

Fetische des modernen Schulwesens

Episode 2: In der Rasterfalle – der ausgeklügelteste Beurteilungsbogen

Von Philipp Loretz



Als Fetisch bezeichnet man einen verehrten Gegenstand, dem man geradezu magische Eigenschaften zuschreibt. Auch im sich als modern verstehenden Schulbetrieb gibt es einige Elemente, deren Einsatz in immer stärkerem Masse als unverzichtbares Qualitätsmerkmal gepriesen wird. Es ist daher an der Zeit, ebendiese Elemente einem kritischen Nachdenken zu unterziehen.

M

Magie auf dem Bildschirm

Als Kind der 1980er Jahre erinnere ich mich noch gut daran, wie fasziniert ich eine der ersten Sendungen von «Wetten, dass ...?» verfolgte; dies aber nicht wegen einer besonders skurrilen Wette, sondern aufgrund meiner ersten «Begegnung» mit einem Simultandolmetscher. Ich konnte mir schlicht nicht erklären, wie der unsichtbare Mann im Hintergrund *gleichzeitig* zuhören, übersetzen und sprechen konnte – und es dabei auch noch schaffte, völlig locker zu bleiben und freundlich zu klingen.

Dieser Moment hatte für mich etwas Magisches. Also beschloss ich, später einmal selbst Dolmetscher zu werden.

M

Multitasking?

Und Sie? Beherrschen Sie Multitasking? Wenn ja, wie viele Dinge können Sie *gleichzeitig* tun? Hemden zu

bügeln und gleichzeitig Musik zu hören, stellt wohl noch kaum eine unlösbare Herausforderung dar, handelt es sich doch um eine *automatisierte* Tätigkeit einerseits und eine *passive* «Leistung» andererseits. Das lässt sich prima kombinieren.

Eine Prüfung zu entwerfen und *gleichzeitig* Ihrem Kind seine Französischhausaufgaben zu erklären, dürfte Sie dagegen bereits an die Grenzen Ihrer Multitasking-Künste bringen, da das menschliche Hirn nicht darauf ausgelegt ist, zwei *bewusste* Handlungen *gleichzeitig* zu bewältigen. Formulieren ist eine *aktive* «Leistung», bewusstes Zuhören und Erklären ebenfalls. Übrigens: Frank Schirrmacher, Mitherausgeber der «Frankfurter Allgemeinen Zeitung», bezeichnete Multitasking schon 2009 in seinem Buch «Payback» als Körperverletzung.

Machen Sie die Probe aufs Exempel mit Hilfe dieses Miniexperiments: Schliessen Sie die Augen und versuchen Sie, innerlich die 16er-Reihe aufzusagen und gleichzeitig den Lufthauch an den Nasenflügeln beim Ein- und Ausatmen *permanent* wahrzunehmen! Sie werden feststellen, dass

es sich dabei um ein unmögliches Unterfangen handelt: Man kann lediglich *entweder* rechnen *oder* den Lufthauch registrieren – genau so, wie man nur *entweder* eine Diskussionsendung aufmerksam verfolgen *oder* die neuesten Posts auf Facebook checken kann.

S

Selbstüberschätzung

Was passieren kann, wenn wir unsere Wahrnehmungsfähigkeiten überschätzen, zeigt eindrücklich ein im Netz kursierender Clip mit dem Slogan «You can't even text and walk, so why do you text and drive?»¹ Indes behaupten nicht wenige Leute in meinem Freundeskreis, dazu imstande zu sein, *gleichzeitig* eine E-Mail zu verfassen, einen Telefonanruf entgegenzunehmen und die neusten Termine, auf die sie Ihre Partner respektive Partnerinnen en passant hinweisen, zuverlässig abzuspeichern. In Wahrheit dürften die Chancen gut stehen, aus den Mündern dieser Personen dereinst den folgenden Satz zu hören zu bekommen: «Meine Partnerin (respektive mein Partner) wirft mir zwei Dinge vor: Ich

würde ihr (ihm) nicht richtig zuhören und ... noch etwas anderes.»

«Meine Partnerin wirft mir zwei Dinge vor: Ich würde ihr nicht richtig zuhören und ... noch etwas anderes.»

Diese alltäglichen Beispiele zeigen auf, dass es sich beim vermeintlichen Multitasking lediglich um die Fähigkeit handelt, blitzschnell zwischen zwei oder mehreren Tätigkeiten hin- und herzuschalten. Sie zweifeln? Überzeugen Sie sich selbst anhand eines aufschlussreichen 80-Sekunden-Clips, **bevor** Sie weiterlesen. Tippen Sie in Ihrem Browser exakt den Begriff «selective awareness» ein und klicken Sie anschliessend auf den [Clip von Daniel Simons](#)². Sie werden sodann dazu aufgefordert werden, zu zählen, wie oft die Spieler in den weissen T-Shirts sich gegenseitig den Ball zuspielen. Bitte lesen Sie den vorliegenden Artikel erst weiter, **nachdem** Sie diesen Wahrnehmungstest durchgeführt haben – andernfalls bringen Sie sich um

eine erhellende Erkenntnis. Also dann: Schnappen Sie sich jetzt Ihren Laptop oder Ihr Smartphone und bis gleich!

