

Editorial: Die Urteilsbildung im Unterricht

Franz Baeriswyl und Lucie Mottier Lopez

In den thematischen Beiträgen dieser Nummer wird die Genese der Urteilsbildung im Unterricht untersucht. Gegenstand und Funktion der Beurteilung im Unterricht sind je mannigfaltig. So wird die fachspezifische Leistungsfähigkeit, das Wissen und Können aber auch überfachliche Kompetenzen wie Arbeitshaltung, Engagement oder Kommunikationskompetenz beurteilt. Die Beurteilungsfunktion umfasst die Zielsetzung und die Beurteilungsform. Die lernprozesssteuernde (formative) Beurteilung orientiert die lernende Person über Fortschritte und zeigt das Gekonnnte und die Lücken auf. Sie liefert der Lehrperson aber auch wichtige Erkenntnisse über die Bedürfnisse der Lernenden und hilft ihr den Unterricht zu differenzieren. Die lernsteuernde Beurteilung ist also kein abschliessendes Urteil, wie der Fall der summativen oder zertifizierten Beurteilung ist. Letztere bescheinigt den Abschluss eines Ausbildungsganges oder eines zeitlich strukturierten Abschlusses wie beispielweise eines Schuljahres. Bei Schulübergängen wird von den Lehrpersonen oft eine holistische Schulzuweisungs- oder Schulgangsempfehlung abgefordert, die einen speziell prognostischen Charakter hat.

Mit der Übertragung der psychometrischen Qualitätsanforderungen auf die Beurteilung im Schulkontext entstand in den siebziger und achziger Jahren des letzten Jahrhunderts eine Flut von empirischen Untersuchungen, welche die Unzulänglichkeiten der schulischen Beurteilung aufzeigten. Nur selten waren hingegen die Stimmen, die auf den diesen Studien zugrunde liegenden paradigmatischen Irrtum aufmerksam machten, dass die Gütekriterien der Psychometrie nicht für die Beurteilung im Klassenzimmer gelten dürfen, wenn der pädagogische Auftrag erfüllt werden soll. Trotzdem haben diese oft schulzuweisenden Forschungsergebnisse in der Bildungslandschaft Wesentliches bewirkt, indem eine eigenständige Forschung mit einer angepassten Forschungsmethodik ausgelöst wurde.

Aspekte dieser erziehungswissenschaftlichen Forschung erscheinen in den thematischen Beiträgen. Dabei wird ein neuer Forschungsfokus gesetzt: Es interessiert die Genese der Urteilsbildung einer Lehrperson über verschiedene Beur-

teilungsgegenstände. Um diesen neuartigen Gesichtspunkt zu klären, wird im Folgenden eine begriffliche Diskussion publiziert.

Alle Beiträge dieser Ausgabe befassen sich mit der Urteilsbildung der Lehrperson in Situationen zertifizierender oder prognostischer Beurteilung, in Situationen von grosser Bedeutung für den Ausbildungsgang der Schülerinnen und Schüler. Solche Übergangssituationen sind nicht nur besonders empfindlich, insbesondere, wenn die Schulsysteme nur schwer Möglichkeiten von Neuorientierungen erlauben (Baeriswyl et al; Maaz et al., in dieser Ausgabe), sondern sie rufen darüber hinaus auch prognostische Beurteilungen hervor, die immer ein gewisses Mass an Unsicherheit beinhalten, welches niemals vollständig ausgeschaltet werden kann. Dies sind Formen von «Wetten auf die Zukunft», deren sich die Lehrpersonen bewusst sind, wie Marcoux und Crahay in dieser Ausgabe festhalten.

Die im ersten Artikel von *Lucie Mottier Lopez* und *Linda Allal* vorgestellte Forschung untersucht die Merkmale der professionellen Schülerbeurteilung von Genfer Lehrpersonen am Ende der Primarschule. Die Ergebnisse der Forschung beschreiben die Verfahren ihrer Notengebung und die Art der von ihnen vorgenommenen Beurteilungsanpassungen, für jene Schüler die auf der Kippe stehen. Die professionelle Schülerbeurteilung wird in diesem Artikel als ein kognitiver Prozess und eine an verschiedene Kontexte geknüpfte soziale Praxis beschrieben, die die Autorinnen dazu führen, die individuellen und sozialen Dimensionen der professionellen Urteilsbildung zu beschreiben.

Dany Laveault schafft einen Literaturüberblick über den Begriff der «professionellen Schülerbeurteilung», indem er sowohl nordamerikanische (französisch- und englischsprachige) als auch europäische Beiträge auswertet. Der Autor beleuchtet verschiedene Beurteilungskategorien und -indikatoren und diskutiert diese im Hinblick auf professionelle Praxis und Reformanforderungen in der Bildung. Die Komplexität der zertifizierenden und prognostischen Beurteilung wird hervorgehoben (z.B. wichtige Herausforderungen, multiple Ziele, mögliche Loyalitätskonflikte), sowie die «Synergien», die entwickelt werden müssten, um die Qualität der professionellen Urteilsbildung der Lehrpersonen zu verbessern.

