

Wir brauchen formatives Assessment mit Tests, die dem Lernen wirklich nützen

Texte:
Wolfgang Beywl,
John Hattie

Lehrpersonen wollen Heranwachsende auf das Leben vorbereiten, damit sie die komplexe Welt sachkundig und reflektiert mitgestalten können. Die Schule soll allen Kindern und Jugendlichen ermöglichen, das hierfür erforderliche Wissen und Können zu erwerben. Formatives Assessment, unter anderem in Verbindung mit Tests, kann zu einem Unterricht beitragen, der dem Ziel individueller Lernförderung optimal gerecht wird.

Die Vielfalt von Unterrichtsmethoden ist gewachsen. Wir wissen aus der Forschung, welche Methoden bei welchen Schülergruppen besser wirken. Viel Potenzial bergen demnach Methoden, die auf wechselseitige Rückmeldungen zwischen Lehrenden und Lernenden setzen, zum Beispiel Feedback. Professionelle Lehrpersonen kennen diese Forschung. Sie vergewissern sich mit Hilfe mehrperspektivisch gewonnener Daten ihres Einflusses auf Lernen und Lernresultate. Hierfür nutzen sie beispielsweise unterrichtsintegrierte Kurzerhebungen mit dem schulinternen Weiterbildungsangebot Luuise, aber auch standardisierte Tests.

Vergleichend messbar sind Fachkompetenzen

Tests sind Zusammenstellungen von Aufgaben, deren Bearbeitung gute Schlüsse auf den Lernstand ermöglicht. «Standardisiert» bedeutet, dass die Tests nicht willkürlich verändert werden dürfen. Ausserdem müssen sie im Hinblick auf Dauer, Nutzung von Hilfsmitteln und dergleichen regelgetreu eingesetzt werden. Für die Messung von fachlichen Kompetenzen sind Standardtests weit verbreitet, so beispielsweise seit 1964 der TOEFL (Englishtest für Nicht-Muttersprachler). TIMMS- und PISA-Schulleistungsvergleiche werden weltweit fast identisch eingesetzt. Die hohen Entwicklungskosten sprechen dafür, das Messen auf ausgewählte Bereiche fachlicher Kompetenz zu konzentrieren. Aber: Das darf nicht dazu führen, dass nicht oder nur schwer messbare Aspekte, wie etwa Persönlichkeitsbildung oder Wissen und Können aus anderen Fachbereichen, an Wichtigkeit verlieren.

Formative und summative Rolle von Tests

Tests können zwei verschiedene Rollen spielen: die formative und die summative. Dies hängt vom Einsatzzeitpunkt und von der Verwendungsabsicht ab. Solange Lehr- und Lernprozess andauern, können Tests helfen zu optimieren. Sie können für das gezielte, womöglich sogar personalisierte Weiterlernen Orientierung bieten, also eine formative Rolle einnehmen. Ist ein Unterrichtsthema abschliessend behan-

delt, kann hingegen mit Tests «bilanziert» oder summativ evaluiert werden (meint: «die Summe unter dem Strich bilden»). Beide Rollen sind wichtig. Sie klar auseinanderzuhalten, ist entscheidend. So muss der Zweck des Tests eindeutig kommuniziert sein: Handelt es sich beispielsweise um einen formativen Test zur Lernförderung oder um einen summativen zur Notengebung? Formativ und summativ dürfen nicht vermischt oder gar erst im Nachhinein deklariert werden. Denn sobald Tests für Entscheidungen verwendet werden, mit denen für den weiteren Lebenslauf tatsächliche oder phantasiert hohe Einsätze verbunden sind, so etwa in der Promotion, leisten sie formativ weniger. Dann werden sich Schülerinnen und Schüler auf gutes Abschneiden konzentrieren – notfalls durch Reduktion auf Oberflächenwissen, Einnahme leistungssteigernder Medikamente, Täuschen usw. Jedenfalls hat sich das formative Fenster dann geschlossen.

Da das Abschneiden zum summativen Endpunkt massgeblich von der formativen Begleitung während des Lernprozesses abhängt, sollten wir besonders in Letztere investieren. Neben emotionaler Sicherheit im Klassenzimmer ist Folgendes wichtig: Lehrpersonen nutzen die Daten für die Entwicklung des Unterrichts sicher und gekonnt. Dazu kann ein Lernförder- und Testsystem massgeblich beitragen.

Das neuseeländische asTTle

Ein Beispiel für nutzungsfreundliche Tests ist das vor 15 Jahren mit einem Aufwand von rund 10 Millionen Franken entwickelte asTTle für die 2500 Schulen Neuseelands. Es ist eines unter mehreren Tools mit hohem formativem Potenzial. Das «Assessment tool for teaching and learning» wird auf freiwilliger Basis von 70 Prozent aller Schulen mit steigender Tendenz eingesetzt. Seine etwa 20000 validierten Testaufgaben für die Klassenstufen 1 bis 10 ermöglichen Tausende massgeschneiderter Testversionen. Mittels der ICT-Schnittstelle des unterdessen webbasierten e-asTTle können Lehrpersonen massgeschneiderte Tests anfordern: zu dem, was gerade unterrichtet wird und gelernt werden

soll. Für den Einsatz als Online-Test kann er in einer adaptiven Version heruntergeladen werden, die sich fortlaufend auf den im Testverlauf erfassten Lernstand der jeweiligen Lernenden anpasst. Auch Paralleltests, also gleichwertige Tests mit verschiedenen Aufgaben, können erstellt werden. Es lassen sich individuelle Lernzuwächse bestimmen, als Basis und Ergebnis eines differenzierten Unterrichts. «Testen» kann zu mehr führen als zu isolierten Informationen zum Leistungsstand. Formativ können Testdaten dazu beitragen, dass Lehrende und Lernende ihr Lehren und Lernen aufeinander abstimmen. Die Zusammenstellung, Auswertung und Interpretation der Tests sowie die darauf aufbauende Unterrichtsplanung erarbeiten Lehrpersonen der jeweiligen Klassen im Team. Hierfür müssen Berichte, die ein System wie asTTle bereitstellt, leicht nutzbar sein.

