

# Leistungsbeurteilung in Deutschland und Italien – Bezugsnormorientierung im kontrastiven Vergleich

Yannik Wilke<sup>2</sup> & Valerio Rigo<sup>1</sup>

<sup>1</sup> *Freie Universität Bozen, Fakultät für Erziehungswissenschaft Universität  
Bielefeld, Fakultät für Erziehungswissenschaft,  
Wissenschaftliche Einrichtung Laborschule  
Kontakt: [yannik.wilke@uni-bielefeld.de](mailto:yannik.wilke@uni-bielefeld.de), [valerio.rigo@student.unibz.it](mailto:valerio.rigo@student.unibz.it)*

**Zusammenfassung:** Während das deutsche Schulsystem immer noch auf einer über 100 Jahre alten Grundidee fußt und nur langsam auf (neue) Anforderungen reagiert, hat Italien schon in den 1970er Jahren begonnen sein Schulsystem zu transformieren. Entsprechend sind die Voraussetzungen für die Umsetzung inklusionsbezogener (Weiter-)Entwicklungen in Italien deutlich anschlussfähiger, als dies in Deutschland der Fall ist. Das zeigt sich insbesondere in der Perspektive auf die schulischen Leistungen von Schülerinnen und Schülern. Im folgenden Beitrag wird insbesondere die Bezugsnormorientierung im Kontext von Inklusion und deren Implementation in den beiden Systemen verglichen, ehe darauf eingegangen werden kann, welche Potenziale und Hürden sich für beide Schulsysteme zeigen.

**Schlagwörter:** Leistungsbeurteilung, Inklusion, Italien, Nordrhein-Westfalen



## 1 Einleitung

Das Erfassen, Beurteilung und das Rückmelden von Schüler\*innenleistungen stellen zentrale Bereiche des alltäglichen Handelns von Lehrkräften dar. Entsprechende Methoden und Verfahren sind prägend für die Unterrichtsgestaltung und didaktische Entscheidungen (Jürgens, 2019). Jedoch herrscht in verschiedenen Diskurslinien um die Beurteilung von Leistungen wenig Einigkeit darüber, was qualitativ hochwertige und pädagogisch sinnvolle oder gar vertretbare Orientierungspunkte dafür sind, was einerseits Leistung ausmacht, und andererseits, mit welchen Methoden und Verfahrensweisen diese erhoben, beurteilt und rückgemeldet werden können (u.a. Seitz & Wilke, 2021; Thurn, 2011; Jürgens & Lissmann, 2015). So fokussiert Leistungsbeurteilung in den verschiedenen Betrachtungsweisen Fragen der Zuteilung zu verschiedenen Bildungsgängen und Kurssystemen als Teil von gesellschaftlichen Anforderungen an schulische Ausbildungen (Jürgens & Lissmann, 2015), wohingegen die eher reformpädagogisch orientierten Konzepte nach einer Potenzialentfaltung, Steigerung der Lernmotivation und der Lernleistungen fragen (u.a. Prengel, 2016; Seitz & Wilke, 2021). Die Problematik dabei liegt vor allem darin, dass gesellschaftliche und pädagogische Funktionen widersprüchlich zueinander sind, jedoch beide gleichzeitig gültig sind (Jürgens, 2012).

Entsprechend entfalten sich die maßgeblichen Unterschiede in den verschiedenen Diskurslinien zur Leistungsbeurteilung in den

- a) Funktionen der Leistungsbeurteilung
- b) Verfahren der Rückmeldungen (u.a. Pro und Contra von Noten)
- c) herangezogenen Vergleichsmaßstäben der Leistungsbeurteilung.

Die Unterschiede in den herangezogenen Vergleichsmaßstäben werden im folgenden Kapitel erläutert und kontextualisiert, ehe die Bezugsnormorientierungen in den Ländern Deutschland (am Beispiel von Nordrhein-Westfalen) und Italien (am Beispiel Trentino-Südtirol) beleuchtet werden.

## 2 Bezugsnormorientierung – Was sind Maßstäbe für die Lernentwicklung von Schüler\*innen?

Im Kontext der Leistungsbeurteilung spielen Bezugsnormen eine große Rolle, da jeder Beurteilungsprozess – ob bewusst oder unbewusst – vor dem Hintergrund von Normen und Vergleichsmaßstäben abläuft. Diese dienen der Einordnung und Klassifikation von Leistungen in verschiedenen Kontexten (s. Kasten).

Beispiel: Ein\*e Läufer\*in legt eine definierte Strecke innerhalb einer bestimmten Zeit zurück. Ob diese Leistung als ‚gute Leistung‘ anerkannt werden kann, hängt vom Maßstab ab, der angelegt wird:

- Handelt es sich bei der gemessenen Zeit um eine Steigerung der Einzelperson auf dieser Strecke?
- Wie verhält sich die Zeit zu den Vorgaben der Bundesjugendspiele/einem Sportabzeichen/...?
- Gelingt die Person vor einer (oder mehreren) anderen ins Ziel? Bzw. gibt es Menschen, die die Strecke langsamer oder schneller schaffen? Handelt es sich bei der gelaufenen Zeit vielleicht um einen neuen Rekord?
- Reicht die erzielte Zeit für eine Qualifikation für einen anderen Wettbewerb oder die Zulassung zu einem bestimmten Beruf/einer bestimmten Ausbildung?

Kasten 1: Das Beispiel vom Läufer zur Illustration von Bezugsnormen. Heckhausen (1974) beschreibt vier verschiedene Maßstäbe, die für die Beurteilung einer Leistung herangezogen werden können:

- die individuelle Bezugsnorm – verstanden als der Vergleich einer Leistung mit den vorherigen Leistungen derselben Person;
- die sachliche Bezugsnorm – verstanden als der Vergleich einer Leistung mit dem Anspruch, den der Gegenstand selbst festlegt – z.B. Vollständigkeit, sachliche Richtigkeit von Darstellungen;
- die soziale Bezugsnorm – verstanden als der Vergleich einer Einzelleistung mit den Leistungen einer Bezugsgruppe;
- die fremdgesetzte Bezugsnorm – verstanden als der Abgleich einer Leistung mit Ansprüchen, die von einer Institution vorgegeben werden.