Willkommen zurück! Ich hoffe, es hat geklappt! Wenn ich diesen Test jeweils mit meinen Schülerinnen und Schülern durchführe, decken sich die Resultate mit Simons' Ergebnissen, die am Schluss des Kurzfilms eingeblendet werden: Rund die Hälfte der Klasse lacht inmitten des Films auf, die andere Hälfte bleibt fokussiert und zählt konzentriert weiter. Die «Gorilla-Spotter» freuen sich über ihre vermeintlich hochstehenden Wahrnehmungsfähigkeiten; die Fokussierten, welche richtig gezählt haben, ebenfalls. Nur korrekt zählende «Gorilla-Spotter» finden sich keine.

U

Und Sie persönlich? Und wir Lehrpersonen?

Stellt sich die Frage: Zu welcher Kategorie von Probanden gehören Sie? Zu denjenigen, welche das Unerwartete

entdecken, sich dadurch aber ablenken lassen und deshalb die eigentliche Aufgabe aus den Augen verlieren? Oder zu jenen, denen es zwar gelingt, fokussiert zu bleiben und so den Auftrag kriteriengetreu zu lösen, dabei jedoch das an sich Unübersehbare – einen Gorilla mitten unter den Basketballspielern! – vollständig ausblenden?

Auf der [Webseite](#) von Daniel Simons und Christopher Chabris³ finden sich eine anspruchsvollere Variante der sogenannten *monkey business illusion* und weitere Aufmerksamkeits-Tests, die uns mit Nachdruck vor Augen führen, wie erstaunlich selektiv, um nicht zu sagen: beschränkt, unsere Wahrnehmung ist. Für die Sportlehrpersonen unter Ihnen dürfte dieser Test⁴ (ab Minute 01:07) interessant sein. Egal, wie sehr wir uns konzentrieren, «wir sind nicht in der Lage, eine Situation vollständig zu erfassen.» Dies allein wäre an sich auch gar nicht weiter problematisch. Der Umstand, dass wir aber irrtümlicherweise davon überzeugt sind, unserer Aufmerksamkeit entgehe kaum etwas, hingegen schon.

Doch was hat das eigentlich mit dem Lehrberuf zu tun? Unheimlich viel! Was für Konsequenzen auftreten können, wenn wir Lehrerinnen und Lehrer – ob gewollt oder ungewollt – unserer eigenen Aufmerksamkeitsillusion auf den Leim gehen, erläutere ich anhand der folgenden Schilderungen aus der Unterrichtspraxis, welche Bezug nehmen auf mündlich zu erbringende Leistungsnachweise.

F

Fähigkeiten und Fertigkeiten

Während meiner Ausbildung am Sekundarlehramt der Universität Bern in den 1990er Jahren wurde uns Studierenden in den Kursen der angewand-



ten Didaktik (eindringlich) vermittelt, wie wichtig es sei, die «Fähigkeiten und Fertigkeiten» – so lauteten im damaligen Prä-Kompetenzen-Zeitalter die einschlägigen Zauberbegriffe – der Schülerinnen und Schüler konsequent zu fördern und sie auch gebührend zu beachten und zu bewerten. Mit den «Fähigkeiten und Fertigkeiten» waren sowohl Arbeits-, Lern- und Präsentationstechniken als auch schulisch relevante Elemente des Sozialverhaltens gemeint. Folgerichtig wurde uns Studierenden bedeutet, in den Praktika darauf zu achten, immer wieder Lernanlässe zu ermöglichen, die den Jugendlichen dabei helfen sollten, ihr Repertoire an Skills und Techniken kontinuierlich zu verbessern und zu erweitern.

Wer über einen grossen Reigen an Lern- und Arbeitstechniken verfügt, hat grössere Chancen, ganz unterschiedliche Lernsituationen zu meistern und sich so neues Wissen und Können geschickter, kreativer und gezielter anzueignen.

Damit bin ich noch heute durchaus einverstanden: Wer beispielsweise über einen grossen Reigen an Lern- und Arbeitstechniken verfügt, hat grössere Chancen, ganz unterschiedliche Lernsituationen zu meistern und

sich so neues Wissen und Können geschickter, kreativer und gezielter anzueignen. Als junger Lehrer nahm ich mir den überzeugenden Appell meiner Didaktik-Dozierenden deshalb zu Herzen und versuchte, auch bei Leistungsnachweisen aller Art nicht nur das «nackte» fachliche Wissen, sondern eben auch die Fähigkeiten und Fertigkeiten der Lernenden zu fördern und bewertend zu erfassen – und zwar ausgiebig, wie ich während meiner Recherche für den vorliegenden Artikel realisiert habe. Beim Stöbern in meinem privaten Datenarchiv stiess ich zu meinem Entsetzen auf komplizierte Beurteilungsraster, die mehrere A4-Seiten füllten.

