

Kurt Vogelsberger

Appelle eines Mathematikers gegen die Mathematik

... bei der Notengebung und der Ermittlung von Zeugnisnoten

- *Ein „Blütenmeer“*
- *Gewichtung von erstem zu zweitem Halbjahr*
- *Die angeblich so genaue Mathematik ist problematischer als man denkt ...*
- *Ist doch sonnenklar: „3.46 ist 3!“*
- *„Gerecht ist: 3 Minus ist gleich 3.25!“*
- *„Eine Epochalnote zählt doch mehr als ...“*
- *Das Heft nicht aus der Hand geben!*

Ein „Blütenmeer“

Dieser Tage, es war gegen Ende des Schuljahres 1995/96, häuften sich - wie regelmäßig alle Jahre wieder - die Diskussionen um Praktiken der Notengebung, um Techniken zur „Berechnung“ der Zeugnisnoten und um Verfahrensweisen zur Ermittlung der „Kopfnote“. Kolleginnen und Kollegen reden sich die Köpfe heiß, welche Mechanismen denn die rechten seien, und inwieweit Konventionen oder gar Konferenzbeschlüsse hierzu denn verbindlich sein könnten; Realschullehreranwärterinnen und -anwärter tragen im Studienseminar Erfahrungen und Beobachtungen aus Konferenzen vor und fragen nach, weil sie Dissensen entdecken zu dem, was in der Seminararbeit als sinnvoll vorgestellt und propagiert wird.

Dabei wurde mir schmerzlich deutlich, wie sehr doch Mathematik und Numerik in vielen Bereichen mißbraucht werden, um Noten und Zeugnisnoten ein Gewand zu geben, das ihnen im Grunde nicht zukommt. Vielerorts geht man in der Schulpraxis von der falschen Annahme aus, das Problem gerechter Bewertungen sei mit Hilfe der Mathematik zu lösen, und zugleich muß die Mathematik dafür herhalten, um über die

Praxis der Notengebung den Mantel der Unangreifbarkeit zu legen. Wer würde sich denn trauen, bei einem nachgewiesenen Notendurchschnitt von 3,23 in Zweifel zu ziehen, daß das die Gesamtnote „befriedigend“ ergibt? (Nebenbei bemerkt: ich sehr wohl!)

Es ist meines Erachtens an der Zeit, der Mathematik zu Leibe zu rücken, die Pädagogik an ihre Stelle zu setzen und mathematische Strategien auf das zu reduzieren, was sie in diesem Zusammenhang leisten können - und das ist erheblich weniger, als viele denken und auch praktizieren.

Im Zuge der Auseinandersetzung mit dem zugegebenermaßen immens schwierigen Problem gerechter Bewertungen und fundierter Zeugnisnoten wurden viele Ideen und Strategien oder einfach nur Rezepte entwickelt, fanden Anhänger, wurden verbreitet, gewannen den Status dogmatischer Lehrsätze, an denen kaum jemand mehr zu rühren sich traut. In dieser Schule gedeiht so via Konvention eine „Blüte“, an jener Schule eine andere, und inzwischen ist die gesamte Schullandschaft ein wahres Blütenmeer, wuchernd zum gleichsam „mathematischen Urwald“, in dem die Pädagogik kaum noch Licht und Luft findet.

So sagte letzt eine Anwärterin zu mir: „An meiner Schule ist bei einem Test die Grenze zwischen 4 und 5 bei 40% festgelegt“.

(Anmerkung: Mit „Test“ meinte sie wohl ein schriftliches Abfragen der Hausaufgabe, eine schriftliche Überprüfung resp. eine Klassenarbeit, andere Formen kennt und legitimiert unsere Schulordnung nicht - und trotzdem gibt es Schulen mit „Extemporale“ oder „3-Stunden-Test“ oder sonstigen exotischen Auswüchsen schriftlicher Leistungskontrollen, womit wir wieder beim „Blütenmeer“ wären, in diesem Fall sogar den „Mohnblüten“, weil im Grunde nicht statthaft.)

Diese Festlegung ist, ebenso wie andere ähnlicher Couleur, sinnig wie unsinnig zugleich. Denn nirgendwo steht geschrieben, daß dieser Wert 40% das Maß aller Dinge sei - eben eine „Blüte“.

Damit ich nicht mißverstanden werde: es erscheint durchaus sinnvoll, im Sinne der Vergleichbarkeit Bewertungsvorstellungen auszutauschen und gewisse Konventionen und Abstimmungen zu treffen, nur muß ihnen eben der rechte Stellenwert zugemessen werden als ungefährer und zugleich unverbindlicher Orientierungsrahmen.

Zwischen der zentralen Grundaussage der Schulordnung, Leistungsfeststellung und -beurteilung seien durch die pädagogische Verantwortung und Freiheit des Lehrers bestimmt [§ 45 (1)], und der mehr peripheren Auflage, der Schulleiter habe auf die Koordination der Notengebung zu achten [§ 56 (1)], ist ein ausgewogenes Verhältnis herzustellen.