Insbesondere im schulpädagogischen Kontext trugen die Arbeiten von Rheinberg (1977, 1980, 1982, 2001a, b) maßgeblich zur Aufklärung der Bezugsnormorientierungen von Lehrkräften und deren Effekte auf das Verhalten von Schüler\*innen- und Lehrer\*innen bei. Rheinberg (1977, 2001a) beschreibt im Gegensatz zu Heckhausen (1974) lediglich drei Bezugsnormen, die bei verschiedenen Autor\*innen zwar unterschiedlich benannt werden, aber inhaltlich identisch sind:

- individuelle oder individualisierte Norm – der Vergleichsmaßstab für eine gezeigte Leistung sind die vorherigen Leistungen des Kindes und seine (formativ erhobene) Entwicklung (u.a. Bohl, 2009; Thurn, 2012);
- soziale oder sozialvergleichende Norm – der Vergleichsmaßstab für Leistungen ist die soziale Gruppe, in der die Leistung erhoben wird (Klasse, Jahrgang, Bundesland etc.) (Jürgens, 2012; Bohl 2009);
- kriteriale oder curriculare Norm – der Vergleichsmaßstab für Leistungen wird anhand des Lerngegenstandes (oder durch das Curriculum) vorgegeben (Prenzel, 2016; Bohl, 2009).

Wie Studien zeigen, prägen Bezugsnormorientierungen (im Weiteren: BnO) von Lehrkräften die Kausalattribution von Schüler\*innenleistungen sowie den Umgang mit didaktischen Maßnahmen wie Sanktionen und Individualisierung (Jürgens & Lissmann, 2015). So halten Jürgens und Lissmann (2015) fest, dass Lehrer\*innen mit stärkerer sozialer BnO Leistungen von Schüler\*innen überwiegend mit zeitstabilen Faktoren wie Begabung, Arbeitshaltung und primärem Herkunftshabitus begründen, wohingegen Lehrkräfte mit individueller BnO zeitvariable Faktoren wie Interesse, Unterrichtsgestaltung und -inhalte als Faktoren benennen. Letztere vertrauen ebenso weniger auf langfristige Prognosen oder halten diese nur in sehr begrenztem Maße für möglich. Allerdings liegen derzeit noch wenige Erkenntnisse über die Mischungen von Bezugsnormen vorliegt (Jürgens & Lissmann, 2015). Vielmehr stellt Bohl (2009) die Hypothese auf, dass es in der alltäglichen Handlungspraxis von Lehrkräften vermehrt zu einer Mischung von sozialer und kriterialer Bezugsnorm kommt, was er mit der eher konservativen Struktur der Leistungsbeurteilung in Kontexten von Schulgesetzen und Kernlehrplänen in Deutschland begründet. Insbesondere stark rahmende curriculare Vorgaben (Lehrpläne, zentrale Vergleichs- und Abschlussarbeiten) sorgen vielfach dafür, dass Lehrkräfte Spannungen zwischen pädagogischen Überzeugungen und administrativen Vorgaben zugunsten einer kriterialen Normen wahrnehmen und folglich ein ‚teaching to the test‘ in der Unterrichtskultur fördern (ebd.).

Reflektiert man dies unter den Maßgaben einer inklusionsorientierten Perspektive auf schulische Leistung, so zeigt sich, dass das Heranziehen einer sozialen BnO für das pädagogische Handeln keinerlei Mehrwert bringt (Thurn, 2012; Jürgens & Lissmann, 2015). Vielmehr verhindert ein dauerhafter sozialer Vergleich Partizipation, Lernmotivation und Leistungsbereitschaft, insbesondere dann, wenn bestimmte Schüler\*innen häufiger als die vergleichsweise schwächeren adressiert werden (Thurn, 2012; Jürgens & Lissmann, 2015). So wird insbesondere im Kontext von Inklusion eine verstärkte Fokussierung von individuellen Entwicklungen und Stärken sowie eine didaktische Diagnostik entlang des Lernprozesses empfohlen (hierzu: Wilke, Knerndel & Schmidt, 2023;

Seitz & Wilke, 2021; Prengel, 2016). Die zahlreichen methodischen Überlegungen zur Begleitung und Evaluation von Lernprozessen (u.a. Winter, 2018; Beutel & Pant, 2020) zeigen sich an die Forderungen zu stärkeren individuellen Bezügen und Partizipation der Lernenden anschlussfähig, sind jedoch in inklusionspädagogischer Perspektive nicht vollständig evaluiert (Holder & Kessels, 2019). Zusätzlich erfordern derartige Methoden eine pädagogische Haltung, die befürwortend gegenüber neuen Methoden der Diagnose und Förderung ist (Schäfer und Rittmeyer, 2021). Bezieht man dieses zurück auf die Bezugsnormen, so stellt sich die Frage, wie sich eine BnO denken lässt, die den Ansprüchen einer inklusionsorientierten Didaktik und den vorherrschenden administrativen Vorgaben in Bezug auf schulische Inklusion entspricht und gerecht wird?

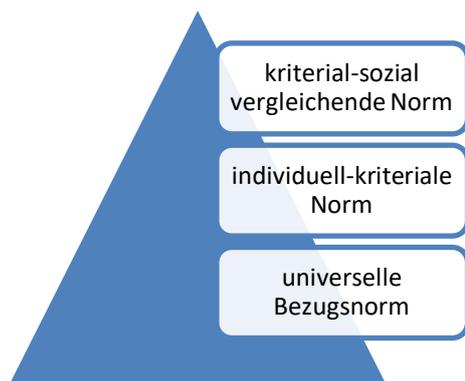


Abbildung 1: Darstellung des Bezugsnormmodells nach Prengel (2017). Eigene Darstellung.

Zur Beantwortung dieser Frage schlägt Annedore Prengel (2016) eine Weiterentwicklung der von Rheinberg (2001a) herausgearbeiteten Bezugsnormen vor, die sich enger als bisher an ein inklusionsorientiertes Leistungsverständnis anlehnt und vor diesem Hintergrund sowohl anerkennungs- und verfassungstheoretisch fundiert ist (Prengel, 2017, modelliert in Abb. 1). Hierfür formuliert Prengel (2016) die *universelle Bezugsnorm* als die Anerkennung jeder bzw. jedes einzelnen Lernenden als wertvolles und vor allem gleichwertiges Mitglied der Lerngruppe, vollkommen unabhängig