Ausladende Raster allenthalben

Mit zunehmender Berufserfahrung musste ich feststellen, dass sich solch ausschweifende Listen nicht als praktikabel erwiesen. Aber – ich gestehe! – auch mein jüngstes Bewertungsraster erachte ich im Lichte meines persönlichen Scheiterns an *sämtlichen* zuvor erwähnten simplen Wahrnehmungstests noch immer als überladen.

Wenigstens befinde ich mich damit in guter Gesellschaft: Die meisten meiner Kolleginnen und Kollegen sind nämlich ebenfalls bemüht, beispielsweise die Qualität eines Schülervortrages mit bis zu 10 Bewertungskategorien in jeweils vier Abstufungen zu beurteilen. Wir Lehrerinnen und Lehrer bil-

den uns folglich allen Ernstes ein, wir seien dazu befähigt, *gleichzeitig* ...

- die Körpersprache (Gestik, Mimik, Körperhaltung, Augenkontakt, Nutzung des Raums),
- die Stimmführung (Lautstärke, Geschwindigkeit, Betonung, Artikulation),
- das Engagement (Ausstrahlung, Überzeugungskraft),
- die sprachliche Gewandtheit (Satzbau, Wortschatz, Grammatik, eigene Formulierungen statt Plagiate),
- den Medieneinsatz (aussagekräftige Folien, anschauliches Wandtafelbild),
- den Aufbau (Klarheit, Stringenz, Schwerpunkte),
- die Verständlichkeit (Einfachheit, Erläuterung von Fachausdrücken, Überleitungen),
- den Vertiefungsgrad (fundierte Kenntnisse)
- und die Effizienz (Zeitlimite, Inhalt)

... zu verfolgen und gar korrekt zu beurteilen – der zuvor beschriebenen *monkey business illusion* zum Trotz.

U

Unangenehme Realitäten

Die traurige Wahrheit indes lautet: Wir können nicht auf die Mimik achten und *zeitgleich* den Variantenreichtum des Satzbaus bewerten. Wir machen uns etwas vor, wenn wir meinen, wir könnten *im selben Moment* die Stimmführung samt Artikulation und Betonung sowie den Anschauungsfaktor einer Grafik erfassen. Es ist illusorisch, zu glauben, wir könnten die Reaktionen in den Gesichtern des Publikums registrieren und *derweil* die Verständlichkeit der Erläuterung eines Fachbegriffes einschätzen.

Wenn wir dies tun respektive zu tun versuchen, dann verhalten wir uns wie Wertungsrichter, die an Tanzturnieren entscheiden müssen, welche der 10 sich auf dem Parkett *simultan* konkurrierenden Tanzpaare in die nächste Runde kommen. Pro Paar verbleiben den Juroren rund 10 Sekunden, in



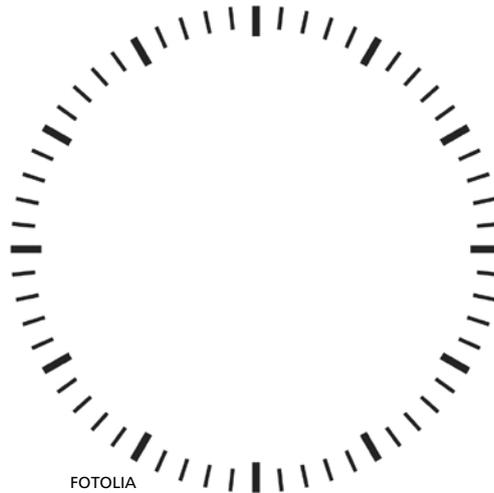
denen sie unter anderem die Choreographie, das Taktgefühl sowie die Ausstrahlung erfassen und beurteilen müssen. Auch wir Lehrpersonen hüpfen mit unseren breitgefächerten Beurteilungsrastern, metaphorisch gesprochen, von einem «Tanzpaar» zum andern, fokussieren und urteilen innert weniger Sekunden, bevor wir uns auf die Suche nach dem nächsten Gesichtspunkt machen – denn was auf der Kriterienliste steht, muss schliesslich abgearbeitet werden.

Natürlich kann man dem Problem bis zu einem gewissen Grad entgegenwirken, indem man den Vortrag aufnimmt und sich ihn ein zweites oder gar drittes Mal ansieht respektive anhört. Wenn man sich dann aber nicht nur auf das Vergeben von Kreuzchen beschränkt, sondern zusätzlich einen Kommentar schreibt, stimmt das Verhältnis von Aufwand und Ertrag definitiv nicht mehr. Nur schon zwei Vorträge pro Schuljahr (zwei Deutsch-Klassen, je ein Vortrag) würden mit diesem Vorgehen locker zu einem Bewertungsaufwand von mehr als 40 Stunden führen. Trotzdem würde sich im Fach Deutsch diese Vortragsbeurteilung in Bezug auf die Zeugnisnote lediglich mit einer Wertigkeit von ungefähr 10% niederschlagen. Verhältnismässig wäre das nicht.