Gewichtung von erstem zu zweitem Halbjahr

Ich möchte den relativen Mißbrauch der Mathematik an einem ersten Beispiel verifizieren, nämlich der Gewichtung von erstem zu zweitem Halbjahr im Zuge der Festsetzung (ausdrücklich nicht: Berechnung) der Zeugnisnoten und der Vorgabe der Schulordnung, daß die Leistungen des zweiten Halbjahres stärker zu berücksichtigen (ausdrücklich nicht: gewichten) seien. Die berechtigten Fragen, wie denn die Schulordnung dies meine und welches Verhältnis für die Berechnung zugrunde zu legen sei, finden natürlich sehr viele unterschiedliche Antworten, von 1 : 2 über 2 : 3 bis hin zu 3 : 5 oder 4 : 7 oder gar 1 : 3.

(Anmerkung: Die Frage, welches dieser Zahlenverhältnisse denn dem zweiten Halbjahr das stärkste Gewicht beimesse, wußten in einem Seminar nicht alle zu beantworten - was ebenfalls zu denken geben sollte.)

Bei dieser Diskussion gerät völlig aus dem Blick, daß die Schulordnung diese Auflage nicht mathematisch meint, sondern wohl in erster Linie pädagogisch! Daß nämlich Leistungssteigerungen bzw. -abfälle einzubeziehen sind und daß

die aus der „Zwischeninformation“ in Form der Halbjahresnote resultierenden Anstrengungen besonders honoriert werden sollten. Und in der Ausformung dessen hat jeder Beurteiler unter Ausschöpfung seines ihm (glücklicherweise) zugebilligten Ermessens- und Beurteilungsspielraums eigene Entscheidungen zu treffen.

Und die (für das Ergebnis sehr gewichtige, vgl. unten) Frage ist, ob jede einzelne Note des zweiten Halbjahres ein höheres Gewicht bekommen soll oder nur der Durchschnitt des zweiten Halbjahres.

Ein Kollege berichtete jüngst, daß es Konferenzbeschuß an seiner Schule sei, diese Gewichtung im Verhältnis 1/3 zu 2/3 vorzunehmen, daß er dies aber nicht für sinnvoll halte und er sich deswegen nicht an den Konferenzbeschuß gebunden fühle und anders berechnen werde. Dieser Konflikt offenbart doch einige eklatante Fehleinschätzungen.

Erstens gibt es absolut kein einziges sachliches Argument, weswegen 1/3 zu 2/3 (also 1 : 2) nicht sinnvoll ist, und weswegen 2/5 zu 3/5 (also 2 : 3) oder auch ein anderes Verhältnis sinnvoller sein soll. Im Grunde ist, weil anheimgestellt, jedes beliebige Verhältnis $m : n$ ($m < n$) sinnvoll oder , auch sinnlos. Denn wie gesagt: diese Auflage ist nicht als in erster Linie mathematisches Konstrukt zu verstehen.

Zweitens hätte an der Schule ein solcher Konferenzbeschuß gar nicht erst gefaßt werden dürfen, denn er stellt tatsächlich eine im Grunde unzulässige Einschränkung des vom Gesetzgeber ausdrücklich eingeräumten pädagogischen Freiraums dar.

Drittens jedoch steht es dem Kollegen nicht zu, sich über Konferenzbeschlüsse, die ohne Wenn und Aber verbindlich sind, unter Berufung auf seinen Freiraum einfach hinwegzusetzen. Er müßte sich stattdessen darum bemühen, dies auf die Tagesordnung einer Gesamtkonferenz zu bringen mit dem Ziel, daß der verbindliche Beschuß in eine Empfehlung umgewandelt wird.

Die angeblich so genaue Mathematik ist problematischer als man denkt ...

Das mathematische Modell, das ein bestimmtes Verhältnis zur „Notenberechnung“ verwendet,

bedarf einer genaueren mathematischen Betrachtung.

Angenommen, es soll die Durchschnittsnote aus fünf Klassenarbeiten (Zahlenbeispiel: 5, 4, 3, 2, 3) berechnet werden, wovon zwei im ersten und drei im zweiten Halbjahr geschrieben wurden.

• Jahresdurchschnitt

Berechnet man den Durchschnitt ohne Gewichtung aus $(5+4+3+2+3) : 5 = 3.4$, so geht jede Klassenarbeit gleichgewichtig in das Ergebnis ein.

• Halbjahresdurchschnitte

Berechnet man ihn jedoch, was viele Kolleginnen und Kollegen wohl tun, getrennt nach Halbjahren $(5+4) : 2 = 4.5$ und $(3+2+3) : 3 = 2.7$ und dann den Gesamtschnitt aus

$$(4.5+2.7) : 2 = 3.6,$$

dann gehen die beiden Noten des ersten Halbjahres mit je 25%, die drei des zweiten Halbjahres nur mit je 16,7% in den Gesamtdurchschnitt ein, d.h. rein aus der Tatsache der unterschiedlichen Notenanzahl resultieren bei diesem Berechnungsmodus bereits unterschiedliche Gewichtungen, und zwar eine geringere der Noten des zweiten Halbjahres - ein mathematisches Faktum, was wohl so manchen Beurteilern und „Rechenstrategen“ nicht bewußt ist und verborgen bleibt.