von ihren bzw. seinen (aktuellen) Leistungen. Diese universelle Bezugsnorm fußt dabei insbesondere auf menschenrechtlichen Argumentationen innerhalb des Inklusionsdiskurses (u.a. Seitz, 2020) und ist die Grundlage für das Heranziehen der zweiten Bezugsnorm, der *individuell-kriterialen Norm*. Dabei wird jedes Kind als „auf seiner Stufe kompetent“ (Prengel, 2016, S. 53) adressiert und in seinem Lernen und Leisten unter der Berücksichtigung seiner spezifischen Lernausgangslagen und Lernschritte gefördert. Vor dem Hintergrund dieser beiden Normen schließt sich in der Rahmung von Prengel die *kriterial-sozial vergleichende Norm* an. Diese dient dazu, den eigenen Lernstand mit anderen Lernenden zu vergleichen, nimmt dabei aber Rücksicht auf die vorgegangenen Normen und kann zu bestimmten Zeitpunkten dazu genutzt werden, den eigenen Lernstand mit den Vorgaben für das Abschließen einer Jahrgangsstufe, eines Kurses oder mit den Zielen eines Bildungsganges abzugleichen. Der Fokus der Betrachtung von Leistungen im Kontext einer inklusiven Didaktik liegt auf einer individuell-kriterialen Norm. Das heißt, dass der Fortschritt aller Lernenden auf jedem individuellen Niveau anerkannt und wertgeschätzt wird. Dies ist dabei keinesfalls willkürlich oder unfair, weil die Lernschritte verschiedener Schüler\*innen in ihrer Komplexität unterschiedlich sind – was in der Argumentation eine rein sozialen oder kriterialen Norm angenommen werden könnte –, sondern durch Transparenz und den fakultativen Vergleich zwischen den Schüler\*innen ist jederzeit auch eine soziale Vergleichbarkeit der Schüler\*innenleistung gegeben. Dieser soziale Vergleich erscheint weitgehend verzichtbar, da auch im allgemeindidaktischen Diskurs deutlich herausgestellt wird, dass der soziale Vergleich keinerlei neue Informationen hervorbringt und entsprechend weder für die Beurteilung noch für die Ausgestaltung des Unterrichts notwendig ist. Ferner ist er keine „positive Hilfe bei der Implementation des schulischen Erziehungs- und Bildungsauftrags“ (Jürgens, 2012, S. 35) und wird im Kontext inklusiver Schulentwicklung als Hürde beschrieben (Thurn, 2019). Im folgenden Abschnitt wird danach gefragt, in welcher Weise die verschiedenen

Bezugsnormen Einfluss auf die Leistungsbeurteilung in Deutschland und Italien nehmen.

## 2 Leistungsorientierung in Deutschland am Beispiel von NRW<sup>1</sup>

In Deutschland ist der Begriff der Leistungsbeurteilung im schulischen Kontext weiterhin geprägt von einer Diskussion über Noten und leistungsorientierte Zuteilung zu in verschiedener Weise qualifizierenden Sekundarschulformen und Förderschulen für verschiedene sogenannte sonderpädagogische Förderbedarfe. Vielfach gehen dabei Fragen der Objektivierung von Leistungsmessungen und Gerechtigkeit von Urteilen einher. Dabei ist längst klar, dass die Zuteilungen zu bestimmten Bildungsgängen nach der vierten Klasse kaum passgenau sind ihr prognostischer Wert gering ist und vielfach andere Variablen (z.B. sozio-kultureller und sozio-ökonomischer Hintergrund) Einfluss auf Beurteilungen und Empfehlungen nehmen, weshalb sie kaum als (leistungs-)gerechte Zuteilungen verstanden werden können (u.a. Steinmayr et al., 2017, Thurn 2017). Ebenso ist die Ineffizienz der bestehenden Förderschulen hinlänglich erforscht (Schnell, Sander & Federolf, 2011). Entsprechende Entwicklungen auf bildungspolitischer Ebene bleiben jedoch aus und inklusive Systeme werden in NRW zum Teil aktiv zurückgebaut (u.a. Seitz, 2020; Seitz et al., 2020). Entsprechend erklären sich auch Beschlüsse einiger politischer Parteien, die in breiten Linien auf alltagstheoretische Konzepte und Argumentationen zurückgreifen, welche bereits seit mehreren Jahrzehnten empirisch widerlegt sind (u.a. Ingenkamp, 1989; Ingenkamp & Lüdes, 2001; Lübke, 1996; Ziegenspeck, 1979):

„Eine Abkehr vom Leistungsprinzip – etwa durch das Abschaffen von Schulnoten – halten wir als Freie Demokraten für einen Schritt in die völlig falsche Richtung. Wer Schulnoten abschaffen will, nimmt Schülerinnen und Schülern Erfolgserlebnisse, Leistungsanreize und Möglichkeiten zur Selbsteinschätzung. Noten erleichtern es auch den Eltern und dem Umfeld, Kinder auf dem individuell passenden Bildungsweg gezielt zu unterstützen.“ (Präsidium FDP, 2023, S. 3)

In der gesellschaftlichen Diskussion um den schulischen Umgang mit Leistungen erscheinen die Fronten rund um die Beurteilung verhärtet. Betrachtet man die Pole des Diskurses fallen zwei Positionen deutlich auf: Gesellschaftliche und politische Positionen deuten dabei in Richtung einer Stärkung der Leistungsorientierung, die weiterhin an Förderbeschulung und einem stratifizierten Schulsystem mit Haupt-, Realschulen und Gymnasien festhalten und dabei insbesondere die Selektion nach Leistungen und Schulnoten beibehalten möchten (u.a. Präsidium FDP, 2023). Dem entgegen stehen Positionen, die ein nicht vertikal gegliedertes Schulsystem wie beispielsweise in Italien fordern, die auf weniger summative und selektierende Verfahren in der Diagnostik setzen und Beurteilungen durch Ziffernnoten ablehnen. In dieser Argumentationslinie steht eine stärkenorientierte, prozessbezogene und partizipative Ausrichtung schulischer Leistungsbeurteilung im Fokus. Entsprechend finden sich widersprüchliche administrative Vorgaben und Ziele, was dazu führt, dass der Ausbau inklusiver Schulkulturen in Bezug auf den Umgang mit schulischen Leistungen in Deutschland nur schleppend vorankommt (u.a. Seitz et al., 2020). Daher verwundert es kaum, dass zwar die positiven Effekte eines Verzichts auf Ziffernnoten (u.a. Winter, 2018) ebenso wie die Schwächen der

---

<sup>1</sup> Die Bildungslandschaft in Deutschland fußt auf eine vierjährige Grundschulzeit. Anschließend ist das Bildungssystem vertikal nach (prognostizierter) Leistungsfähigkeit strukturiert. Die unterschiedlichen Schulen des Sekundarbereichs qualifizieren anschließend zu unterschiedlichen Schulabschlüssen. In diesem Kontext sind bildungspolitische Entscheidungen in Deutschland jedoch nicht einheitlich auf Bundesebene geregelt, sondern obliegen der Administration der einzelnen Bundesländer (beispielsweise NRW), was zu nicht einheitlichen Ausdifferenzierung von Regel- und Förderschulen weit über die Dreigliedrigkeit hinaus führt und Gesamt- und Sekundarschulen lediglich als Parallelstruktur vorhanden sind.