W

Wunsch und Wirklichkeit

Zur Erinnerung: Für die Orientierungsarbeit im Fach Französisch mussten vor wenigen Jahren sämtliche im Kanton Basel-Landschaft zur Verfügung stehenden Fremdsprachenlehrpersonen abgezogen werden, damit die mündlichen Kompetenzen aller Schülerinnen und Schüler der 9. Klassen in einem 15-minütigen Test mit Hilfe eines vergleichsweise kurzen Bewertungsrasters überprüft und bewertet werden konnten. Der enorme organisatorische und personelle Aufwand zog flächendeckend Unterrichtsausfälle und Ersatzstundenpläne mit fachfremder Betreuung nach sich. Und obwohl die Anstrengungen sehr gross



waren, vermochte das Ergebnis niemanden zufriedenzustellen: Die jeweils zwei Examinatoren/-innen mussten sich auf wenige, leicht überschaubare Kriterien beschränken, Zeit für aussagekräftige Kommentare fehlte.

Trotzdem hat diese Form der Beurteilung momentan Hochkonjunktur. So beispielsweise in den obligatorischen Passepartout-Fortbildungskursen, wo den Teilnehmenden immer wieder gepredigt wird, ihren Schülerinnen und Schülern künftig anhand von Bewertungsrastern *regelmässig* formative Feedbacks zu ihren mündlichen Kompetenzen zu geben – *zusätzlich* zu den summativen Beurteilungen, versteht sich. Und dies notabene im Wissen darum, dass die Kursleitungen diese Forderung mangels Ressourcen in ihren eigenen Klassen selbst nicht leisten können.

Man rechne: Ein 15-minütiges, formatives Feedback zu einem Sprech Anlass bei einer Klassengrösse von 24 Schülerinnen und Schüler ergibt – im Idealfall, d.h. ohne krankheitsbedingte Absenzen o.ä.! – 8 Lektionen, in denen der Rest der Klasse selbständig arbeiten muss. Bei fünf Kursen und nur einem einzigen formativen Feedback pro Semester ergibt sich damit ein jährlicher Zeitaufwand von 80 Lektionen. Zum Vergleich: Der Kanton setzte für eine einzige mündliche Orientierungsarbeit den Regelbetrieb an den Schulen ausser Kraft, den Lehrpersonen im Berufsalltag aber scheint man Zauberkräfte zuschreiben zu

Das Forcieren gewohnheitsmässiger formativer Beurteilungen mithilfe komplexer Raster ist faktisch nicht nur ein Problem der beschränkten Wahrnehmung der Bewertenden, sondern schlicht und ergreifend auch der zeitlichen Reserven.

wollen, die es ihnen erlauben, einen niemals versiegenden Quell an Unterrichtszeit zu entdecken. Das Forcieren gewohnheitsmässiger formativer Beurteilungen mithilfe komplexer Raster ist faktisch nicht nur ein Problem der beschränkten Wahrnehmung der Bewertenden, sondern schlicht und ergreifend auch der zeitlichen Reserven.

Ich persönlich kann mich nicht entsinnen, während meiner Schulzeit an der Kanti Solothurn in den 1980er Jahren jemals von einer Lehrperson ein Bewertungsraster erhalten zu haben. Trotzdem wussten wir in der Regel rasch, ob eine Präsentation oder ein anderweitig erbrachter mündlicher Leistungsnachweis gelungen war oder nicht; etwa wenn der Deutschlehrer nach der Rezitation eines Gedichts aufsprang und dem verdutzten Schüler die Hand schüttelte; oder wenn die Augen der Biologielehrerin ob des selbstgebastelten Biberbaumodells aufleuchteten; oder wenn die Französischlehrerin ihr mündliches Feedback hinsichtlich der persönlichen Buchbesprechung mit den Worten «Je vous félicite ...» eröffnete.

D

Die Frage nach dem Warum

Es stellt sich die Frage, warum wir fast alle dennoch immer wieder das Unmögliche versuchen und weiterhin Bewertungskataloge aushecken, die unsere Wahrnehmungsfähigkeit um

das x-fache übersteigen. Ich erkenne dafür drei Hauptgründe:

1. Wie bei so vielem handelt es sich bei den Bewertungsrastern um eine Modeerscheinung: «Alle anderen tun es, also tun wir es auch», gab ein ranghoher Verfechter der Kompetenzorientierung anlässlich eines Lehrplan-21-Podiums unumwunden zu. Corporate Identity also als Begründung für eine Beurteilungsmethode, deren Ungenauigkeit mittels einfachster Experimente aus der Angewandten Psychologie bewiesen werden kann?
2. Die Lehrperson ist bestrebt, die Leistungen der Lernenden gebührend zu würdigen. Sie nimmt ihre Schülerinnen und Schüler ernst, indem sie ihnen transparent und möglichst konkret aufzuzeigen versucht, in welchen Teilbereichen die Lernziele übertroffen, erfüllt oder eben (noch) nicht erfüllt sind. Was eignet sich dazu besser als ein detaillierter Kriterienkatalog, mit Hilfe dessen die Fähigkeiten und Fertigkeiten facettenreich beschrieben werden können?
3. Wie bereits in der ersten Fetisch-Episode zu den schriftlich abgegebenen, detaillierten Lernziellisten⁵ aufgezeigt wurde, dürfte auch in

diesem Fall der Rechtfertigungsdruck gegenüber den Erziehungsberechtigten eine zentrale Triebfeder für die Rasterfalle sein, aus der wir uns ganz offensichtlich nicht zu befreien wagen. Lehrpersonen möchten sich nicht dem Vorwurf ausgesetzt sehen, Bewertungen à la «Handgelenk mal Pi» vorzunehmen, zumal eine wachsende Zahl von Erziehungsberechtigten die Beurteilungen der schulischen Leistungsnachweise ihrer Kinder einer genaueren Prüfung unterzieht als ihre definitive Steuerveranlagung.