(Wurden die drei Klassenarbeiten im ersten und nur zwei im zweiten Halbjahr geschrieben, dann sind die Quoten gerade umgekehrt, es ergibt sich gleichsam automatisch eine stärkere Gewichtung aufgrund der geringeren Anzahl von Noten.)

„Rechenkünstler“, welche jetzt die Halbjahresdurchschnitte 2 : 3 gewichten, also einen Gesamtdurchschnitt mit

$$(4.5 \cdot 2 + 2.7 \cdot 3) : 5 = 3.42$$

berechnen, haben damit jetzt jede einzelne Note zu 20% eingebracht, also mit der Gewichtung 2 zu 3 das „Unverhältnis“ 25% zu 16,7% gerade eben rückgängig gemacht und wieder zum absoluten Gleichgewicht korrigiert. Um eine tatsächlich stärkere Gewichtung der Noten des zweiten Halbjahres zu erzielen, müßte also eine noch stärkere Quote für das zweite Halbjahr gewählt werden, z.B. 1 : 2. Dann ergibt sich ein Schnitt von

$$(4.5 \cdot 1 + 2.7 \cdot 2) : 3 = 3.3,$$

und die Noten des ersten Halbjahres sind mit 16,7% und die des zweiten mit 22,2% eingegangen, also im Verhältnis von tatsächlich „nur“ 1 zu 1.3, also etwa 3 : 4 statt der vermeintlichen 1 zu 2.

Da insbesondere im Bereich der anderen Leistungsnachweise in der Regel aber von unterschiedlichen Anzahlen von Noten in beiden Halbjahren ausgegangen werden muß, ist – als erstes gewichtiges Fazit - von einer Berechnung des Gesamtdurchschnitts aus zwei Halbjahresdurchschnitten dringendst abzuraten. Denn dann entstehen durch die differierenden Notenanzahlen die abenteuerlichsten undurchsichtigen Gewichtungsquoten.

Die „Missverhältnisse“ werden weiter potenziert, das „Gewichtungsgestrüpp“ wird absolut undurchsichtig, wenn, wie im Bereich der Klassenarbeiten eher unüblich, im Bereich der anderen Leistungsnachweise aber durchaus Usus, jetzt auch noch Einzelnoten unterschiedlich gewichtet werden sollen.

Die weitaus bessere Technik, die einzelnen Noten des gesamten Schuljahres zu einer Durchschnittsnote aufzurechnen, entspricht auch eher dem „Geist“ der Schulordnung, wo dem Halbjahreszeugnis lediglich der Stellenwert einer „Zwischeninformation“ zukommt, und wonach die Jahreszeugnisnote die Leistungen des gesamten Schuljahres einbeziehen soll.

Zur Entscheidung über die Gesamtnoten der jeweiligen Leistungsbereiche (Klassenarbeiten resp. andere Leistungsnachweise) sollte aus allen Noten des gesamten Schuljahres ein Durchschnittswert ermittelt werden. Halbjahresdurchschnitte sollten nicht berechnet und keinesfalls weiterverarbeitet werden.

In vielen Fällen (so z.B. im obigen Notenbeispiel 5, 4, 3, 2, 3 mit dem ungewichteten Durchschnitt von 3.4) erübrigt sich eine irgendwie geartete Gewichtung, weil das Notenbild mit den eindeutig besseren Leistungen im zweiten Halbjahr die Entscheidung für die Gesamtnote „befriedigend“ erzwingt. Und jedwede Numerik mit einer stär-

keren Gewichtung der Noten des 2. Halbjahres führte zu einem Schnitt besser als 3.4! Also wozu in diesem Fall einen Taschenrechner oder gar eine Tabellenkalkulation bemühen?

Die Berechnung eines gewichteten Durchschnitts ist nur bei „undurchsichtigeren“ Notenverteilungen erforderlich, wobei jede einzelne Note mit einem Faktor versehen wird.