Beurteilung mittels Ziffernoten (zusammenfassend: Ingenkamp & Lissmann, 2008) hinlänglich bekannt sind, jedoch neben den ersten zwei Jahren der Grundschule – die in der gesamten Bundesrepublik notenfrei sind – nur wenige Schulformen auf Noten als zentrales Element der Bewertung und Rückmeldung von schulischen Leistungen verzichten. Dabei ist jedoch anzumerken, dass alle Schulen – insbesondere im Sekundarbereich – spätestens mit dem Ende der achten Jahrgangsstufe Ziffernoten einführen.

Wirft man einen Blick in die Ausbildungs- und Prüfungsordnungen der verschiedenen Schulformen, wird vielfach der Fokus auf die Ergebnisse summativer Überprüfungen gelegt (Schulgesetz NRW § 48–50). In diesem Zuge finden sich in verschiedenen Paragraphen Kriterien, die im Kern auf die Beurteilung mit Ziffernoten verweisen und rein sachliche Kriterien (kriteriale Bezugsnorm) für die Beurteilung von Leistungen liefern, weshalb hier individuelle und soziale Bezugsnormen ausgespart werden. Vertieft man hier den Blick auf die verschiedenen Schulfächer, so finden sich unterschiedliche Operationalisierungen von Leistung.

So heißt es in den Lernplänen für die Grundschule im Land Nordrhein-Westfalen:

„Die Primarstufe ist einem pädagogischen Leistungsverständnis verpflichtet, das Leistungsanforderungen mit individueller Förderung verbindet. Für den Unterricht bedeutet dies, Leistungen nicht nur zu fordern, sondern sie auch zu ermöglichen und zu fördern. Deshalb geht der Unterricht von den individuellen Voraussetzungen der Kinder aus und leitet sie dazu an, ihre Leistungsfähigkeit zu erproben und weiterzuentwickeln. Grundlage hierfür ist die Ermittlung der Lernausgangslage.“ (Lehrplan Grundschule, 2021, S. 195)

In diesem Zitat deutet sich neben der rein kriterialen BnO ebenso eine individuelle BnO an, die nach den Lernausgangslagen der Schüler\*innen fragt und postuliert, dass Unterricht an diesen entlang geplant wird, um zu ermöglichen, dass Schüler\*innen ihre eigenen Fähig- und Fertigkeiten weiterentwickeln. Im Weiteren wird auf die verschiedenen Dimensionen von Leistung Bezug genommen, sodass entgegen der rein kriterial gestalteten Passagen aus dem Schulgesetz hier individuelle, gruppenbezogene und kriteriale Ziele verbunden werden und Leistung als ‚mehr‘ verstanden wird als das dargelegte Wissen und die unter Beweis gestellten Kompetenzen in Prüfungssituationen (Lehrplan Grundschule).

Einen anderen Fokus setzen die Lehrpläne in der Sekundarstufe: Zwar formulieren sie die Berücksichtigung der individuellen Förderung und Transparenz (Lehrplan Mathematik Sek. I), orientieren sich aber deutlich stärker an den Kriterien der Sekundarstufe.

„Die Leistungsbewertung ist so anzulegen, dass sie den in den Fachkonferenzen gemäß Schulgesetz (§ 70 Abs. 4 SchulG) beschlossenen Grundsätzen entspricht, dass die Kriterien für die Notengebung den Schülerinnen und Schülern transparent sind und die Korrekturen sowie die Kommentierungen den Lernenden auch Erkenntnisse über die individuelle Lernentwicklung ermöglichen.“

In dieser Perspektive fokussieren Bewertungen auf Inhaltsbereiche und Kompetenzen, die als Teil der summativen Überprüfung – auch in Form zentraler Prüfungen – in Form von Noten bewertet werden. Somit dominiert in der Sekundarstufe auf einer formalrechtlichen Ebene ein starker kriterialer Bezugsrahmen. Hierzu finden sich zahlreiche kritische Positionen im deutschsprachigen Raum (u.a. Thurn, 2019; Jürgens, 2012), welche jedoch nur in Ausnahmefällen Einzug in die Ausgestaltung der Leistungsbeurteilung im allgemeinbildenden Schulsystem finden. Als (fast) notenfreie Schulen verbleiben damit bezogen auf Nordrhein-Westfalen nur die PRIMUS-Schulen (PRIMUS NRW, o.J.) und die Laborschule Bielefeld sowie einige Schulen in privater Trägerschaft. In anderen Bundesländern gilt diesbezüglich andere Regelungen (beispielsweise an niedersächsischen Gesamtschulen), in denen eine Beurteilung in der Eingangsstufe auch auf Konferenzbeschluss andere Formen der Leistungsrückmeldung möglich sind.

### 3 Orientierungen und Entwicklungen in Norditalien

Im Gegensatz zu Deutschland hat Italien seit den 1970er Jahren sein Schulsystem konsequent umstrukturiert. In diesem Zuge wurde die vertikale Strukturierung der Schulen im Primar- und Sekundarbereich abgebaut und die Zeit des gemeinsamen Lernens bis zum Ende der achten Jahrgangsstufe etabliert. Eine gesonderte Beschulung von Schüler\*innen mit sogenannten sonderpädagogischen Unterstützungsbedarfen findet nicht statt. Weiter wurde von 1977 bis 2008 auf die Beurteilung mit Ziffernnoten verzichtet. Die Steuerung des Schulsystems und bildungspolitische Entscheidungen sind maßgeblich vom Ministerium für Unterricht, Universitäten und Forschung gesteuert und somit für Italien in der Breite einheitlich. Den einzelnen Provinzen obliegen daher weniger Spielräume in der Umsetzung als es in Deutschland der Fall ist. Eine Ausnahme bilden hierbei autonome Provinzen.

In diesem Abschnitt wird die jüngste Änderung des Bewertungssystems an italienischen Grundschulen erörtert. Wir werden uns dann kurz auf den spezifischen Fall der Region Trentino-Südtirol konzentrieren und versuchen, die wesentlichen Aspekte zu erfassen, die heute die Bewertungsmethoden der autonomen Provinz Bozen von denen in Trient unterscheiden.