S

Schein und Sein

Vorsicht, Sarkasmus: Man decke mit möglichst vielen, akribisch genau formulierten Kriterien die denkbar breiteste Palette von Kompetenzen ab, bestimme ihre Wertigkeit in Form von Punkten und kreuze dann den zutreffenden Qualitätsgrad (gut erkennbar; erkennbar; wenig erkennbar; nicht erkennbar) mit schlafwandlerischer Sicherheit an. Nun müssen nur noch die Punkte zusammengezählt werden, und den Rest erledigt das Notenprogramm, das standardmässig die «magische» 60%-Formel (ein weiterer potenzieller Fetisch für eine spätere Episode) verwendet. Und fertig ist die

einzig wahre, wissenschaftlichen Ansprüchen genügende, mathematisch exakt errechnete Note – auf den Hundertstel genau, wenn es denn sein muss. Bestechend, nicht wahr?

In Wirklichkeit jedoch verhält es sich leider tendenziell genau umgekehrt: *Je mehr Kriterien wir definieren, desto ungenauer wird unsere Bewertung,*

Je mehr Kriterien wir in den Katalog aufnehmen, desto geringer wird die Wahrscheinlichkeit, dass ein Schüler respektive eine Schülerin die Note 6 erreicht.

weil wir erstens kein Multitasking können und zweitens vor lauter Kriterien weder den Gorilla noch das Känguru noch den Ozelot sehen – ganz zu schweigen vom Elefanten im Raum. Und je mehr Kriterien wir in den Katalog aufnehmen, desto geringer wird die Wahrscheinlichkeit, dass ein Schüler respektive eine Schülerin die Note 6 erreicht – weil bei einer zweistelligen Anzahl Faktoren am Ende immer etwas gefunden wird, an dem man noch herumäkeln und die Lernenden zur Selbstoptimierung hindrängen kann.

Wenn dann auch noch schwammig-schwülstige Kompetenzformulierungen – eine Art Pädagogenlyrik – Eingang in die Rasterfalle finden («Die Schülerinnen und Schüler können das Zusammenspiel von Verbalem, Non-verbalem und Paraverbalem zielorientiert einsetzen»⁶), wird die Messbarkeitsillusion vollends auf die Spitze getrieben.

K

King, Ogi und Treasure

Bedenken Sie: Mit bis in den Nanobereich zerteilten und dadurch zeitweise fast neurotisch anmutenden Bewertungsrastern hätte selbst Martin Luther Kings weltberühmte Rede die Maximalpunktzahl verfehlen müs-

Eine wachsende Zahl von Erziehungsberechtigten unterzieht die Beurteilungen der schulischen Leistungsnachweise ihrer Kinder einer genaueren Prüfung als ihre definitive Steuerveranlagung.



sen. Begründung: kein Medieneinsatz, zu wenig Körpersprache, störende Wiederholung der Formulierung «I have a dream»⁷. Alt-Bundesrat Adolf Ogi hingegen hätte für seine legendäre Neujahrsansprache vor dem Lötschbergtunnel⁸ dank seines netten Accessoires – es handelte sich um ein Tannenbäumchen – beim Kriterium «wirkungsvoller Medieneinsatz» mit Sicherheit abgeräumt, wohl aber bei der Intonation, der Sprechmotorik und der Rhetorik wertvolle Punkte eingebüsst.

Sogar der renommierte Geräuschexperte Julian Treasure hätte für seinen höchst sehenswerten TED-Talk «How to speak so that people want to listen»⁹ nicht die Bestnote erreicht. Sprache: genial; Aufbau: exzellent; Intonation: perfekt; Bilderwahl: wunderbar; Einbezug des Publikums: hervorragend; Inhalt: überzeugend; Micro Expressions: absolut synchron. Aber: Entgegen der eindringlichen Empfehlung der meisten Auftrittskompetenz-Gurus versäumte es der Profi, *sich bewusst im Raum zu bewegen*. Stattdessen verharrte er stur auf dem roten Punkt wie unsichere Kandidaten in den sattsam bekannten Castingshows. Schade: erreichte Punktzahl geteilt durch Maximalpunktzahl mal 5 plus 1 = Note 5.85. Und nein, ich kann kein Multitasking. Ich habe mir Treasures TED-Talk einfach dreimal hintereinander angeschaut, hätte ihm aber alleine für den famosen Inhalt und die charismatische Ausstrahlung eine 6 gegeben.

