

Stellungnahme zur Bürgerinitiative 25/BI betreffend „Die Ermöglichung der alternativen Leistungsbeurteilung ohne Noten im Rahmen der Schulautonomie“

In ihrem Schreiben vom 1. Juli 2020 (Zl. 25/BI-NR/2020) hat die Parlamentsdirektion die Bitte des Ausschusses für Petitionen und Bürgerinitiativen um eine Stellungnahme zum Gegenstand der Bürgerinitiative „Die Ermöglichung der alternativen Leistungsbeurteilung ohne Noten im Rahmen der Schulautonomie“ weitergegeben, welcher wir hiermit nachkommen.

Die vorliegende Stellungnahme bezieht sich auf Ziffernnoten im Kontext schulischer Leistungsbeurteilung und bezieht sich dabei auf einschlägige Forschungsergebnisse zu diesem Thema. Noten sind in der Schule verknüpft mit der Bewertung von Einzelleistungen von Schülerinnen und Schülern (Schularbeiten, Tests) sowie mit den Schulnachrichten am Ende des ersten Semesters und den Jahreszeugnissen. Sie sollen Schülerinnen und Schülern bescheinigen, wie ihr Leistungsstand in den verschiedenen Fächern ist.

Bei der Beurteilung von Ziffernnoten ist zuerst zu betrachten, wofür sie dienen sollen. Noten werden in der Regel drei große Funktionen zugeschrieben, eine Rückmeldefunktion, eine Berichtsfunktion und eine Selektionsfunktion. Im Rahmen der Rückmeldefunktion sollen Noten der Förderung des Lernens dienen, indem die Lernenden Rückmeldung zu ihren Lernerfolgen erhalten und sie als Rückmeldungen eine Basis bilden, sich über die Leistungen zu verständigen. Die Berichtsfunktion dient der Information – insbesondere der Eltern – darüber, wo ein Schüler oder eine Schülerin in den verschiedenen Fächern steht. Dazu sollen Noten den Lernstand möglichst genau und gerecht beschreiben. Mit der Selektionsfunktion ist umfasst, wenn Leistungsbeurteilungen als Grundlage für die Berechtigung zum Aufsteigen in die nächste Schulstufe bzw. den Übertritt in das nächste Segment des Schulsystems oder die Zulassung zum Besuch einer bestimmten Schulart dienen. Darüber hinaus führt Winter (2015) als weitere Intentionen, die manchmal mit Noten verbunden sind, an, dass Noten Schülerinnen und Schüler motivieren sollen sich anzustrengen und nicht selten „Noten auch als Disziplinierungsmittel gebraucht [werden], obwohl dies eigentlich nicht korrekt ist“ (ebd., S. 60).

Noten sind Teil der komplexen und umfangreichen Bereiche Leistungsfeststellung und Leistungsbeurteilung. Im Rahmen von Aktivitäten der Leistungsfeststellung erfassen Lehrerinnen und Lehrer den Lernstand ihrer Schülerinnen und Schüler, entsprechend der LBVO (vgl. BGBl. Nr. 139/1974 i.d.g.F.) zum Beispiel anhand von unterrichtsbegleitenden Maßnahmen der Leistungsfeststellung oder mithilfe von punktuellen Leistungsfeststellungen wie mündlichen Prüfungen, Tests oder Schularbeiten. Erst in einem nächsten Schritt findet eine Leistungsbeurteilung statt, indem die Beobachtungen der Leistungsfeststellung interpretiert und bewertet und in eine Beurteilung überführt werden. Noten sind dann die von der LBVO für punktuellen Leistungsfeststellungen sowie als zusammenfassende Beurteilung im Zeugnis vorgesehenen Beurteilungsstufen, durch die das Ergebnis dieses Prozesses gefasst und kommuniziert wird. Auch wenn die vorliegende Stellungnahme explizit zur Ausformung der Beurteilung in

Form von Ziffernnoten (im Vergleich zu alternativen Formen der Leistungsbeurteilung) erbeten wurde, ist es erforderlich, einen Blick auf den dahinterliegenden größeren Prozess der Leistungsfeststellung als Grundlage für die Leistungsbeurteilung zu werfen.