Der Beitrag von *Géry Marcoux* und *Marcel Crahay* hat zum Ziel, in einem psychologischen Ansatz die Komplexität der Maßnahme zu analysieren, die die Lehrperson zur Entscheidung der Klassenrepetition bewegt- und dies trotz der zahlreichen Forschungsergebnisse, die die Unwirksamkeit dieser Maßnahme aufzeigen. Durch Interviews versuchen die Forscher, das Lehrerurteil sowohl im Falle einer Promotionsentscheidung als auch im Falle einer Klassenrepetition zu verstehen. Auf der Grundlage ihrer Forschungserkenntnisse diskutieren die Autoren vor allem die Professionalität und nicht die Urteile der Lehrer, wenn diese trotz allem ihre Schüler weiterhin wiederholen lassen.

Kai Maaz, Marko Neumann, Ulrich Trautwein, Wolfgang Wendt, Rainer Lehmann und Juergen Baumert untersuchen die Lernkompetenzeinschätzung der Lehrpersonen beim Übergang nach der sechsten Klasse der Grundschule in die gegliederte Sekundarstufe I in Berlin. Die Autoren diskutieren die verschiedenen «Quellen» oder Elemente, welche für die Bildung der Lernkompetenzeinschätzung herangezogen werden. Dabei haben sie einen systematischen Geschlechtereffekt und einen Effekt des sozioökonomischen Status festgestellt. Ebenso beeinflusst die Leistungsstärke der Klasse die Urteilsbildung der Lehrperson, was als Referenzgruppeneffekt bezeichnet werden kann.

Eine ähnliche Fragestellung bearbeiten *Franz Baeriswyl, Christian Wandeler und Oliver Christ*, indem sie die Stabilität der Zusammensetzung der Lehrerempfehlung beim Übertritt in die Sekundarstufe I untersuchen. Obwohl ein Grossteil der Lehrpersonen von einem Jahr zum anderen variiert, stellen sie keine statistisch interpretierbare Veränderung der Lehrerempfehlung über sieben Jahre fest. Von besonderem Interesse ist der Vergleich dieser Lehrerempfehlung mit der Lernkompetenzeinschätzung von Maaz et al., die einen rigorosen Geschlechtereffekt konstatieren, der bei Baeriswyl et al. nicht systematisch feststellbar ist. Die zeitliche Stabilität der Lehrerempfehlung wird als positives Merkmal für Gerechtigkeit interpretiert.

Der dritte deutschsprachige Beitrag von *Anke Treutlein, Jeanette Roos und Hermann Schöler* untersucht den Referenzgruppeneffekt einer Klasse bezüglich der individuellen Notengebung in Rechtschreiben, Lesen und der Gesamtbeurteilung Deutsch. Entgegen der Annahme und auch der Ergebnisse von Maaz et al. zeigte sich ein nur geringer oder kein Referenzgruppeneffekt. Dieses eher erwartungswidrige Ergebnis diskutierten sie ausführlich und zeigen mehrere Erklärungsmöglichkeiten auf.

Die sechs thematischen Beiträge zur Urteilsbildung im Unterricht fokussieren verschiedene Beurteilungsformen und -arten. So beginnt sich ein Puzzle zur Genese der Urteilsbildung zu formen, das allein mit diesen Beiträgen aber noch kein differenziertes Bild erkennen lässt. Die Artikel dieser Ausgabe erlauben jedoch, die Konzeptualisierung und das Verständnis der Prozesse in der pädagogischen Beurteilung weiterzuverfolgen, indem vor allem die schwierige Frage des Verhältnisses zwischen Beurteilung und Bewertung gestellt wird. Von Interesse ist unter anderem, die Bewertungsprozesse in der regulären Beurteilungspraxis zu untersuchen, insbesondere in Situationen, die sich, wie eingangs erwähnt wurde, durch hohe Einsätze auf individueller und kollektiver Ebene auszeichnen.

Die Problematik der Urteilsbildung ist erheblich, besonders unter Berücksichtigung von aktuellen bildungspolitischen Tendenzen, die versuchen, den Spielraum und die Professionalität der Lehrpersonen zu verringern, indem mechanische Verfahren zur Entscheidungsfindung eingeführt werden. Wir denken dabei an Verfahren, die sich im Wesentlichen auf Berechnungen der Durch-

schnittsnoten basieren, und die das Problem der ungenügenden Berücksichtigung anderer Informationen aufwerfen, von denen die (in dieser Ausgabe) im Folgenden dargestellten Forschungsarbeiten zeigen, dass diese Informationen für die Entscheidungen zertifizierender und prognostischer Bewertung unabdinglich und informativ sind.