E

Ermessensspielraum statt Punktefixierung

Derartige Zahlenspiele gaukeln eine Genauigkeit vor, die mit keinem Bewertungsraster erreicht werden kann, und sei dieses noch so ausgeklügelt. Mehr noch: Anlässlich meiner ersten Versuche mit «Punkte-Bewertungsrastern» musste ich feststellen, dass die mit der «magischen Formel» berechnete Note nicht selten schlicht nicht der erbrachten Leistung gerecht wur-

de, weil das Unvorhergesehene zu keinem der exakt definierten Kriterien passte und somit aus Gründen der Gerechtigkeit nicht in die Bewertung einfließen konnte. Erfahrene und mit dem Dilemma bestens vertraute Kolleginnen und Kollegen rieten mir damals, entweder die *eigentlich* erreichte Punktzahl bei der nachträglichen Bewertung dergestalt anzupassen,

*Mit bis in den Nano-Bereich
zerteilt und dadurch
zeitweise fast neurotisch
anmutenden*

*Bewertungsrastern hätte selbst
Martin Luther Kings
weltberühmte Rede die
Maximalpunktzahl verfehlen
müssen.*

dass sich mit der dadurch frisierten totalen Punktzahl jene Note ergibt, welche die *tatsächliche* Leistung widerspiegelt – oder aber auf das Vergeben von Punkten ganz zu verzichten.

Ich entschied mich für eine aus meiner Sicht «ehrliche» Variante: abgespeckte Raster ohne Fixierung auf starre Punktzahlen, stattdessen hin zu mehr Ermessensspielraum und flexiblerer Gewichtung. Dieses Vorgehen erlaubt es mir seither, aussergewöhnliche Fähigkeiten – eine packende Passage, eine *native-speaker*-ähnliche Aussprache, eine bestechende Eloquenz, ein verblüffendes Experiment, eine ansteckende Lebendigkeit – besonders zu würdigen. Dabei verlieren allfällige Unzulänglichkeiten in anderen Bereichen an Bedeutung, treten in den Hintergrund und lassen berechtigterweise dem Besonderen, dem Genialen, dem Berührenden den Vortritt.

Individuelle Feedbacks in Worten

Bei dieser Form der Beurteilung kann es folglich passieren, dass Lernende auf ihren nunmehr reduzierten Bewertungsrastern die Kreuzchen zwar am selben Ort stehen haben, aber

trotzdem nicht die gleiche Note erhalten. Damit die Schülerinnen und Schüler dies nicht als Ungerechtigkeit empfinden, ist es natürlich wichtig, sie vor den Vorträgen oder mündlichen Tests transparent über die Art der Bewertung zu informieren und sie von den Vorteilen der Kombination aus vorgegebenen Rastern und individuellen Rückmeldungen zu überzeugen. Für mich steht ausser Zweifel, dass bewertende Feedbacks dieser Art auch mit den vordergründig durchdachtesten Bewertungsrastern nicht adäquat einzufangen sind.

Einem besonders begabten Redner in meiner Klasse teilte ich einst mit: «*Du wirkst präsent und locker zugleich, hast die Lage jederzeit im Griff und führst die Zuhörenden durch dein Programm wie ein Profi. Die humorvollen Einlagen, gepaart mit den fundierten Ausführungen, hinterlassen einen bleibenden Eindruck. Ist eine Karriere als Moderator geplant?*» Jahre später traf ich ebendiesen (mittlerweile ehemaligen) Schüler zufälligerweise an einer Grillstelle. Unglücklicherweise war mir sein Name entfallen, weshalb ich ihn danach fragen musste. Seine Freunde spöttelten: «*Die Namen ihrer besten Schüler vergessen Lehrer aber normalerweise nicht.*» Ich konterte sofort: «*Ich konnte mich zwar nicht mehr an den Namen erinnern, an den bestechenden Vortrag zum Thema «Atomare Bedrohung» aber umso besser.*» Der ehemalige Schüler lächelte, während seine Freunde staunten.

Einer anderen Schülerin, einer Secunda, die in ihrem Vortrag «Siamo italiani», die bewegende Geschichte italienischer Gastarbeiterfamilien in der Schweiz geschildert hatte, schrieb ich: «*Deine persönliche Auseinandersetzung mit der Geschichte deiner Familie hat mich berührt. Es ist dir gelungen, eine dunklen Teil der jüngsten Schweizergeschichte eindrücklich zu beleuchten.*» Und einen Schüler, der über das Leben in der DDR berichtet hatte und dem es dabei souverän gelungen war, die Erfahrungen seiner Eltern subtil einzubauen, liess ich wissen: «*Du erzählst ausserordentlich lebendig; besonders gefallen hat mir die*

persönliche Note. Es ist dir damit gelungen, die Klasse in deinen Bann zu ziehen. Das machte die sprachlichen Ungenauigkeiten mehr als wett.»

Drei konkrete Beispiele zur Veranschaulichung der gleichen Aussage: Solche individuellen Rückmeldungen sind für die Schülerinnen und Schüler eindeutig wertvoller als eine nackte Punktzahl in der fiktiven Kategorie «autobiographische Elemente» – ganz abgesehen davon, dass Authentizität nun wirklich nicht als verbindliches Kriterium in einem Bewertungsraster taugen würde. Darüber hinaus stellen solche Feedbacks hervorragende Elemente einer funktionierenden Beziehungsebene zwischen Lehrkraft und Schüler/Schülerin dar.