Beispiel 1:

	1. Halbjahr	2. Halbjahr
Klassenarbeiten	3, 4	2, 5, 4

Gewichtung	Rechnung	Ø
1 : 1 (=1.0)	$(3 \cdot 1 + 4 \cdot 1 + 2 \cdot 1 + 5 \cdot 1 + 4 \cdot 1) : 5$	3.60
4 : 5 (=0.8)	$(3 \cdot 4 + 4 \cdot 4 + 2 \cdot 5 + 5 \cdot 5 + 4 \cdot 5) : 23$	3.61
2 : 3 (≈0.7)	$(3 \cdot 2 + 4 \cdot 2 + 2 \cdot 3 + 5 \cdot 3 + 4 \cdot 3) : 13$	3.62
3 : 5 (=0.6)	$(3 \cdot 3 + 4 \cdot 3 + 2 \cdot 5 + 5 \cdot 5 + 4 \cdot 5) : 21$	3.62
4 : 7 (≈0.6)	$(3 \cdot 4 + 4 \cdot 4 + 2 \cdot 7 + 5 \cdot 7 + 4 \cdot 7) : 29$	3.62
1 : 2 (=0.5)	$(3 \cdot 1 + 4 \cdot 1 + 2 \cdot 2 + 5 \cdot 2 + 4 \cdot 2) : 8$	3.63
1 : 3 (≈0.3)	$(3 \cdot 1 + 4 \cdot 1 + 2 \cdot 3 + 5 \cdot 3 + 4 \cdot 3) : 11$	3.64

Spätestens bei diesem Rechenbeispiel muß auch dem hartnäckigsten Verfechter mathematischer Zahlenspielerien deutlich werden, daß es mehr als unsinnig ist, über Gewichtungsquoten zu diskutieren. Selbst die extreme Spannweite von 1 zu 1 (gleichgewichtig) bis 1 zu 3 (2. Halbjahr 3fach) liefert im Ergebnis nur Abweichungen von vier Hundertstel (und lediglich bei stärkeren Leistungsunterschieden zwischen 1. und 2. Halbjahr werden diese Abweichungen größer sein).

Die Berechnung eines gewichteten Durchschnitts erbringt also in den meisten Fällen kaum Entscheidungshilfen.

Anmerkung: Im obigen Beispiel wäre es fernab der Mathematik eher angebracht, über den extremen Sprung (2 auf 5) nachzudenken resp. über die etwas „aus der Reihe tanzende« Zwei; denn mögliche Ursachen liefern weitere Entscheidungshilfen pädagogischer Provenienz.

Zur Festsetzung der Gesamtnote ist es sinnvoll, sich zunächst am einfachen Durchschnitt zu orientieren. Dem kann in den wenigen verbleibenden „Zweifelsfällen“ eine Extremgewichtung (z.B. 1 zu 2) gegenübergestellt werden, um festzustellen, welche Veränderungen sich damit ergeben, ob sie unerheblich oder erheblich sind und ob sie demzufolge nach Maßgabe der Schulordnung in die Überlegungen und die Entscheidung für die Gesamtnote einbezogen werden müssen.

Beispiele 2:

a)	1. Halbjahr	2. Halbjahr	Ø 1 : 1	Ø 1 : 2
	2, 2	3, 5, 4	3.2	3.5

Hier wäre jenseits der Numerik zu bedenken, wie der extreme Leistungsabfall im 2. Halbjahr zustande kommt. Aus Gründen, die einzig der Schüler oder die Schülerin zu verantworten hat? Wenn nicht, dann kommt eine Festsetzung auf 4 wohl kaum in Betracht. Auch wäre einzubeziehen, wie die Halbjahresnote ausfiel und ob ein Sprung von 2 Noten sachlich begründet werden kann.

b)	1. Halbjahr	2. Halbjahr	Ø 1 : 1	Ø 1 : 2
	3, 3	4, 4, 3	3.4	3.5
c)	1. Halbjahr	2. Halbjahr	Ø 1 : 1	Ø 1 : 2
	3, 3, 4	4, 3	3.4	3.4

Hier ist dieselbe Notenfolge unterschiedlich auf die Halbjahre verteilt. Aus der Tatsache, daß die dritte Klassenarbeit dem sinnvollen Arbeitsplan der Lehrkraft folgend „zufällig“ Ende Januar oder Anfang Februar gefertigt wurde, dürfen doch wohl keine Nach- oder Vorteile erwachsen?

d)	1. Halbjahr	2. Halbjahr	Ø 1 : 1	Ø 1 : 2
	4, 3	3, 2, 2	2.8	2.6
e)	1. Halbjahr	2. Halbjahr	Ø 1 : 1	Ø 1 : 2
	2, 3	2, 3, 4	2.8	2.9

In Beispiel d) und e) wurden dieselben Noten erzielt, lediglich in unterschiedlicher Abfolge. Die Numerik spricht in Beispiel e) für „befriedigend“, in Beispiel d) sollte wegen der eindeutig positiven Tendenz und der besseren Leistungen des 2. Halbjahres die Gesamtnote „gut“ in Erwägung gezogen werden.

(Anmerkung: Man berichtete mir, daß zur rechnerischen Ermittlung der Jahreszeugnisnote auch Halbjahreszeugnisnoten oder Gesamtnoten des ersten Halbjahres herangezogen werden - womit wir wieder bei den „Mohnblüten“ wären. Denn diese Noten sind im allgemeinen nivellierte Setzungen und dürfen keinesfalls weiterverrechnet werden.)