Daraus ergeben sich zwei Analysestränge, die als Grundlage für die Einschätzung von Noten und alternativen Formen der Leistungsbeurteilung in Bezug auf die oben genannten Funktionen benötigt werden: Noten im Kontext des Gesamtprozesses der Leistungsfeststellung und -beurteilung; Leistungsbewertung im Kontext des Unterrichts und als solche ihre Wirkungen auf das Lehren und Lernen.

Lehrer*innen-Urteile als Grundlage der Leistungsbewertung

Zur Qualität von Ziffernnoten auf der Grundlage von Lehrer*innen-Urteilen liegen eine Reihe von empirischen Studien und zusammenfassenden Analysen vor. Sehr umfangreich haben Brügelmann et al. (2006) den Forschungsstand mit einem Fokus auf den Bereich der Grundschule aufgearbeitet. Spezifisch für die Situation in Österreich arbeitet der Beitrag von Eder, Neuweg und Thonhauser (2009) im nationalen Bildungsbericht relevante Studienergebnisse auf. Die wichtigsten Erkenntnisse werden im Folgenden zusammengefasst. Dies erfolgt zunächst entlang der wissenschaftlichen Gütekriterien Objektivität, Validität, Reliabilität (Hesse & Latzko, 2017; Ingenkamp & Lissmann, 2008) und wird ergänzt durch Hinweise zu den zugrundeliegenden Vergleichsmaßstäben der Beurteilung (Bezugsnormen).

Objektivität:

Mit Objektivität ist gemeint, dass Leistungsbeurteilungen unabhängig von der beurteilenden Person sind. Noten sind als Einschätzungen von Leistungen in hohem Ausmaß personenabhängig, weil sie auf Lehrer*innen-Urteilen beruhen. Zwar wird eine bewusste Empathie der Lehrperson als Vorteil in Bezug auf förderorientierte Rückmeldungen angesehen, in Bezug auf die Berichts- und insbesondere Selektionsfunktion der Leistungsbeurteilung ist die Beeinflussung durch sachfremde Informationen allerdings als fatal zu werten (Brügelmann, 2006, S. 25). So weist Ingenkamp bereits in den 1970er Jahren (Ingenkamp, 1976 (6. Aufl.)) darauf hin, dass dieselbe Schüler*innen-Arbeit von unterschiedlichen Lehrpersonen völlig unterschiedlich beurteilt wird, so dass teilweise die gesamte Notenskala für die Beurteilung ein und derselben Arbeit ausgeschöpft wird. Dies gilt nicht nur für offenere Formate wie Deutsch-Aufsätze, sondern etwa auch für Arbeiten aus dem Fach Mathematik. Diese Befunde aus den frühen 1970er Jahren konnten in den letzten Jahrzehnten mehrfach repliziert werden (Brügelmann et al., 2006). Sie weisen darüber hinaus auf den Einfluss von sachfremden Informationen (wie Sprachstil bei naturwissenschaftlichen Arbeiten oder Schriftbild bei Deutschaufsätzen) sowie eine systematische Verzerrung der Notengebung durch Merkmale wie Geschlecht oder soziale Herkunft hin. So sind teils z.B. deutliche Unterschiede in den Noten zwischen Mädchen und Burschen zu beobachten, die keine Entsprechung in gemessenen Leistungen haben. Zudem gibt es in Österreich wesentliche regionale Unterschiede in der Notengebung (Eder et al., 2009).

Validität:

Die Validität einer Messung oder Bewertung betrifft deren Gültigkeit. Wird auch tatsächlich das gemessen, was zu messen vorgegeben wird? Die Gültigkeit der Leistungsbeurteilung ist grundsätzlich schwer zu beurteilen, weil die Außenkriterien umstritten sind (Brügelmann et al., 2006). Allerdings deuten wesentliche Befunde auf mangelnde Validität der Notengebung hin.