Während der durch die COVID-19-Pandemie bedingten Zäsur war es für italienische Lehrkräfte besonders schwierig, die Aktivitäten zu bewerten, die ihre Schüler\*innen im Rahmen des Fernunterrichts durchführen sollten. Diese Schwierigkeit ist zum Teil darauf zurückzuführen, dass im Jahr 2008 numerische Noten wieder eingeführt wurden, die zuvor 1977 abgeschafft worden waren (Piscozzo & Stefenel, 2022). Dies führte im Dezember 2020 zur Formulierung der Ministerialverordnung 172 (O.M. 172), einem Dokument, das den formativen Charakter der Beurteilung als ein Element bekräftigen soll, das es den Schüler\*innen ermöglichen soll, ihre Stärken zu verstehen und ihre persönliche Identität zu entwickeln.

Eine der Hauptveränderungen, die den Übergang von der früheren Beurteilungsmethode zu der durch den O.M. 172 eingeführten Methode kennzeichnet, ist der Übergang von numerischen Noten zu so genannten „beschreibenden Beurteilungen“ (*giudizi descrittivi*); die im deutschsprachigen Raum als Wortgutachten beschrieben werden. Um besser verstehen zu können, wie solche Beurteilungen von Lehrer\*innen vorgenommen werden sollten, ist es notwendig, auf einige grundlegende Aspekte des Wandels einzugehen, auf den wir uns beziehen. Die Absicht der Gruppe von Forscher\*innen, Grundschullehrer\*innen, Schulleiter\*innen und Universitätsprofessor\*innen, die den Leitfaden zur Veröffentlichung des Ministerialerlasses 172 verfasst haben, bestand in erster Linie darin, den Übergang zu einer kriteriengestützten Bewertungsperspektive zu unterstützen. Der bisherige Bewertungsmodus maß die individuelle Leistung der Schüler\*innen an der Durchschnittsnote der Klasse. Diese Art der Bewertung wird im Italienischen als normative Bewertung (*valutazione normativa*)<sup>2</sup> bezeichnet. Im Gegensatz dazu suchte das Forschungsteam, das die Leitlinien erstellte, nach einer Möglichkeit, jedem einzelnen Schüler bzw. jeder einzelnen Schülerin deutlich zu machen, wo er oder sie auf dem Weg zur Erreichung der gesetzten Ziele steht (Agrusti, 2021A). Diese Art der Bewertung, die die Leistung der Schüler\*innen mit Kriterien vergleicht, die zu Beginn des Lernprozesses festgelegt werden, wird als kriterienbasiert (*valutazione criteriale*) bezeichnet<sup>3</sup>. Dementsprechend wurde beschlossen, die Erstellung der beschreibenden Beurteilungen auf zwei Hauptelemente zu stützen: zum einen auf die Formulierung von Lernzielen, die als Grundlage für die Lehrtätigkeit dienen, und zum anderen auf die Verwendung von vier verschiedenen Lernniveaus, um die Qualität des Lernfortschritts eines jeden Schülers bzw. einer jeden Schülerin zu ermitteln.

<sup>2</sup> Diese ist vergleichbar mit der sozialen Norm, wie sie von Bohl (2009), Rheinberg (2001A) oder Prengel (2017) beschrieben wird.

<sup>3</sup> Vergleichbar mit der kriterialen Norm (Agrusti, 2021A)

Konzentrieren wir uns zunächst auf den ersten Aspekt: die Definition der Lernziele. Deren Wortlaut bezieht sich auf ein anderes wichtiges ministerielles Dokument, die so genannten nationalen Angaben (*Indicazioni Nazionali*). Artikel 34 der italienischen Verfassung besagt, dass der Unterricht in den Künsten und den Wissenschaften frei ist. Es gibt also keinen Lehrplan, den die Lehrer in einem bestimmten Zeitraum unterrichten müssen, sondern nur Hinweise auf die vom Ministerium vorgegebenen Ziele. In dem 2012 veröffentlichten Dokument *Indicazioni Nazionali* ist von Zielen (*obiettivi*) und Zwecken (*traguardi*) die Rede, die die wünschenswerten Ergebnisse der verschiedenen Bildungsstufen angeben (Trincherò, 2021). Das Erreichen der Lernziele wird dann zum Gegenstand der Bewertung (Agrusti, 2021B). Um den Lehrkräften die Beobachtung der Schüleraktivitäten zu erleichtern, müssen diese Lernziele so operationalisiert werden, dass nicht nur der Inhalt, den die Schüler lernen sollen, sondern auch die Art des Handelns, durch das sich dieses Lernen manifestiert, deutlich wird. Die Geografie eines bestimmten Ortes zu verstehen, kann vieles bedeuten: Es ist eine Sache, eine einfache Liste von Flüssen und Bergen auswendig zu wiederholen, eine ganz andere, zu erklären, wie die wirtschaftliche Entwicklung durch die orografische Beschaffenheit begünstigt wurde oder nicht. Die Lehrkraft ist daher aufgefordert, sich zu fragen, durch welche Art von Aktivität sich ein bestimmter Lerninhalt manifestieren kann (Trincherò, 2021). Auflisten, Erklären, Annahmen treffen, eine bestimmte Veränderung vorhersagen, etwas planen – all dies sind Tätigkeiten, die ein unterschiedliches Maß an Beherrschung dessen erfordern, was der/die Schüler\*in erworben hat, und daher zielt das neue italienische Bewertungssystem darauf ab, diese Unterschiede in die Bewertung einzubeziehen (Agrusti, 2021B).

Sobald feststeht, was bewertet werden soll – d. h. die Lernziele operationalisiert –, ist es auch wichtig zu verstehen, auf welche Weise ein bestimmtes Ziel erreicht werden kann. Dies wird durch zwei verschiedene Gruppen von Indikatoren angezeigt: einerseits die vier Niveaus des Lernens (*livelli di apprendimento*), andererseits die so genannten vier Dimensionen (*dimensioni*) des Lernens (Agrusti, 2021A). Die Niveaus geben nur an, inwieweit ein Lernziel erreicht worden ist. Sie können wie folgt übersetzt werden: Anfang (*in via di prima acquisizione*), Basis (*base*), Mittelstufe (*intermedio*) und Fortgeschrittene (*avanzato*). Im Gegensatz zu früher sieht der kriteriengestützte Bewertungsansatz vor, dass die Vergabe dieser Niveaus nicht allein von der Entscheidung des Lehrkräfte abhängt, sondern von der Anerkennung der oben genannten vier Dimensionen (Girelli, 2022, Tabelle 1).