Ist doch sonnenklar: „3.46 ist 3!“

Bereits in den vorstehenden Beispielen und den jeweiligen Argumentationen wurde deutlich, welcher Stellenwert der Mathematik meines Erachtens zukommen darf: Keinesfalls ein leitender, alles determinierender. Numerische Betrachtungen dürfen lediglich eine dienende Hilfsfunktion haben, um zu pädagogisch verantworteten Entscheidungen zu gelangen.

Die „Flucht“ in die Mathematik wird doch nur deshalb angetreten, weil ihr der Touch der absoluten Exaktheit anhaftet; denn wer wird schon vermuten wollen, daß ein solch exaktes, reales Konstrukt wie 3.46 etwas eher Ungenaues und Irrales sein könnte? Und solche angeblich exakten Zahlen scheinen fast unangreifbar, und sie sind eher justitiabel als pädagogische Entscheidungen, die von den Betroffenen (Schülerinnen und Schüler, Eltern) und Verwaltungsgerichten womöglich in Frage gestellt werden könnten, eine Gefahr, der man mit Zahlen weitaus weniger anheim fällt. Unzulänglichkeiten in unserer Leistungsbewertung - zu denen wir offen stehen sollten, denn sie liegen vielfach in dem schwierigen Phänomen und weniger in unseren mangelnden Kompetenzen begründet - werden mit dem Deckmantel der Mathematik kaschiert. Sobald die Schülerleistungen als Zahlen in unserem Notenbuch gelandet sind, wird mit ihnen hemmungslos mathematisch operiert und argumentiert.

Also: 3.46 ist eben nicht per se gleich 3!

Erstens: Die von einer Lehrkraft erteilten Noten kommen, auch bei noch so intensiven Anstrengungen zur Qualitätssicherung, häufig über den Charakter einer „Schätzung“ nicht hinaus. Dabei spielt es kaum eine Rolle, ob es sich um Noten aus schriftlichen Leistungskontrollen handelt oder um Epochalnoten oder Noten aus mündlichen Überprüfungsformen, bei denen sich naturgemäß sehr viel mehr Unsicherheitsfaktoren

einschleichen. Und aus Schätzwerten exakte Durchschnitte zu bilden, auch noch auf zwei Stellen genau, ist mathematisch gesehen schlicht und einfach falsch und eine pädagogische Perversion.

Zweitens: Selbst Durchschnitte aus gleichen Leistungsformen wie z.B. Klassenarbeiten zu bilden, ist mathematisch äußerst problematisch und fragwürdig, weil von Arbeit zu Arbeit häufig die Meßplatten wie Hemden gewechselt werden. Eine Durchschnittsbildung ist aber nur dann legitim und ergibt eine zuverlässige Aussage über die „durchschnittliche Leistung“, wenn die Meßplatten einigermaßen identisch waren.

Drittens: Noch sehr viel problematischer wird die Sache, wenn im Bereich der anderen Leistungsnachweise Noten aus den verschiedensten Ecken (schriftliches Abfragen der Hausaufgabe, Epochalnoten, Referate, Bewertungen von Heftführung, mündliche Überprüfungen u.a.m.) miteinander „verwurschtelt“ werden. Der Durchschnittswert sagt alles mögliche aus, leistet aber kaum das was er soll, nämlich ein zuverlässiges Indiz für die „Durchschnittsleistung“ zu sein.

Daraus folgt, daß 3.46 eben nicht automatisch gleich 3 ist. Daraus folgt weiterhin, daß es pädagogischer Nonsens ist und besser heute als morgen untersagt werden müßte, Durchschnitte auf zwei Dezimalen „genau“ zu berechnen. Daraus folgt ebenso, daß $3.4 = 4$ sein kann und $3.3 = 4$ und $3.5 = 3$ und $3.6 = 3$ und $3.7 = 3$, und womöglich sind die Spannen des Ermessens noch weiter zu setzen. Allerdings nur dann, wenn der Durchschnitt als Hilfsgröße unter vielen weiteren Entscheidungskomponenten verstanden wird. Bildet er die einzige Entscheidungsinstanz, dann bleibt wirklich nur, sich auf die „nackte“ Mathematik zu reduzieren, und dann ist ab 3.5 tatsächlich der Ofen aus, der Ermessensspielraum konvergiert gegen Null.

Der Notendurchschnitt ist nicht als absolut bindende Instanz, sondern lediglich als Entscheidungshilfe zu verstehen. Die Entscheidung über die Gesamtnote ist nur in Verbindung mit anderen Kriterien (z.B. Notentrend, Tendenzzzeichen, Entwicklung, individuelle Prämissen [aber nicht: Disziplin] etc.) zu treffen.

„Gerecht ist: 3 Minus ist gleich 3.25!“

Es ist wohl allgemein übliche Praxis, Noten mit Tendenzzeichen zu versehen. Die Motive hierfür sind einerseits lauter (weil sachlich begründet), wenn sie nämlich gesetzt werden, um dem Schüler bzw. der Schülerin die Nähe zur besseren oder schlechteren Note zu signalisieren. Sie sind eher unlauter, wenn sie als „Beruhigungsspillen“ rangieren; hatte man auf eine 2 gehofft, so lebt sich's mit der 3+ besser als mit der 3, und eine 4– schluckt man eher als eine 5. Letztere Motive sind besonders häufig in jenen Bereichen anzutreffen, wo die Bewertung schwieriger und unwägbarer ist, also z.B. in mündlichen Leistungskontrollen.