E

Ein gangbarer Weg – wie lange noch?

Mit der Kombination aus individuellen Rückmeldungen einerseits und der flexiblen Handhabung «vernünftiger»

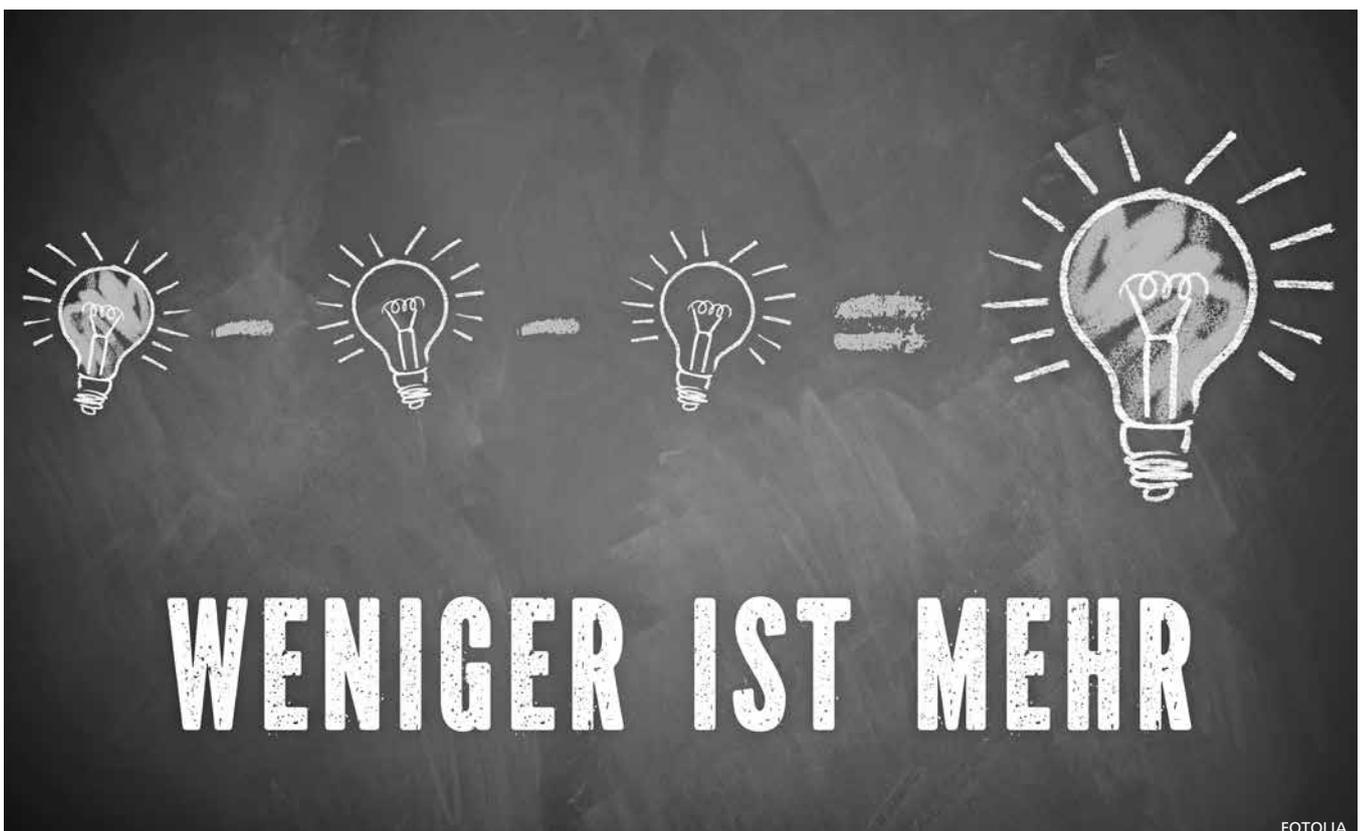
Bewertungsraster andererseits habe ich bis anhin gute Erfahrungen gemacht. Nur selten wurde die Bewertung seitens Eltern oder Lernender angezweifelt. Und selbst dann fand sich meistens eine Lösung, mit der alle Beteiligten gut leben konnten.

Die Beurteilung mündlicher Leistungen ist keine exakte Wissenschaft. Da braucht einem auch kein Zacken aus der Krone zu fallen, wenn man als Lehrperson aufgrund einer gut begründeten Stellungnahme zur Einsicht gelangt, dass die Note durchaus etwas besser hätte ausfallen können. *«Ich kann gut verstehen, dass dich das Resultat in Anbetracht der investierten Zeit enttäuscht hat»*, schrieb ich einst einem Schüler. *«Es ist mir bewusst, dass eine Vortragsnote trotz der definierten Bewertungskriterien und meiner differenzierten Rückmeldungen (z.B. «Das ist dir besonders gut gelungen; hier kannst/musst du dich noch verbessern») nie unantastbar sein kann. Als Folge deiner sauberen Argumentation habe ich deinen Vortrag noch einmal studiert [...]. Ich erhöhe die Note auf 4.5 und hoffe, dass sich*

diese Anpassung positiv auf deine Zeugnisnote auswirken wird.»