Zum einen gehen Eder et al. (2009) davon aus, dass in der Praxis beim Zustandekommen der Noten quantitative Aspekte vorrangig bestimmend sind, sodass die Noten die qualitativen Abstufungen der Notendefinitionen nicht abzubilden vermögen. Sie stellen darüber hinaus in Frage, „ob Ziffernnoten überhaupt in der Lage sind, das Leistungsniveau von Schüler/innen im Hinblick auf die (impliziten) Kompetenzmodelle einzelner Fächer angemessen abzubilden“ (Eder et al., S. 257). Damit sprechen sie ein grundlegendes Problem der Notengebung im Hinblick auf die Validität an, nämlich, dass Ziffernnoten die Komplexität der angestrebten Fähigkeiten, über die sie eine Aussage treffen sollten, nicht angemessen abbilden können (vgl. dazu auch Brügelmann, 2006, S. 31).

Den Versuch der Überprüfung von externer Validität machen Studien, die Zeugnisbeurteilungen durch Lehrkräfte in Form von Noten im Vergleich mit externen Leistungsmessungen analysieren. Wie für Deutschland in diversen Studien basierend auf verschiedenen Datensätzen belegt (Brügelmann et al., 2006) gilt auch für Österreich, dass der Zusammenhang zwischen extern gemessenen Leistungen/Kompetenzen und den Noten der Schüler*innen nicht hinreichend groß ist. Insbesondere zwischen verschiedenen Klassen und Schulen sind die von den Lehrpersonen vergebenen Noten nicht vergleichbar. Trotz teils massiver Unterschiede in den Leistungen zwischen Schulen des gleichen Typs fällt die Benotung der Schüler/innen sehr ähnlich aus (für zusammenfassende Darstellungen vgl. Eder et al., 2009, Winter, 2015, Brügelmann et al., 2006; eine empirische Studie auf Basis aktuellerer Daten aus den nationalen österreichischen Standardüberprüfungen findet sich bei Paasch, Schmid, Kallinger-Aufner, & Knollmüller, 2019). Die geringe Vergleichbarkeit von Noten zwischen verschiedenen Klassen führt dazu, dass Schüler*innen, die laut den Noten in einer Klasse zu den Besten gehören, in einer anderen Klasse oder Schule nach ihren Leistungen zu den Schwächsten zählen würden.

Die Validität der Notengebung kann auf anderem Weg in Form von prognostischer Validität, ihrer Vorhersagekraft, geprüft werden. Wie gut geben Noten darüber Auskunft, wie erfolgreich Schüler*innen in späteren Abschnitten des Bildungssystems sein werden? Die prognostische Qualität der Beurteilung in Volksschulen und der Sekundarstufe I ist gering. Die Beurteilungen am Ende jeweils eines Abschnitts des Schulsystems können nur bedingt den Erfolg in den nachfolgenden Segmenten des Schulsystems vorhersagen (Eder et al., 2009). Auch Brügelmann kommt auf Basis der Analysen zu einem ähnlichen Schluss: Die prognostische Validität ist über alle Stufen hinweg gering (2006, S. 23 ff.).

Reliabilität:

Mit Reliabilität ist die Zuverlässigkeit einer Messung oder Beurteilung gemeint. Die Beurteilung soll von äußeren Umständen wie Tageszeit oder Reihenfolge der Prüflinge unabhängig sein. Dies ist weder bei Tests noch informellen Leistungsproben gewährleistet (Brügelmann et al., 2006). Die Zahl von Prüfungen und deren Reihenfolge haben wesentliche Einflüsse auf die Noten. Auch konnte gezeigt werden, dass dieselben Arbeiten von derselben Lehrperson unterschiedlich beurteilt werden, wenn diese sie in zeitlichem Abstand erneut korrigiert. Darüber hinaus zeigen Studien, dass die Reihenfolge, in der

Schüler*innen-Arbeiten korrigiert werden, einen Einfluss auf die Beurteilung haben. Die Beurteilung einer konkreten Arbeit fällt zum Beispiel anders aus, je nachdem ob diese nach einigen Arbeiten sehr guter Qualität oder nach einigen weniger zufriedenstellenden Arbeiten korrigiert wird. Diese mangelnde Reliabilität bezieht sich auf das Lehrer*innen-Urteil, das den Noten zugrunde liegt.