Die Frage ist, wie diese Tendenzzeichen bei der Ermittlung der Gesamtnote zu verwerthen sind. Und mancherorts hat sich, um die anfängliche Metapher aufzunehmen, die „Blüte“ entwickelt, 3+ als 2.75 oder 2.8 und 5– als 5.25 oder 5.2 in die Berechnung des Durchschnitts einfließen zu lassen, mit der Intention, daß sich die Tendenzzeichen auswirken, die Probanden etwas davon haben sollen.

So hehr diese Absicht auch ist, und auch wenn diese Praxis von den Betroffenen im Schnitt wohl begrüßt wird, das ist ein unzulässiges und zudem auch unsinniges Verfahren.

Zum einen steht diese Praxis eindeutig im Widerspruch zu den Bestimmungen der Schulordnung, die ein sechsstufiges Notensystem zwingend vorschreibt. Tendenzzeichen so zu verwerthen, hebt diese Auflage aus und schafft ein 16stufiges (oder gar 18stufiges, wenn 1+ und 6– erteilt und mit 0.75 bzw. 6.25 gerechnet werden, was sich meiner Kenntnis entzieht).

Des weiteren gibt es einschlägige VerwaltungsgerichtsUrteile, welche eine solche Praxis als ordnungswidrig einstufen und zur Auflage machen, daß Tendenzzeichen nicht verrechnet, sondern lediglich als Entscheidungshilfe bei der Festsetzung der Gesamtnote herangezogen werden dürfen.

(Anmerkung: Gerichtsurteile sind mitunter pädagogisch weiser als man denkt!)

Und zuletzt tritt wiederum die Mathematik an die Stelle pädagogischer Sinnhaftigkeit. Wenn es uns schon schwerfällt und mit vielen Fehlern behaftet ist, Schülerinnen und Schüler in nur 6 Kategorien sicher und zuverlässig einzuordnen,

dann ist mit 16 Stufen diese Problematik potenziert. Und wer bis hierhin den Kampf „Pädagogik versus Mathematik“ verfolgen konnte, vermag nachzuvollziehen, daß man mit dieser numerischen Eskalation und angeblichen Präzision das Kind im wahrsten Sinne des Wortes mit dem Bade ausschüttet.

Tendenzzeichen dürfen nicht dezimal verrechnet werden, für die Durchschnittsbildung ist von glatten Noten auszugehen. Die Tendenzzeichen sind als Entscheidungshilfe bei der pädagogischen Bilanzierung der Gesamtnoten heranzuziehen.

„Eine Epochalnote zählt doch mehr als ...“

... so die Äußerungen einer Kollegin, die sich über die Praxis eines Kollegen mokierte, eine Epochalnote ebenso stark zu gewichten wie eine Note aus einem Abfragen von 10 Vokabeln.

Tatsächlich gibt die Schulordnung vor, daß Noten unterschiedlich gewichtet werden können, wenn Umfang und Schwierigkeitsgrad dies nahelegen, was wohl in diesem Fall gegeben sein dürfte. Allerdings „können“ sie unterschiedliches Gewicht erhalten, sie „müssen“ nicht; d.h. es ist dem Kollegen in der Ausschöpfung seines Beurteilungsspielraums anheimgestellt, dies zu tun oder auch nicht.

„Wer es noch nicht mal schafft, 10 Vokabeln zu lernen, der soll hierfür eine kräftige Quittung bekommen, also zähle ich diese Note 3fach“, wäre ein zwar unpädagogischer, aber dennoch nichtvöllig abwegiger Standpunkt, den hoffentlich niemand bezieht, der aber deutlich macht, daß hier viele Auffassungen und Interpretationen möglich sind. Jeder Beurteiler muß das mit sich selbst ausmachen, und niemand sollte sich aufschwingen, die Wahrheit oder den Stein der Weisen gepachtet zu haben.

Daß in der Praxis von dem pädagogisch überaus sinnvollen Instrument der Gewichtung recht wenig Gebrauch gemacht wird, mag auch daran liegen, daß man den „Rechenaufwand“ bei der Durchschnittsbildung scheut. Denn eine Note doppelt, eine andere dreifach zu gewichten, wiederum andere nur einfach, und dann auch noch die Noten des zweiten Halbjahres stärker zu

zählen als die des ersten, das ist wirklich schon eine Mathematikaufgabe höheren Schwierigkeitsgrades. Zumindest zeigen einschlägige praktische Übungen im Rahmen unserer Seminararbeit, daß man sich schwer damit tut und hierzu mitunter recht abenteuerliche und leider auch falsche Rechenstrategien praktiziert werden. Deshalb sollte man sich mit dem einfachen „Trick“ behelfen, eine doppelt gewichtete Note auch gleich doppelt ins Notenbuch einzutragen, zum Beispiel in der Form 2-2, eine 3fach gewichtete z.B. als 4-4-4. Bei der Durchschnittsbildung können sie dann einfach wie zwei oder drei einzelne Noten behandelt werden, und die Rechnung wird um einiges schlichter.