Tabelle 1: Lernziele und ihre Niveaustufen				
	<i>Autonomie</i>	<i>Kontinuität</i>	<i>Art der Situation</i>	<i>Eingesetzte Mittel</i>
<b>Fortgeschritten</b>	Autonom	Kontinuierliche Zielerreichung	Bekannte und unbekannte Situationen	Von der Lehrkraft bereitgestellte und spontan gefundene Mittel
<b>Mittelstufe</b>	Autonom, wenn die Situation bekannt ist	Kontinuierliche Zielerreichung, wenn die Situation bekannt ist	Bekannte und unbekannte Situationen	Nur von der Lehrkraft bereitgestellte Mittel

<b>Basis</b>	Begrenzt autonom, wenn die Situation bekannt ist	Diskontinuierliche Zielerreichung	Bekannte Situation	Nur von der Lehrkraft bereitgestellte Mittel
<b>Anfang</b>	Nicht autonom	Diskontinuierliche Zielerreichung	Bekannte Situation	Nur von der Lehrkraft bereitgestellte Mittel

Wie die obige Tabelle zeigt, sollten die Niveaus und Lernziele entsprechend der tatsächlichen Manifestation der vier Dimensionen verstanden werden. Für jedes Ziel ist es wichtig, dass der Grad der Autonomie bestimmt wird, mit dem es erreicht wird (*autonomia*); die Tatsache anerkannt wird, dass der Grad Beherrschung der Lernschritte sich nicht isoliert manifestiert (*continuità*); dass die Art der Aufgabe oder Situation dem/der Schüler\*in nicht schon vorher bekannt war (*tipologia di situazione*) und dass der Grad zu ermitteln ist, zu dem der/die Schüler\*in der Lage ist, spontan alle verfügbaren Ressourcen zur Lösung einer Aufgabe zu nutzen (*risorse mobilitate*).

Die Lehrkräfte müssen daher alle diese Elemente berücksichtigen, um zu verstehen, wie ein\*e Schüler\*in ein Lernziel erreicht hat (Agrusti, 2021A). Dies führt uns einen weiteren wesentlichen Aspekt von O.M.172 vor Augen, nämlich die Tatsache, dass sie für die Durchführung von Bewertungen über einen großen Zeitraum konzipiert ist: das erste Halbjahr und das Ende des Schuljahres (Castoldi, 2021). Die Ministerialdokumente enthalten dagegen keine besonderen Hinweise darauf, wie die *in itinere* (prozessbezogene) Bewertung für jede einzelne Aufgabe oder Tätigkeit durchgeführt werden soll (Nigris & Balconi, 2021; Piscozzo & Stefanelli, 2022). Dieser Aspekt, der zum Teil auch mit der bereits erwähnten Freiheit des Unterrichts zusammenhängt, birgt die Gefahr, dass falsche Vorstellungen entstehen und die Lehrer beispielsweise Bewertungsmethoden anwenden, die der formativen Perspektive von O.M.172 nicht gerecht werden. Wenn wir zum Beispiel an die Dimension der Kontinuität denken, wird schnell deutlich, dass sie nicht in einer einzigen Ausprägung zu finden ist, sondern dass es notwendig ist, die Prozesse und nicht die Produkte des Lernens zu betrachten.

Ein letzter Aspekt, der hervorgehoben werden muss, ist die Aufwertung der Selbsteinschätzung als wesentliches Moment in der Bildungsentwicklung der Schüler\*innen. Dieser Aspekt wird insbesondere in den Leitlinien der O.M.172 aufgegriffen, die zum ersten Mal in der Geschichte der italienischen Schulen vorschlagen, die Selbsteinschätzung der Schüler\*innen als wichtiges Element in das Jahresendzeugnis aufzunehmen (Sorgato, 2021). Vor diesem Hintergrund muss auch der prägende Charakter von O.M.172 verstanden werden.

In Artikel 7 des Ministerialdokuments heißt es, dass es den autonomen Provinzen Trient und Bozen freisteht, die Bestimmungen des Ministerialerlasses so anzuwenden, wie es in ihren jeweiligen Provinzgesetzen und Durchführungsbestimmungen vorgesehen ist. Während die Provinz Bozen die erwähnten Änderungen bereits ein Jahr nach dem Erlass der Verordnung vorgenommen hat, ist die Autonome Provinz Trient bis heute die einzige in ganz Italien, die ein völlig anderes Bewertungssystem anwendet, das auf rein summativen Methoden beruht. Im Jahr 2009 hatte die Provinz Trient ihre eigenen Richtlinien für die Schüler\*innenbeurteilung erlassen, und obwohl in jenen Jahren innovative Änderungen wie Selbst- und Fremdeinschätzung gefördert wurden, basierte die Beurteilungsmethode immer auf der Verwendung von Noten, wenn auch durch ein Wort und nicht durch eine Zahl ausgedrückt. Heute wird ein\*e Schüler\*in, der/die nur wenige Kilometer von der Provinz Bozen, der Lombardei oder Venetien in die Provinz Trient zieht, in einem institutionellen Rahmen beurteilt, der durch die Verwendung von Noten den Vergleich mit Gleichaltrigen und die Bildung von Durchschnittswerten fördert, ohne

den Schüler\*innen die Mittel an die Hand zu geben, über ihren eigenen Entwicklungs- und Leistungsweg nachzudenken.

#### 4 Leistung inklusive – Potenziale und Entwicklungsfelder

Vergleicht man die Leitlinien zur Leistungsbeurteilung in den ausgewählten Provinzen Italiens und Nordrhein-Westfalen fällt auf, dass deutlich unterschiedliche Beurteilungsformen und Definitionen von schulischen Leistungen prägend sind. So fokussieren die ausgewählten italienischen Provinzen ebenso wie die Grundschulen in Nordrhein-Westfalen auf entwicklungs- und prozessorientierte Formen der Beurteilung, die sich deutlich an einer individuellen Bezugsnorm orientieren, jedoch in Kombination mit kriterialen Ansprüchen an Leistungen stehen. Dies lässt sich in Bezug auf die von Prenzel (2016) vorgestellten Bezugsnormen auf der individuell-kriterialen Ebene verorten und scheint somit – eine menschenrechtlich fundierte Anerkennung aller Lernenden als wertvoller Teil der Lerngruppen und Gesellschaft vorausgesetzt – den Ansprüchen an eine inklusionsorientierte Leistungsbeurteilung gerecht zu werden.<sup>4</sup> Kritisch erscheint hier jedoch weiterhin die Rückmeldung der Ergebnisse der Leistungsüberprüfung mittels Ziffernoten, welche einen sozialen Vergleich erbrachter Leistungen nahelegen und unterstützen, obgleich dieser insbesondere in Bezug auf stark heterogene Lerngruppen als kritisch beurteilt werden kann. Umso mehr, da bekannt ist, dass Schüler\*innen ihr eigenes Selbstwertgefühl zum Teil an erhaltene Rückmeldungen und Noten knüpfen (Thurn, 2012; Steiner, 1981).

