; Hrsg.), *Zur Entwicklung nationaler Bildungsstandards. Expertise* (Bildungsforschung, Band 1). Zugriff am 27.10.2015 unter https://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf

Klieme, E., Bürgermeister, A., Harks, B., Blum, W., Leiß, D. & Rakoczy, K. (2010). Leistungsbeurteilung und Kompetenzmodellierung im Mathematikunterricht. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. Zeitschrift für Pädagogik, 56. Beiheft*, 64–74.

Kobarg, M., Thoma, G., Dalehefte, I. M., Seidel, T. & Prenzel, M. (2011). Lernwirksame Unterrichtsbedingungen in der Unterrichtsplanung berücksichtigen. In K. Zierer (Hrsg.), *Jahrbuch der Allgemeinen Didaktik 2011* (S. 46–58). Baltmannsweiler: Schneider.

Kuhl, J. (2001). *Motivation und Persönlichkeit. Interaktionen psychischer Systeme*. Göttingen: Hogrefe.

Kuhl, J., Schwer, C. & Solzbacher, C. (2014). Professionelle pädagogische Haltung: Persönlichkeitspsychologische Grundlagen. In C. Schwer & C. Solzbacher (Hrsg.), *Professionelle pädagogische Haltung. Historische, theoretische und empirische Zugänge zu einem viel strapazierten Begriff* (S. 79–106). Bad Heilbrunn: Klinkhardt.

Kuhlmann, K. & Hartke, B. (2011). FE-RS 2. *Formative Erfassung der Rechtschreibleistung im 2. Schuljahr*. Rostock: Universität Rostock.

Leahy, S., Lyon, C., Thompson, M. & Wiliam, D. (2005). Classroom assessment: Minute-by-minute and day-by-day. *Educational Leadership*, 63 (3), 18–24.

Mahlau, K., Blumenthal, Y., Diehl, K., Schöning, A., Sikora, S., Voß, S. et al. (2014). Das Rügener Inklusionsmodell (RIM) – RTI in der Praxis. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufdiagnostik* (S. 101–125). Göttingen: Hogrefe.

Maier, U. (2014). Formative Leistungsdiagnostik in der Sekundarstufe – Grundlegende Fragen, domänenspezifische Verfahren und empirische Befunde. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufdiagnostik* (S. 19–39). Göttingen: Hogrefe.

Mietzel, G. (2001). *Pädagogische Psychologie des Lernens und Lehrens*. Göttingen: Hogrefe.

Narciss, S. (2014). Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 43–82). Münster: Waxmann.

Neuweg, G. H. (2014). *Schulische Leistungsbeurteilung: Rechtliche Grundlagen und pädagogische Hilfestellungen für die Schulpraxis* (5., aktualisierte und erw. Aufl.). Linz: Trauner.

Nicholls, J. G. (1984). Achievement motivation: Conception of ability, subjective experience, task choice and performance. *Psychological Review*, 91, 328–346.

Rechnungshof. (2015). *Bericht des Rechnungshofes zu den Schulversuchen im Schuljahr 2012/13*. Zugriff am 08.08.2015 unter http://www.rechnungshof.gv.at/fileadmin/downloads/_jahre/2015/berichte/teilberichte/bund/Bund_2015_01/Bund_2015_01_3.pdf

Republik Österreich. (2013). *Arbeitsprogramm der österreichischen Bundesregierung 2013–2018. Erfolgreich. Österreich*. Zugriff am 04.08.2015 unter <https://www.bka.gv.at/DocView.axd?CobId=53264>

Rieß, C. & Zuber, J. (2014). *BIST-Begleitforschung 2/2014: Rezeption und Nutzung von Ergebnissen der Bildungsstandardüberprüfung in Mathematik auf der 8. Schulstufe unter Berücksichtigung der Rückmeldemoderation*. Salzburg: BIFIE. Zugriff am 20.05.2015 unter <https://www.bifie.at/node/2658>

Roos, M. (2001). *Ganzheitliches Beurteilen und Fördern in der Primarschule*. Chur: Rüegger.

Roos, M., Huber, S., Sandmeier, A., Sempert, W. & Schuler, P. (2007). *Beurteilen und Fördern auf der Sekundarstufe I des Kantons Zug*. Zugriff am 18.07.2015 unter http://www.bildungsmanagement.net/pdf/IBB_Evaluation_BF.pdf

Ruf, U. & Gallin, P. (2011). Erkennen und bewerten von Leistungen im Dialogischen Unterricht. In W. Sacher & F. Winter (Hrsg.), *Diagnose und Beurteilung von Schülerleistungen* (S. 231–249). Baltmannsweiler: Schneider.