Natürlich kann man mit dieser Technik Faktoren wie 1.5-fach oder 0.5-fach nicht bewältigen. Allerdings ist es meines Erachtens auch nicht sinnvoll, in der Gewichtung zu viele Differenzierungen vorzunehmen und damit wieder die Mathematik über Gebühr einkehren zu lassen. Es ist besser, sich auf ganzzahlige Faktoren zu beschränken.

Im Zusammenhang mit der Gewichtung sind noch zwei weitere Sachverhalte von enormer Bedeutung.

Zum einen sind dies wiederum die „heimlichen“ Gewichtungen, die sich automatisch aus der Anzahl der erteilten Noten ergeben. Wenn in einem Schuljahr 5 Klassenarbeiten geschrieben und im Bereich der anderen Leistungsnachweise nur 4 weitere Noten vergeben wurden, dann bedeutet dies wegen der 50-50-Regelung (die Zeugnisnote ist aus den beiden Gesamtnoten arithmetisch zu bilden), daß de facto jede Klassenarbeitsnote ein geringeres Gewicht hatte als jede Note aus den anderen Leistungsnachweisen, z.B. einem schriftlichen Abfragen der Hausaufgaben. Um dies zu verhindern, müßte (und muß?) die Anzahl der Noten in der Kategorie „andere Leistungsnachweise“ um einiges höher sein als die Zahl der Klassenarbeitsnoten. Einschränkung und abschwächend muß man allerdings zugestehen, daß die Schulordnung eben die beiden Kategorien „Klassenarbeiten“ und „Andere Leistungsnachweise“ streng trennt, und daß es nicht a priori legitim ist, Noten aus diesen streng getrennten Bereichen in ihrer Gewichtung gegeneinander aufzurechnen. Zum zweiten stecken in Gewichtungen ein erhebliches Unwägbarkeitspotential für die Betrof-

fenen, also die Schülerinnen und Schüler resp. die Eltern. Der aktuelle Leistungsstand muß aber im Grunde jederzeit kalkulierbar sein, und Lehrerinnen und Lehrer sind fairerweise verpflichtet, hierüber Auskunft zu geben. Deshalb dürfen nicht erst im Zuge der Festlegung der Zeugnisnote unvermittelt Gewichtungen aus dem Hut gezaubert werden, sie müssen vorher bekannt und den Betroffenen geläufig sein. Dies ist ein weiteres Argument dafür, einerseits mit Gewichtungen sparsam umzugehen, andererseits aber auch einschlägige Regeln festzulegen (wie z.B. „Epochalnoten zählen dreifach“) oder ggfs. Abweichungen vom Faktor 1 vorher bekanntzugeben (z.B. „die Klassenarbeit Nr. 3 zählt doppelt, weil sie sehr viel mehr Stoff umfaßte, weil sie schwierig war und weil sie 3stündig geschrieben wurde“).

Mit Gewichtungen muß sparsam umgegangen werden. Es sollten nur wenige Faktoren verwendet werden, damit der Leistungsstand überschaubar und kalkulierbar bleibt. Die Gewichtungen müssen den Beurteilten rechtzeitig bekannt sein und begründet werden,

Die Auswirkungen von Gewichtungen auf den Durchschnitt sollen im folgenden noch an einem Beispiel verdeutlicht werden. Ein Notenbuch enthält folgende Eintragungen:

Manuela Muster / Erdkunde	1. Halbjahr	2. Halbjahr
Schriftliche Überprüfung*	3-3-3	5-5-5
Epochalnote	4-4, 5-5	3-3
Schriftliches Abfragen Hfg.	2, 3	3
Mündliche Überprüfung	3	3, 4
Heft	2	3
Referat	-	4-4

* für die schriftliche Überprüfung schreibt die Schulordnung das höhere Gewicht vor. Der Lehrer hat bestimmt, daß sie 3fach zählt, und das gleich entsprechend eingetragen. Epochalnoten hat er als doppelt zählend festgelegt, das Referat soll ebenfalls doppeltes Gewicht haben.

Manuela hatte im Halbjahreszeugnis eine „befriedigend“.