Externe, standardisierte Tests lösen dieses Problem zudem nur für Aussagen über größere Gruppen, wie sie zum Beispiel für bildungspolitische Entscheidungen genutzt werden (etwa im Rahmen von PISA, TIMSS, PIRLS oder den nationalen Standardüberprüfungen). „Ihre Ergebnisse sind aber nicht zureichend verlässlich für die Erfassung und Bewertung individueller Leistungen von SchülerInnen“ (Brügelmann, 2006, S. 27). Dies hat mit dem Messfehler auf Individualebene („Schwankungsbreite“ der Messung; vgl. Gugerell, Kriechmayr, Pacher, Breit, & Wiesner, 2020) zu tun, die sich bei größeren Gruppen ausgleichen, auf der Individualebene aber voll zum Tragen kommen.

Bezugsnormen:

Eine Bewertung von Leistungen oder Kompetenzen ist nur unter Heranziehung eines Vergleichsmaßstabs möglich. Drei Arten von sogenannten Bezugsnormen kommen in Betracht (vgl. Rheinberg, 2001): Die soziale Bezugsnorm bewertet Leistungen eines Individuums im Kontext der Leistungen anderer Personen (einer Bezugsgruppe), z.B. im Vergleich zu den anderen Schülerinnen und Schülern der gleichen Klasse, der gleichen Schule oder einer größeren Bezugsgruppe, z.B. der Altersgruppe. Die kriteriale Bezugsnorm (oder Sachnorm) bewertet die Leistungen eines Individuums im Vergleich zu einem inhaltlichen Kriterium, z.B. im Vergleich zu inhaltlich definierten Kompetenzstufen, dem Erreichen eines definierten Lernziels. Die individuelle Bezugsnorm bewertet die Leistungen eines Individuums zu einem Zeitpunkt, indem diese mit den Leistungen des gleichen Individuums zu einem früheren Zeitpunkt verglichen werden, und hat damit die individuelle Leistungsentwicklung im Fokus.

Abhängig vom gewählten Maßstab wird dieselbe Leistung unterschiedlich bewertet. Bei der Vergabe von Noten dominiert im Allgemeinen der Durchschnitt der eigenen Klasse (also eine eng begrenzte soziale Bezugsnorm, vgl. Brügelmann et al., 2006). Empirische Indizien für diese Tatsache berichten auch Eder et al. (2009, S. 256 f.): In allen Schularten (mit Ausnahme der Volksschule, wo i.d.R. nur etwa eine Hälfte der Notenskala ausgenützt wird), finden sich annähernd normalverteilte Noten, was auf eine starke Berücksichtigung der sozialen Bezugsnorm (innerhalb der jeweiligen Klasse) schließen lässt.

Als lernförderlich gelten Rückmeldungen in Bezug auf inhaltliche Kriterien (kriteriale Bezugsnorm) sowie hinsichtlich individueller Lernfortschritte, insbesondere dann, wenn diese den Schülerinnen und Schülern konkrete Hinweise für die weitere Arbeit geben. Die individuelle Bezugsnorm ist bei Ziffernnoten allerdings nicht möglich (Brügelmann, 2006, S. 29). Eine kriteriale Bezugsnorm als Grundlage von Ziffernnoten kann nur unter extremer Verdichtung von Informationen eines Fachs zu einer einzelnen Ziffernote erfolgen, wodurch jedoch die für den Lernprozess erforderlichen konkreten Informationen verloren gehen.

Zwischenfazit:

Basierend auf diesen Analysen erweisen sich Ziffernnoten als mangelhaft. Einige der Argumente gegen Noten (insbesondere die mangelnde Reliabilität und Objektivität) beziehen sich auf das der Beurteilung zugrundeliegende Lehrer*innen-Urteil und besitzen deshalb auch für alternative Formen der Leistungsbeurteilung grundlegend Gültigkeit. Ziffernnoten werden allerdings Verständlichkeit und

Vergleichbarkeit unterstellt und die Form der Darstellung in Form von Zahlen suggeriert eine Genauigkeit und Objektivität des Urteils, die durch die Leistungsbewertung nicht erfüllt werden können.