Test-Konzeption aus der Perspektive der Nutzung

Daher wurde das asTTle von der Nutzungssituation her «rückwärts» konzipiert. In Fokusgruppen diskutierten Lehrpersonen Musterberichte, die dann immer wieder verbessert wurden. So blieb es den Lehrpersonen erspart, sich in das Denken der Messexperten einzuarbeiten. Stattdessen haben die Testkonstrukteure ihre Angebote auf Praktikabilität für die Nutzenden hin gestaltet. Derart entlastete Lehrende können sich weiter auf ihre herausfordernde Hauptaufgabe, die Gestaltung des Unterrichts, konzentrieren. Wie bei asTTle

sollten schulinterne Tests nicht für regionale oder nationale Rechenschaftslegung genutzt werden. Dies erleichtert es, eine von Lehrpersonen und Schulleitungen getragene formative Kultur zu stabilisieren. Die zu anderen Zeitpunkten des Schuljahres erforderlichen summativen Nutzungen sind davon klar zu trennen. Formatives Assessment kann so als zentraler Bestandteil der Lehrprofession verstärkt werden. ■

Weiter im Text

Beywl, Wolfgang und Odermatt, Miranda (2016): Luuise – Lehrkräfte untersuchen und reflektieren ihren eigenen Unterricht. *Journal für Schulentwicklung*, 20(4), 33–39.

Hattie, John A. C. (2015): *Lernen sichtbar machen*, erw. 3. Auflage. Schneider Hohengehren.

Scriven, Michael (1972): «Die Methodologie der Evaluation». In: Wulf, Christoph (Hg.): *Evaluation*. Piper, S. 60–91.

Gavin T. L. Brown und Harris, Lois R.: *Handbook of Human and Social Conditions in Assessment*. New York: Routledge 2016 – Das Handbuch bereitet die aktuelle Forschungsbasis zum Thema auf.

Weiter im Netz

e-asttle.tki.org.nz – Der Practice Test vermittelt einen Eindruck und findet sich auf der rechten Seite als Link.

www.bildungbern.ch/publikationen/schulpraxis – Themenheft «Lernen sichtbar machen»

www.lernensichtbarmachen.ch

Assessment

Dieses englische Fachwort sperrt sich gegen Übersetzungsversuche wie «Einschätzung» oder «Beurteilung». Assessments enthalten Beschreiben, Feststellen, Diagnostizieren, Messen und nachfolgendes kriteriengestütztes Bewerten. Lernende können sich selbst (Selbst-Assessment) oder gegenseitig einschätzen (Peer-Assessment). Assessments durch Lehrpersonen spüren vor, welche Aufgaben als Nächstes gestellt, wie Fragen in der Klasse beantwortet, wie Schülerarbeiten kommentiert werden (= formative assessment for learning) oder wie faire Noten gegeben werden (= summative assessment of learning). Daher ist das Thema so wichtig.

e-asTTle-Statistik

Im Jahr 2016 nutzten 72 Prozent der Schulen in Neuseeland das Tool e-asTTle, davon haben 47 Prozent mehr als 100 Durchgänge getätigt. Insgesamt wurde e-asTTle 2016 1,4 Millionen Mal angewendet. Dabei wurde das Tool zu 28 Prozent von Lehrpersonen und zu 72 Prozent von Schülerinnen und Schülern für Selbsttests genutzt. Die Tests wurden zu zwei Dritteln am Bildschirm und zu einem Drittel auf Papier ausgeführt.

Standardisierte Tests in der Schweiz

Tests werden seit Mitte der 1990er-Jahre an Deutschschweizer Schulen verbreitet

eingesetzt. Für die Volksschule in vielen Kantonen obligatorisch sind u. a. Klassencockpit (3.–9. Kl.), Stellwerk (8./9. Kl., adaptiv), Checks in der Nordwestschweiz (3., 6., 8., 9. Kl., teils adaptiv). Sie sind auch für summative Nutzungen vorgesehen, zum Beispiel für den Abgleich bei der Notengebung, als Beilage für die Lehrstellenbewerbungen und als Bestandteil des Zeugnisses oder des Abschlusszertifikats. Je nach Kanton werden Schulen zu Letzterem verpflichtet oder es wird sogar «abgeraten» (nur Bern). Zum anderen beanspruchen (die gleichen) Tests auch eine formative Rolle. Mehrere Begleituntersuchungen weisen besonders bezüglich der Unterrichtswirksamkeit auf Schwächen hin, ebenso bezüglich der Relation Aufwand–

Ertrag. Bis 2020 sollen auf den Lehrplan 21 abgestimmte «Lernfördersysteme» einsatzbereit sein, z. B. Mindsteps und Lernlupe für die Volksschule sowie Lernnavi für die Sek II. Zu wünschen wäre, dass diese nicht lediglich, wie die bisherigen, für Summatives «akzeptiert» werden, sondern dass sie tatsächlich formativ direkt von Lehrenden und Lernenden genutzt werden und so zu effektivem Lehren und Lernen beitragen. Wünschenswert wäre auch, dass bei jeder Durchführung eines Tests eindeutig deklariert wird, ob er einem formativen oder einem summativen Zweck dient – und dass dies ausnahmslos eingehalten wird.