Schabmann, A., Landerl, K., Bruneforth, M. & Schmidt, B. M. (2012). Lesekompetenz, Leseunterricht und Leseförderung im österreichischen Schulsystem. Analysen zur pädagogischen Förderung der Lesekompetenz. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 17–69). Graz: Leykam. Zugriff am 27.10.2015 unter <https://www.bifie.at/buch/1915>

Schiefele, U. & Pekrun, R. (1996). Psychologische Modelle des fremdgesteuerten und selbstgesteuerten Lernens. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Pädagogische Psychologie. Band 2: Psychologie des Lernens und der Instruktion* (S. 249–278). Göttingen: Hogrefe.

Schlichtherle, B., Weiskopf-Prantner, V. & Westfall-Greiter, T. (2013). *Kriterienorientierte Leistungsfeststellung mit der 4.0-Skala*. Innsbruck: Zentrum für lernende Schulen. Verfügbar am 14.12.2014 unter <http://www.nmsvernetzung.at/>

Schmidinger, E. (2012). Beurteilen – ein Element des Unterrichts. In E. Schmidinger & R. Vierlinger (Hrsg.), *Zeitgemäße Leistungsbeurteilung* (S. 11–116). Wien: Jugend und Volk.

Schmidinger, E. (2013). Formative Leistungsbeurteilung. *Erziehung und Unterricht*, 163 (9–10), 776–785.

Schratz, M. & Westfall-Greiter, T. (2010). Das Dilemma der Individualisierungsdidaktik. Plädoyer für ein personalisiertes Lernen in der Schule. *Journal für Schulentwicklung*, 14 (1), 18–31.

Schuster, A. (2013). Themenprogramm „Prüfungskultur“. In K. Krainer, H. Senger & I. Andreitz (Hrsg.), *Endbericht zum Projekt IMST 2010–2012* (S. 129–137). Klagenfurt: Alpen-Adria-Universität, Institut für Unterrichts- und Schulentwicklung.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné & M. Scriven (Hrsg.), *Perspectives of curriculum evaluation* (S. 39–83). Chicago, IL: Rand McNally.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A. et al. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21 (4), 295–314.

Shepard, L. A., Hammerness, K., Darling-Hammond, L., Rust, F., Snowden, J. B., Gordon, E. et al. (2005). Assessment. In L. Darling-Hammond & J. Bransford (Hrsg.), *Preparing teachers for a changing world: what teachers should learn and be able to do* (S. 275–326). San Francisco, CA: Jossey-Bass.

Smit, R. (2009). *Die formative Beurteilung und ihr Nutzen für die Entwicklung der Lernkompetenz*. Baltmannsweiler: Schneider.

Souvignier, E., Förster, N. & Salaschek, M. (2014). quop: Ein Ansatz internetbasierter Lernverlaufsdagnostik mit Testkonzepten für Lesen und Mathematik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdagnostik* (S. 239–256). Göttingen: Hogrefe.

Souvignier, E., Förster, N. & Schulte, E. (2014). Wirksamkeit formativen Assessments – Evaluation des Ansatzes der Lernverlaufsdagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdagnostik* (S. 221–238). Göttingen: Hogrefe.

Specht, W. (Hrsg.). (2009). *Nationaler Bildungsbericht Österreich 2009, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen*. Graz: Leykam. Zugriff am 01.12.2015 unter <https://www.bifie.at/buch/1024>

Sperl, H., Kiss, K. & Müllauer-Hager, B. (2013). *QIBB. Bundesqualitätsbericht der kaufmännischen Schulen für den Berichtszeitraum 2010/2011/2012*. Zugriff am 01.08.2015 unter https://www.qibb.at/fileadmin/content/QIBB/Dokumente/Q-Berichte/II3_Bundes-Q-Bericht_HAK_2010-12_April_2013.pdf

Statistik Austria (Hrsg.). (2015). *Bildung in Zahlen 2013/14. Schlüsselindikatoren und Analysen*. Wien: Statistik Austria. Zugriff am 17.08.2015 unter http://www.statistik.at/web_del/services/publikationen/5/index.html?includePage=detailedView&pubId=508§ionName=Bildung%2C+Kultur

Stern, T. (2010). *Förderliche Leistungsbewertung. Österreichisches Zentrum für Persönlichkeitsbildung und soziales Lernen (ÖZEPS)*. Wien: Bundesministerium für Unterricht, Kunst und Kultur (BMUKK).