Alternative Formen der Leistungsbewertung vermögen in der Regel dieser Personenabhängigkeit des Urteils sowie der Ungenauigkeit der Einschätzung besser Ausdruck zu verleihen, bieten Möglichkeiten einer detaillierteren Rückmeldung, die konkrete Bezüge zu Lernzielen herstellen kann, können aber die mit dem Lehrer*innen-Urteil verbundenen Probleme (mangelnde Objektivität und Reliabilität) ebenso i.d.R. nicht überwinden.

Wirkungen von Noten im Kontext des Lehrens und Lernens

Abgesehen von der Qualität der Beurteilungen, auf der Noten beruhen, haben sich wissenschaftliche Studien mit Wirkungen von Noten auf das Lernen der Schüler*innen auseinandergesetzt. Die wichtigsten sind im Folgenden zusammenfassend dargelegt.

Zur Rückmeldefunktion der Leistungsbewertung

Rückmeldung wird als fundamentaler Bestandteil von Lehren und Lernen angesehen. Durch Rückmeldungen sollen Lernende Informationen darüber erhalten, wie erfolgreich ihre Lernanstrengungen waren, was gute Lösungen waren, welche Wege und Strategien zum Erfolg führten. Lernförderliches Feedback soll möglichst unmittelbar erfolgen, konkret auf transparent kommunizierte Lernziele bezogen sein und Hinweise für den weiteren Lernweg enthalten. Darüber hinaus ist Feedback besonders dann hilfreich, wenn es dialogisch angelegt ist, und wird erst wirksam, wenn sich der oder die Lernende auch tatsächlich damit auseinandersetzt (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). In diesem Kontext sind Noten bezüglich ihrer Rückmeldequalität in mehrfacher Hinsicht problematisch. Schon Klafki (1985) bezog sich darauf, dass die herkömmliche Praxis der Leistungsbeurteilung insbesondere durch ihre Ausrichtung auf die Endprodukte hin keine differenzierten, inhaltlich aussagekräftigen Informationen liefern würde, weil sie keine Bezüge zu Lernzielen realisiert und keine oder zu wenig detaillierte Auskunft über Teilleistungen gibt. Winter resümiert, dass „die Kommunikation über Leistungen durch die Ziffernnoten eher erschwert als erleichtert zu werden“ scheint und „sich zudem leicht in den Bereich der negativen ‚Leistungen‘ oder Fehler“ (Winter, 2015, S. 65) verschiebt.

Auswirkungen auf die Motivation

Die Motivationspsychologie beurteilt die Möglichkeiten, Lernende mithilfe von Prüfungen und Noten zu einem tiefer gehenden Lernen zu bewegen, skeptisch. Noten rufen negative affektive Reaktionen hervor und bewirken ein qualitativ schlechteres Lernen (Winter, 2015).

Dies hängt damit zusammen, welche langfristigen motivationalen Orientierungen die Schule bzw. ein Leistungsbewertungssystem hervorbringt. Dabei werden zwei grundlegende motivationale Orientierungen unterschieden: Die Motivation an der Sache (etwa Interesse als intrinsische Motivation), die die Auseinandersetzung mit dem Lerngegenstand und die Bewältigung damit zusammenhängender Herausforderungen anregt; oder die Motivation am Leistungsrang, die Lernbemühungen vor allem

dadurch anregt, dass eine gute Note, ein bestimmter Status innerhalb der Klasse erzielt werden soll (vgl. dazu Deci & Ryan, 1993). „Die Benotung fördert zweifellos die Orientierung der Schülerinnen und Schüler darauf, sich auf ihren Leistungsrang zu beziehen („Ich bin in Englisch ein Zweierschüler“) und darauf zu achten, dass sie im Rahmen ihrer Klasse ihr Niveau behaupten oder verbessern“ (Winter, 2015, S. 66). Positiv auf die Lern- und Leistungsmotivation wirkt sich eine Entwicklungsorientierung der Leistungsbeurteilung aus. Insbesondere bei leistungsschwächeren und ängstlichen Kindern gibt es positive Wirkungen durch einen Notenverzicht. Bei leistungsstärkeren Kindern führen Noten zu einer stärkeren Orientierung an äußerer Anerkennung („extrinsische Motivation“) – auf Kosten persönlicher Interessen und damit eines Verlusts an intrinsischer Motivation. Dabei gilt intrinsische Motivation langfristig als wesentlich wichtiger für den Kompetenzerwerb (vgl. Brügelmann et al., 2006).