Mit Blick auf die Sekundarstufe in Deutschland zeigt sich die Dominanz einer stark kriterial geprägten Leistungsbeurteilung, die im Kern auf Standardisierung und das Erreichen von Vorgaben abzielt und zudem noch in normierten Abschlussprüfungen mündet. Zu diesem Verständnis von Leistung trägt die Aufteilung der Schüler\*innen auf verschiedene Schulformen im Sekundarbereich bei, die einerseits als langfristige Leistungsprognosen verstanden werden und zum anderen entlang einer sozialen Norm erfolgen. In diesem Bereich zeigt sich ein Einheitsschulsystem wie das Italienische deutlich anschlussfähiger an inklusionspädagogische Ansprüche als Nordrhein-Westfalen.

Entsprechend dieses Vergleiches zeigt sich die deutlich günstigere Ausgangslage für inklusionsorientierte Leistungsbeurteilung in Italien, da hier seit den 1970er Jahren ein ungleich breiterer Erfahrungsschatz im Umgang mit (Leistungs-)Heterogenität entwickelt werden konnte, als dies in Deutschland der Fall ist. Zwar können Lehrkräfte in Deutschland auf zahlreiche Konzepte, Methoden und Begründungslinien für einen „neuen“ Umgang mit schulischen Leistungen zurückgreifen (u.a. Seitz & Wilke, 2021) – was im aktuellen öffentlichen Diskurs häufig nicht berücksichtigt wird. Diese sind jedoch aktuell nicht unmittelbar anschlussfähig an regulative Rahmungen, Weisungen, Lehrpläne und das Schulgesetz sowie deren Rekontextualisierung im schulischen Kontext. Vielmehr scheinen sich widerständige Einzellehrkräfte in Formen funktioneller Regelverstöße einen inklusionspädagogischen Umgang mit der Leistungsbeurteilung zu erschließen, obgleich regulative Maßnahmen einen solchen erschweren (Wilke & Reh, 2023). Im Sinne der Schüler\*innen ist hier auf bildungspolitische Novellierungen zu hoffen, die sich in Bezug auf die Sekundarstufe an Italien orientieren könnten. Ein vertiefter bildungspolitischer sowie bildungswissenschaftlicher Austausch wünschenswert wäre.

---

<sup>4</sup> An dieser Stelle bleiben jedoch bezogen auf Nordrhein-Westfalen die Wirkungen verschiedener Erlasse, die als ein aktiver Rückbau in der inklusionsorientierten Entwicklung der Schulkulturen zu sehen sind (u.a. Seitz et al, 2020), fraglich.

## 5 Literatur

- Agrusti, G. (2021A). Approcci criteriali alla valutazione nella scuola primaria. *RicercaAzione*, 13 (1), 25–38. <https://doi.org/10.32076/RA13101>
- Agrusti, G. (2021b). Per un ritorno agli obiettivi. Come cambia la valutazione nella scuola primaria. *Cadmo* (1), S. 5–20.
- Beutel, S.-I., & Pant, H. A. (2020). *Lernen ohne Noten. Alternative Konzepte der Leistungsbeurteilung*. Stuttgart: Kohlhammer.
- Bohl, T. (2009). *Prüfen und Bewerten im offenen Unterricht*. Weinheim und Basel: Beltz.
- Castoldi, M. (2021). *Valutare gli apprendimenti nella scuola primaria*. Mondadori Università.
- Girelli, C. (2022). *Valutare nella Scuola Primaria: Dal Voto al Giudizio Descrittivo*. Carocci.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Holder, K., & Kessels, U. (2019). Unterrichtsgestaltung und Leistungsbeurteilung im inklusiven und standardorientierten Unterricht aus der Sicht von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 22 (2), S. 325–346.
- Ingenkamp, K. (1989). *Die Fragwürdigkeit der Zensurengebung*. Weinheim.
- Ingenkamp K., & Lissmann U. (2008). *Lehrbuch der Pädagogischen Diagnostik*. Weinheim und Basel: Beltz.
- Ingenkamp, K. & Lüders M. (2001). Dispositionsspielräume im Bereich der Schülerbeurteilung. *Zeitschrift für Pädagogik*, 47 (2), S. 217–234.
- Jürgens, E. (2012). Diagnosefunktion: Fehlanzeige. Ist eine Neuorientierung der Beurteilungspraxis nötig? In C. Fischer (Hrsg.), *Diagnose und Förderung statt Notengebung? Problemfelder schulischer Leistungsbeurteilung* (S. 24–44). Münster u.a.: Waxmann.
- Jürgens, E. (2019). Leistungsbewertung. In M. Harring, C Rohlfs & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 505–515). Münster und New York: Waxmann.
- Jürgens, E. & Lissmann, U. (2015). *Pädagogische Diagnostik*. Reihe „BildungsWissen Lehramt“, 27. Weinheim [u.a.]: Beltz.
- Lübke, S.-I. (1996). *Schule ohne Noten. Lernberichte in der Praxis der Laborschule*. Opladen.
- Ministerium des Inneren des Landes Nordrhein-Westfalen (2022). *Schulgesetz für das Land Nordrhein-Westfalen (Schulgesetz NRW – SchulG)*. Zugriff am 24.10.2023 Verfügbar unter: [https://recht.nrw.de/lmi/owa/br\\_bes\\_text?sg=0&menu=0&bes\\_id=7345&aufgehoben=N&anw\\_nr=2](https://recht.nrw.de/lmi/owa/br_bes_text?sg=0&menu=0&bes_id=7345&aufgehoben=N&anw_nr=2)
- Nigris, E., & Balconi, B. (2021). La Valutazione in itinere: Dalla Documentazione alla Valutazione Descrittiva. In E. Nigris & G. Agrusti (a cura di), *Valutare per apprendere. La nuova valutazione descrittiva nella scuola primaria*. Pearson Italia.
- Piscozzo, M., & Stefanel, S. (2022). *La valutazione nella scuola primaria. Obiettivi, curricula, scelte*. UTET.
- Präsidium der FDP (2023). *Leistungsprinzip und Chancengerechtigkeit stärken: Für ein modernes und leistungsorientiertes Bildungssystem*. Zugriff am 24.10.2023 Verfügbar unter: <https://www.fdp.de/beschluss/beschluss-des-praesidiums-leistungsprinzip-und-chancengerechtigkeit-staerken-fuer-ein>
- Prenzel, A. (2016). Didaktische Diagnostik als Element alltäglicher Lehrerverarbeit – „Formatives Assessment“ im inklusiven Unterricht. In B. Amrhein (Hrsg.), *Diagnostik im Kontext inklusiver Bildung – Theorien, Ambivalenzen, Akteure, Konzepte* (S. 49–63). Bad Heilbrunn: Klinkhardt.
- Prenzel, A. (2017). Individualisierung in der „Caring Community“ – Zur inklusiven Verbesserung von Lernleistungen. In A. Textor, S. Grüter, I. Schiermeyer-Reichl &