Strathmann, A. M. & Klauer, K. J. (2012). LVD-M 2–4. *Lernverlaufsdagnostik-Mathematik für die zweiten bis vierten Klassen*. Göttingen: Hogrefe.

Strathmann, A. M., Klauer, K. J. & Greisbach, M. (2010). Lernverlaufsdagnostik – Dargestellt am Beispiel der Entwicklung der Rechtschreibkompetenz in der Grundschule. *Empirische Sonderpädagogik, 1*, 64–77.

Valtin, R. & Wagner, C. (2002). Wie wirken sich Notenzeugnisse und verbale Beurteilung auf die leistungsbezogene Persönlichkeitsentwicklung aus? In R. Valtin (Hrsg.), *Was ist ein gutes Zeugnis? Noten und verbale Beurteilungen auf dem Prüfstand* (S. 113–137). Weinheim: Juventa.

Vierlinger, R. (2012). Die Defizite der Ziffernnote und die Direkte Leistungsvorlage (DLV) als Alternative. In E. Schmidinger & R. Vierlinger (Hrsg.), *Zeitgemäße Leistungsbeurteilung* (S. 117–220). Wien: Jugend und Volk.

Voß, S. & Hartke, B. (2014). Curriculumbasierte Messverfahren (CBM) als Methode der formativen Leistungsdiagnostik im RTI-Absatz. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdagnostik* (S. 83–99). Göttingen: Hogrefe.

Vygotskij, L. (2002). *Denken und Sprechen* (J. Lompscher & G. Rückriem, Hrsg. und Übers.). Weinheim: Beltz.

Walter, J. (2010). *LDL. Lernfortschrittsdiagnostik Lesen. Ein curriculumbasiertes Verfahren*. Göttingen: Hogrefe.

Walter, J. (2013). *VSL. Verlaufsdiagnostik sinnerfassendes Lesen*. Göttingen: Hogrefe.

Weinert, F. E. (1996). Lerntheorien und Instruktionsmodelle. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion* (Enzyklopädie der Psychologie. Pädagogische Psychologie, Band 2, S. 1–48). Göttingen: Hogrefe.

Westfall-Greiter, T. (2012). *Orientierungshilfe Leistungsbeurteilung. Teil 1–3. ZLS-NMSEB*. Verfügbar am 02.12.2014 unter www.nmsvernetzung.at

Wiliam, D. (2006, Juli). *Does Assessment hinder Learning?* Paper presented at ETS Invitational Seminar held at the Institute of Civil Engineers, London.

Wiliam, D. (2013). Assessment: The Bridge between Teaching and Learning. *Voices from the Middle*, 21 (2), 15–20.

Wiliam, D. (2014, April). *Formative assessment and contingency in the regulation of learning processes*. Paper presented in a symposium entitled Toward a Theory of Classroom Assessment as the Regulation of Learning at the annual meeting of the American Educational Research Association (AERA), Philadelphia, PA. Zugriff am 05.06.2015 unter http://www.dylanwiliam.org/Dylan_Wiliams_website/Papers.html

Wilson, M. & Bertenthal, M. (Hrsg.). (2005). *Systems for state science assessment*. Washington, DC: National Academies Press.

Winter, F. (2004). *Leistungsbewertung. Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen*. Baltmannsweiler: Schneider.

Winter, F. (2015). *Lerndialog statt Noten. Neue Formen der Leistungsbeurteilung (Pädagogik-praxis)*. Weinheim: Beltz.

Wollenschläger, M. (2015). Kompetenzorientiertes Schülerfeedback – Forschungsergebnisse und Empfehlungen. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 263–281). Weinheim: Beltz.

Yin, Y., Tomita, M. K. & Shavelson, R. J. (2015). Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *International Journal of Science Education*, 36 (4), 531–552.

Zierer, K., Busse, V., Otterspeer, L. & Wernke, S. (2015). Feedback in der Schule – Forschungsergebnisse. In C. G. Buhren (Hrsg.), *Handbuch Feedback in der Schule* (S. 31–50). Weinheim: Beltz.

Zug, U. (2013). *QIBB. Bundesqualitätsbericht der Bildungsanstalten für Kindergartenpädagogik und Bildungsanstalten für Sozialpädagogik für den Berichtszeitraum 2010/2011/2012 und den Planungszeitraum 2013/2014/2015*. Zugriff am 01.08.2015 unter https://www.qibb.at/fileadmin/content/QIBB/Dokumente/Q-Berichte/I1_Bundes-Q-Bericht_BAKIP-BASOP_Juli_2013.pdf