Insgesamt verschärfen Noten den Selektions- und erhöhen den Konkurrenzdruck (insbesondere in Kombination mit der starken Orientierung an einer sozialen Bezugsnorm; siehe oben), was vor allem im unteren Leistungssegment zu einem Gegeneinander von Lernenden und Lehrenden führt (Brügelmann, 2006). In Klassen ohne Noten fühlen sich Schüler*innen wohler. Ein Verzicht auf Noten führt nicht zu einem Leistungsabfall (ebd., S. 33), die „verbreitete Annahme, dass SchülerInnen nur lernen, wenn ein entsprechender Druck ausgeübt wird, ist falsch“ (ebd., S. 34).

Zusammenfassung

Die vorliegende Stellungnahme fasst vorliegende wissenschaftliche Erkenntnisse zur Verwendung von Ziffernnoten im Rahmen der schulischen Leistungsbeurteilung zusammen. Dabei ist zu beachten, dass – insbesondere aufgrund der Grundlage von Lehrer*innen-Urteilen schulischer Leistungsbeurteilung – die Bewertung nach informellen Proben und Beobachtungen in hohem Maße fehleranfällig ist und demnach die wissenschaftlichen Gütekriterien Objektivität, Validität und Reliabilität (Personenunabhängigkeit, Gültigkeit und Zuverlässigkeit) durch Noten nur unzureichend erfüllt werden können. Zudem werden Ziffernnoten mit einer sozialen Bezugsnorm assoziiert, die Konkurrenzdruck erzeugt, negative Auswirkungen auf die Lernmotivation hat und keinen Bezug zu inhaltlichen Lernzielen aufweist. Alternative Formen der Leistungsbewertung unterliegen ebenso den Bedingungen, dass sie auf Lehrer*innen-Urteilen beruhen und den damit verbundenen Problemen. Sie ermöglichen allerdings, Bezugsnormen größeres Gewicht zu geben, die für den Lernprozess förderlich sind (Individualnorm, Sachnorm) und weisen das Potenzial auf, detailliertere, stärker auf die Lernziele fokussierte Rückmeldungen zu geben, was sowohl in Bezug auf die Motivation als auch die direkte Wirkung auf den Lernprozess günstig ist.

In Bezug auf die einleitend angesprochenen Funktionen der Leistungsbeurteilung – Rückmeldung, Bericht, Selektion – kann Folgendes zusammenfassend festgehalten werden.

Rückmeldung wird als fundamental für Lernen angesehen. Durch Rückmeldungen erhalten Lernende Aufschluss darüber, wie erfolgreich ihre Anstrengungen waren und wo sie im Lernprozess stehen. Da Beurteilungen in Bezug auf die Rückmeldefunktion dann besonders hilfreich sind, wenn sie den Schülerinnen und Schülern konkrete Hinweise für die weitere Arbeit geben, sind Ziffernnoten als lernförderliche Rückmeldungen nicht geeignet.

In ihrer *Berichtsfunktion* sollen Noten Auskunft darüber geben, was Schülerinnen und Schüler in Bezug auf die anzustrebenden Kompetenzen eines Fachs bereits gelernt haben. Allerdings können Ziffern die Komplexität der anzustrebenden Kompetenzen nicht abbilden und sind deshalb wenig aussagekräftig, suggerieren dabei jedoch Genauigkeit, Vergleichbarkeit und Prognosefähigkeit, obwohl sie diese Kriterien nicht erfüllen.

Brügelmann merkt bezüglich der *Selektionsfunktion* der Notengebung an, dass in der Schule „die Förder- und Berichtsfunktion Vorrang haben“ müssen (2006, S. 40). Dies gilt für die ersten Schulstufen in besonderem Maße. Die geringe prognostische Validität und die mangelnde Objektivität und Zuverlässigkeit, die der schulischen Leistungsbewertung in Studien bescheinigt wird, „entziehen ihnen [den Bewertungen] die Grundlage für Selektionsentscheidungen (ebd., S. 27).