Zugegeben: Solche Diskussionen sind zeitintensiv und können nervenaufreibend sein. Sie können sich im besten Fall aber auch ausgesprochen lohnen. Allein schon das Gefühl des Ernstgenommenwerdens auf Schülerseite ist von unschätzbarem Wert. Der betroffene Schüler jedenfalls arbeitete von diesem Tag an merklich motivierter, probierte aufgrund meiner konkreten Hinweise plötzlich neue Lerntechniken aus und schaffte in der Folge den Sprung ans Gymnasium mit Bravour.

Dominoeffekte sind wider Erwarten ausgeblieben. Nur weil ich ab und zu einem Schüler oder einer Schülerin hinsichtlich der Bewertung nachträglich ein Stück entgegenkomme, stehen die Klassenkameraden und -kameradinnen deswegen nicht Schlange, in der Hoffnung, ihre Noten würden, einem Automatismus gleich, ebenfalls angehoben. Ich bin davon überzeugt, dass die allermeisten Schülerinnen und Schüler ein feines Senso-



rium entwickeln, dank dem sie sehr wohl beurteilen können, ob eine Bewertung nachvollziehbar, stimmig und fair ist oder nicht.

Ich bin davon überzeugt, dass die allermeisten Schülerinnen und Schüler ein feines Sensorium entwickeln, dank dem sie sehr wohl beurteilen können, ob eine Bewertung nachvollziehbar, stimmig und fair ist oder nicht.

Es steht allerdings zu befürchten, dass die stetig wachsende Population der sogenannten Helikoptereltern, der chronisch überfüllte Berufsauftrag von uns Lehrpersonen sowie die kapillarenähnlich verästelten Kompetenzraster, die in den Schubladen der Schreibtischpädagogen seit längerem ihrer epidemischen Verbreitung har-

ren, einer gleichwohl anspruchsvollen wie pragmatischen Beurteilungspraxis mittelfristig den Garaus machen werden.

W

Weniger ist mehr

Wie dem auch sei; da ich im Rahmen des *Monkey-Business-Aufmerksamkeits-Tests* lediglich beim Zählen der gespielten Pässe hatte punkten können, habe ich beschlossen, meine Beurteilungspraxis erneut anzupassen. Die Bewertungsraster werden künftig eher den Stellenwert eines Vorbereitungsinstruments erhalten, das den Schülerinnen und Schülern dabei helfen soll, *ausgewählte* Bereiche gezielt zu trainieren. Wer sich Treasures Referat zu Gemüte geführt hat oder dies noch tun wird, weiss, dass nur schon die Optimierung der Stimmführung

eine lohnenswerte, zugleich aber auch zeitintensive Herausforderung darstellt.

Als Lehrer, zu dessen Aufgaben es gehört, transparent, fair und effizient zu beurteilen, habe ich mir auf die Fahne geschrieben, mich von der Aufmerksamkeitsillusion – im wahrsten Sinne des Wortes – nicht länger zum Affen machen zu lassen und deshalb meine ausgeklügelten Bewertungsbögen *klüger* zu gestalten, sprich: auszudünnen. Dies in der Hoffnung, dass sich in meinem Schulzimmer schon bald nicht nur Gorillas, sondern weitere ungeahnte Tierarten tummeln werden. Bestsellerautor Rolf Dobelli, dessen Bücher mich zu diesem Artikel inspiriert haben, formuliert es so: «Etwas Überraschendes kann noch so gross und anders sein, wir sehen es nicht. Gross und anders zu sein genügt nicht. Es muss erwartet werden.»¹⁰

Links zu den im Artikel erwähnten Clips

¹ «You can't even text and walk, so why do you text and drive?»
<https://www.youtube.com/watch?v=cyGUzjMQj98>

Aufmerksamkeitstests

^{2,3} The Monkey-Business-Illusion by Daniel Simons
<https://www.youtube.com/watch?v=vJG698U2Mvo>
<http://www.theinvisiblegorilla.com/videos.html>

King, Ogi, Treasure

⁷ Martin Luther King, «I have a dream»
<https://www.youtube.com/watch?v=3vDWWy4CMhE>
⁸ Adolf Ogi, Neujahrsansprache vor dem Lötschbergtunnel
https://www.limmattalerzeitung.ch/mediathek/videos/1_uw4130li
⁹ Julian Treasure, ««How to speak so that people want to listen»»
https://www.ted.com/talks/julian_treasure_how_to_speak_so_that_people_want_to_listen/transcript#t-6509

⁴ http://www.sportwissenschaft.rub.de/mam/spomed/mitarbeitende/Jendrusch/ak_20_jendrusch__aufmerksamkeit-druckversion_.pdf

⁵ siehe lvb.inform 2017/18-03

⁶ Lehrplan Baselland, D.3.A.1

¹⁰ Rolf Dobelli, Kunst des klugen Handelns, Warum Sie gelegentlich am Brennpunkt vorbeischaun sollten, S. 175