- Zählt man die schriftliche Überprüfung 3fach, Epochalnote und Referatsnote 2fach, dann ergibt der einfache Durchschnitt (beide Halbjahre 1:1) 3,6, das zweite Halbjahr doppelt gerechnet ergibt 3,7

- Zählt man die schriftliche Überprüfung 3fach, alle anderen Noten aber nur 1fach, so ergibt sich 3.5 bzw. 3.6
- Zählt man die schriftliche Überprüfung nur 2fach, alle anderen Noten 1fach, so ergibt sich 3.4 bzw. 3.5

Das Beispiel zeigt, daß die gewichtungsbedingten Unterschiede in den meisten Fällen geringer sind als man intuitiv annimmt. Dennoch: Mit den beiden letzten Modalitäten erhält die Schülerin Manuela wahrscheinlich noch eine „befriedigend“, bei der ersten muß sie womöglich doch mit der Note „ausreichend“ rechnen (rein vom Notenbild her ist diese Entscheidung nicht zu treffen, hier braucht es die Kenntnis der Schülerin, ihrer „Lerngenese“, ihres Lernverhaltens und ihrer Leistungsbereitschaft).

Das Heft nicht aus der Hand geben!

Die Intentionen der Schulordnung sind in erster Linie pädagogisch determiniert. Begriffe und Formulierungen wie „pädagogische Verantwortung und Freiheit“, „Festsetzung“, „... muß nicht der rechnerische Durchschnitt ... sein“, „Berücksichtigung“, „Gesamteindruck“ u.s.f. signalisieren dies eindeutig. Wenn wir uns - nicht ganz unbegründet - darüber beschweren, daß das Korsett von Auflagen, Bestimmungen und Rechtfertigungszwängen immer enger geschnürt wird, dann sollten wir diese Tendenz nicht von uns aus weiter schüren, indem wir in uns belassenen Freiräumen nach Automatismen, Regeln und Regularien - z.B. jenen der Mathematik - verlangen, die uns die Entscheidung letztlich abnehmen; dann werden Noten nicht mehr festgesetzt, sondern schlicht und einfach berechnet. Die Numerik zum Maß der Dinge zu machen, gegen Ende des Schuljahres per Knopfdruck via Tabellenkalkulation die Noten aus einem Computer abrufen zu können, erspart doch manche Bürde oder schwere Last und auch gelegentlichen Argumentationsnotstand.

Das Heft aus der Hand zu geben und Freiräume und Entscheidungsspielräume aufzugeben, reduziert die Verantwortung, und dies mag manchem attraktiv erscheinen, insbesondere dann, wenn eine angeblich so neutrale und überaus objektive Instanz wie die Mathematik selbige übernimmt.

Wer sich nicht auf die Numerik zurückziehen bereit ist, das Heft eben nicht aus der Hand ge-

ben will, trägt ein weitaus höheres Maß an Verantwortung, braucht andere Kompetenzen und Praktiken denn mathematische. Und dazu gehören vor allem Offenheit gegenüber den Beurteilten und Transparenz der Beurteilungsmechanismen, und das wie ein Wertgegenstand gehütete „heimliche Notenbuch“ sollte ebenso der Vergangenheit angehören wie das „Überraschungsei“ am Ende eines Beurteilungszeitraums.

In Carl Zuckmayers „Hauptmann von Köpenick“ fordert Schuster Voigt „zuerst kommt der Mensch - und dann die Menschenordnung!“. In diesem Sinne muß das Bewußtsein der Beurteiler wie auch der Beurteilten gestärkt werden, daß zuerst der Mensch kommt, der Schüler und die Schülerin, die Pädagogik - und dann die Mathematik, und daß Menschen und die Beurteilung ihrer Leistungen nicht auf Ziffern und Rechenergebnisse reduziert werden können und dürfen.

Für die Beurteilung von Notenbildern und deren Verteilung hat die Statistik eine Menge Mittel bereitgestellt; der Durchschnitt als charakterisierende Größe ist nur eine davon, und sogar ein recht bescheidenes Instrument, in der täglichen Praxis hinsichtlich der Aussagekraft bei weitem überschätzt und oft fehlinterpretiert. So heißt es schon in einem Mathematik-Schulbuch der 8. Klasse („Schnittpunkt“, Ausgabe RLP, S. 188, Klett Verlag Stuttgart 1995): „... deshalb sollten in der Regel extreme Einzelwerte (sogenannte Ausreißer) ... weggelassen werden, da sie die Aussagekraft des Mittelwerts stark verfälschen“ (um Gottes willen, wo kämen wir hin, wenn wir Noten weglassen würden?).

Es gibt weitere Größen, die unter Umständen sehr viel besser geeignet wären, auf so etwas wie „durchschnittliche Leistung“ zu schließen, z.B. Zentralwert, Quantile, p%-Quantile, Modalwert (zur Beurteilung „schiefer“ Verteilungen) und lineare Regression (für den Notentrend), um nur einige zu nennen.

Aber ich will nicht - mir selbst widersprechend - für noch mehr Mathematik plädieren (höchstens für sinnvollere). Lassen wir's doch beim Durchschnitt, aber nehmen wir ihn als das, was er ist: ein relativ bescheidenes Mittel, sich über ein Notenbild Klarheit zu verschaffen, und eine Hilfsgröße, die zur Festsetzung der Gesamtnoten einen eher kleinen Beitrag zu leisten vermag.