Die weite Akzeptanz der ziffernnotenbasierten Beurteilung beruht vor allem auf der historisch gewachsenen Vertrautheit mit diesem System über Generationen hinweg. Schüler*innen, Eltern und Lehrpersonen sind mit dem System vertraut und haben mit anderen Ansätzen in der Regel wenig bis keine Erfahrung. Vor dem Hintergrund der empirischen Befunde zur Qualität schulischer Leistungsbeurteilung erscheint dies allerdings nur bedingt als Vorteil.

Für das ILS: Ass.-Prof. Dr. Claudia Schreiner
Innsbruck, am 21. 08. 2020

Quellen:

- Brügelmann, H. (2006). Sind Noten nützlich – Und nötig? Ziffernzensuren und ihre Alternativen im empirischen Vergleich. Eine Expertise der Arbeitsgruppe Primarstufe an der Universität Siegen im Auftrag des Grundschulverbands e. V., Frankfurt (Kurzfassung). In H. Bartnitzky, H. Brügelmann, U. Hecker, & G. Schönknecht (Hrsg.), *Pädagogische Leistungskultur* (S. 17–46). Frankfurt am Main: Grundschulverband – Arbeitskreis Grundschule e.V. 2006.
- Brügelmann, H., Backhaus, A., Brinkmann, E., Coelen, H., Franzkowiak, T., Knorre, S., ... Roth, S. (2006). *Sind Noten nützlich – Und nötig? Ziffernzensuren und ihre Alternativen im empirischen Vergleich. Eine wissenschaftliche Expertise des Grundschulverbandes erstellt von der Arbeitsgruppe Primarstufe an der Universität Siegen*. Frankfurt am Main: Grundschulverband – Arbeitskreis Grundschule e.V.
- Deci, E. L., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 2(39), 223–238.
- Eder, F., Neuweg, G. H., & Thonhauser, J. (2009). Leistungsfeststellung und Leistungsbeurteilung. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich 2009. Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (Bd. 2, S. 247–267). Graz: Leykam.
- Gugerell, S., Kriechmayr, C., Pacher, K., Breit, S., & Wiesner, C. (2020). IKM: Überblick, intendierte und tatsächliche Nutzung. In U. Greiner, F. Hofmann, C. Schreiner, & C. Wiesner (Hrsg.), *Bildungsstandards. Kompetenzorientierung, Aufgabenkultur und Qualitätsentwicklung im Schulsystem* (S. 102–123). Münster: Waxmann.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
- Hesse, I., & Latzko, B. (2017). *Diagnostik für Lehrkräfte*. Opladen, Toronto: Barbara Budrich.

- Ingenkamp, K. (1976). *Die Fragwürdigkeit der Zensurenggebung: Texte u. Untersuchungsberichte* (6., überarb. u. erw. Aufl.). Weinheim, Basel: Beltz.
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik* (6. Aufl.). Weinheim, Basel: Beltz.
- Klafki, W. (1985). *Neue Studien zur Bildungstheorie und Didaktik: Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik* (6. Aufl.). Weinheim: Beltz.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.
<https://doi.org/10.1080/03075070600572090>
- Paasch, D., Schmid, C., Kallinger-Aufner, A., & Knollmüller, R. (2019). Noten und Kompetenzen in verschiedenen Fächern, Schulstufen und Schulformen. In A. C. George, C. Schreiner, C. Wiesner, M. Pointinger, & K. Pacher (Hrsg.), *Fünf Jahre flächendeckende Bildungsstandardüberprüfungen in Österreich. Vertiefende Analysen zum Zyklus 2012 bis 2016* (S. 161–178). Münster: Waxmann.
- Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 59–71). Weinheim: Beltz.
- Winter, F. (2015). *Lerndialog statt Noten: Neue Formen der Leistungsbeurteilung*. Weinheim, Basel: Beltz.