- B. Streesse (Hrsg.), *Leistung inklusive? Inklusion in der Leistungsgesellschaft. Band II: Unterricht, Leistungsbewertung und Schulentwicklung*. Bad Heilbrunn: Klinkhardt.
- PRIMUS NRW (o. J.). *Leistungsbewertung*. <https://www.primus-muenster.de/leistungsbewertung>
- QUALIS NRW (o. J.). *Lehrpläne Mathematik in der Sekundarstufe I*. Zugriff am 24.10.2023 Verfügbar unter: <https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan-navigator-s-i/>
- QUALIS NRW (o. J.). *Lehrplan Grundschule: Sachunterricht*. Zugriff am 24.10.2023 Verfügbar unter: [https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp\\_gs/LP\\_GS\\_2008.pdf](https://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_gs/LP_GS_2008.pdf)
- Rheinberg, F. (1977). *Soziale und individuelle Bezugsnorm*. Unveröffentlichte Dissertation. Bochum: Psychologisches Institut der Ruhr-Universität.
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation*. Göttingen. Hogrefe.
- Rheinberg, F. (Hrsg.) (1982). *Bezugsnormen zur Schulleistungsbewertung: Analyse und Intervention*. Düsseldorf: Schwann.
- Rheinberg, F. (2001a). Bezugsnormorientierung. In D.H. Rost (Hrsg.), *Handwörterbuch der Pädagogischen Psychologie*. 2. Auflage (S. 55–61). Weinheim/Basel: Beltz.
- Rheinberg, F. (2001b). Bezugsnormen und schulische Leistungsbeurteilung. In F.E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 59–71). Weinheim/ Basel: Beltz.
- Schäfer, H., & Rittmeyer, C. (2021). *Handbuch Inklusive Diagnostik*. Beltz Verlag.
- Schnell, I., Sander, A., & Federolf, C. (Hrsg.) (2011). *Zur Effizienz von Schulen für Lernbehinderte. Forschungsergebnisse aus vier Jahrzehnten*. Bad Heilbrunn: Klinkhardt.
- Seitz, S. (2020). Dimensionen inklusiver Didaktik – Personalität, Sozialität und Komplexität. *Zeitschrift für Inklusion* (2). Zugriff am 24.10.2023 Verfügbar unter: <https://www.inklusion-online.net/index.php/inklusion-online/article/view/570>
- Seitz, S., Hamisch, K., Kaiser, M., Slodczyk, N., & Wilke, Y. (2020). Inklusiv Schul-kulturen und widersprüchlichen Vorzeichen In T. Dietze, D. Gloystein, V. Moser, A. Piezunka, L. Röbenack, L. Schäfer, G. Wachtel & M. Walm (Hrsg.), *Inklusion – Partizipation – Menschenrechte: Transformationen in die Teilhabegesellschaft?* (S. 251–258). Bad Heilbrunn: Klinkhardt.
- Seitz, S., & Wilke, Y. (2021). „Dann hab‘ ich das einfach gemacht“ – Leistungsbeurteilung im inklusiven Unterricht der Sekundarstufe I. *Schule inklusiv 2021* (11), S. 35–36.
- Steiner, S. (1981). Das Leistungsprinzip als Glaubensbekenntnis. *Die Sozialistische Erziehung. Monatsschrift für die Bildungs- und Kulturarbeit der sozialistischen Bewegung Österreichs*, S. 99–102.
- Steinmayr R., Michels J., & Weidinger, A. F. (2017). *FA(IR)BULOUS – FAIRe BeUrteilung des LeistungspOtenzials von Schülerinnen Und Schülern*. Zugriff am 24.10.2023 Verfügbar unter: <https://www.stiftung-mercator.de/de/publikationen/studie-fairbulous/>
- Sorgato, S. (2021). L'autovalutazione al centro della valutazione formativa. In E. Nigris & G. Agrusti (a cura di), *Valutare per apprendere. La nuova valutazione descrittiva nella scuola primaria*. Pearson Italia.
- Thurn, S. (2012). Die Bielefelder Laborschule – Leistung in einer Kultur der Vielfalt oder: „Die Würde des heranwachsenden Menschen macht aus, sein eigener ‚Standard‘ sein zu dürfen.“ In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Leistungsbeurteilung* (S. 137–152). Wiesbaden: Springer.
- Thurn, S. (2017). Leistungsbewertung und Vielfalt – Oder: Umgang mit den Widersprüchen des Systems. *Pädagogik* 69, (9), S. 6–9.

- Thurn, S. (2019). Inklusives Schulsystem. In M. Harring, C. Rohlfis & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 103–115). Münster: Waxmann.
- Trincherò, R. (2021). La formulazione degli obiettivi di apprendimento. In E. Nigris & G. Agrusti (a cura di), *Valutare per apprendere. La nuova valutazione descrittiva nella scuola primaria*. Pearson Italia.
- Wilke, Y., Knerndel, H., & Schmidt, T.L. (2023). Lernprozessbegleitung partizipativ gestalten. Konzeptionen der Laborschule und mögliche Entwicklungsfelder Literaturangabe ergänzen
- Wilke, Y., & Reh, A. (2023). Ungenutzte Innovationspotentiale inklusiver Unterrichts- und Schulentwicklung: Mechanismen der Abgrenzung und Regression von Innovationen durch funktionelle Regelverstöße. *Empirische Pädagogik* 37 (3), S. 334–349.
- Winter, F. (2018). *Lerndialog statt Noten. Neue Formen der Leistungsbewertung*. Weinheim [u.a.]: Beltz.
- Ziegenspeck, J. (1979). Zensuren und Zeugnis – Ein Mängelbericht. In: D. Bolscho & C. Schwarzer (Hrsg.). *Beurteilen in der Grundschule* (S. 36–53). München: Urban & Schwarzenberg.