

Ist die Bewertung von Prüfungsleistungen nach dem degressiven 100-Punkte-Schlüssel noch zeitgemäß?

Jens U. Schmidt

*Dr. phil., Dipl.-Psychologe,
wissenschaftlicher Mitarbeiter in der Abteilung 2.2
„Prüfungen und Lernerfolgskontrollen“ im Bundesinstitut für Berufsbildung, Berlin*

Vor über 25 Jahren wurde in Musterprüfungsordnungen für Ausbildungen, Fortbildungen und Meisterprüfungen ein Umrechnungsverfahren für Punkte in Noten festgelegt, der degressive 100-Punkte-Schlüssel. Diese Methode gerät immer wieder in die bildungspolitische Diskussion, obwohl – wie empirisch gezeigt wird – das eigentliche Problem in der Auffassung darüber besteht, wie schwierig Prüfungen sein sollen, welche Durchfallquote und welche Quote guter Leistungen akzeptabel erscheinen, und wie stark die Leistungen der Prüfungsteilnehmer streuen dürfen.

Das Thema „Bewertung und Benotung von Prüfungsleistungen“ ist immer wieder Grund für bildungspolitische Auseinandersetzungen, auch wenn diese Debatten nicht so breit und öffentlich geführt werden wie diejenigen, bei denen es um Innovationen im Prüfungswesen geht wie integrierte Prüfungen, Gruppenprüfungen, Betriebsaufgabe, handlungsorientierte Aufgabengestaltung, computergestützte Prüfungen¹ usw. Im Mittelpunkt der Diskussion der Bewertung steht ein eher technisch anmutender Aspekt, dessen Bedeutung für Qualität und Aussagekraft von Prüfungen möglicherweise überbetont wird: der Umrechnungsschlüssel von Punkten in Noten. Die vorliegende Arbeit basiert auf einer auf Veranlassung des Deutschen Handwerkskammertages (DHKT) und des Deutschen Gewerkschaftsbundes (DGB) erstellten Expertise des Bundesinstituts für Berufsbil-

dung. Im Rahmen der am 1. 1. 1994 in Kraft getretenen Novelle zur Handwerksordnung wurde die Meisterprüfungsordnung geändert und bei dieser Gelegenheit vom DGB eine Veränderung des Umrechnungsschlüssels gefordert. Man einigte sich zunächst auf eine Beibehaltung des bisherigen Verfahrens und eine erneute Diskussion nach Vorliegen der Expertise.

Der praktizierte degressive Umrechnungsschlüssel für Punkte in Noten wurde allerdings auch von anderen Personen und Institutionen immer wieder in Frage gestellt, seit zur Vereinheitlichung der von den zuständigen Stellen zu erlassenden Prüfungsordnungen vom Bundesausschuß für Berufsbildung **Musterprüfungsordnungen**² erlassen wurden.

Hier ist ein Umrechnungsschlüssel festgelegt worden, der eine Bewertung von Prüfungsleistungen mit 0–100 Punkten vorsieht, die dann nach einem degressiven und nicht linearen System in sechs Notenstufen umzurechnen sind. Degressiv bedeutet, daß der Punktbereich zunehmend enger wird, je besser die Leistungen sind. Während für die Note 6 der Bereich 0 bis 29 Punkte (Spannweite 30 Punkte) umfaßt, für die 5 der Bereich 30 bis 49 Punkte (Spannweite 20), sind die Bereiche für die 1 mit 92 bis 100 Punkte (Spannweite 9) und die 2 mit 81 bis 91 (Spannweite 11) erheblich geringer. Ein Prüfungsteil ist bestanden, wenn die Note 4 erreicht wurde. Dies soll dann der Fall sein, wenn die Hälfte der maximal möglichen Punktzahl erreicht ist, also bei 50 Punkten.

Man muß sich vor Augen halten, daß diese Form der Bewertung seit über 20 Jahren in fast allen Ausbildungs- und Fortbildungsprüfungen praktiziert wird, also die Mehrheit der heute im Arbeitsleben stehenden Personen nach diesem Verfahren benotet wurden. Es ist daher auch keine Diskussion für Spezialisten, wenn immer wieder dieses Bewertungsverfahren hinterfragt wird. Welche Argumente lassen sich für eine Änderung des Umrechnungsschlüssels heranziehen? Entscheidend dürfte eine im Einzelfall möglicherweise bestehende Unzufriedenheit mit der Durchfallquote sein bzw. mit dem Anteil von Personen, die besonders erfolgreich in der Prüfung sind, also mit einer 1 oder 2 abschließen. Wie kraß die Auswirkungen einer schlichten Änderung des Umrechnungsschlüssels sind und welche weiteren Konsequenzen sich daraus ergeben würden, wird erst deutlich, wenn man dies wie im vorliegenden Artikel anhand empirischer Daten analysiert.

Das eigentliche Dilemma besteht darin, daß Prüfungen im strengen Sinne nicht **kriteriumsorientiert** erfolgen. Nicht jeder, der eine Prüfung nicht besteht, ist unfähig, in dem entsprechenden Beruf zu arbeiten, und nicht jeder erfolgreiche Kandidat verfügt wirklich über die notwendige Handlungskompetenz. Die Gestaltung einer Prüfung bezogen auf ein eindeutiges Kriterium „Berufseignung“ wird sicher auch langfristig eine Utopie bleiben, da dies für jeden einzelnen Abschluß eine aufwendige Bewährungsanalyse erfordert. Bis dahin muß man sich mit Setzungen behelfen, also mit praktikablen Bestehensgrenzwerten und Umrechnungsschlüsseln.

Als Alternative zum degressiven Umrechnungsschlüssel wird vorgeschlagen, mit gleich großen Punktintervallen für jede Note zu arbeiten, also mit einem linearen Schlüssel. Soll am Bestehenskriterium von mindestens 50 Punkten festgehalten werden, so wären die Leistungen unterhalb und oberhalb von 50 Punkten in jeweils vier Notenstufen

Tabelle 1: **Gegenüberstellung von drei Umrechnungsschlüsseln für Punkte in Noten**
Punkteschlüssel zur Umrechnung von 0–100 Punkten in 6 Noten

Note	Degressiver Schlüssel	Linearer Schlüssel	Teil-linearer Schlüssel
1	100–92	100–83	100–87,5
2	unter 92–81	unter 83–67	unter 87,5–75
3	unter 81–67	unter 67–50	unter 75–62,5
4	unter 67–50	unter 50–34	unter 62,5–50
5	unter 50–30	unter 34–17	unter 50–25
6	unter 30–0	unter 17–0	unter 25–0

aufzuteilen. Eine Differenzierung ungenügender Leistungen in 4 Notenstufen dürfte jedoch kaum sinnvoll sein. Würde der Bestehens-Grenzwert 50 aufgegeben, gleichzeitig an den 6 Noten festgehalten werden, wäre bei einem linearen Schlüssel die Bestehensgrenze zwischen 33 und 34 Punkten anzusetzen.

Als dritte Möglichkeit bietet sich ein teil-linearer Schlüssel an, bei dem am Grenzwert 50 festgehalten, der Bereich unterhalb von 50 in zwei gleiche Intervalle und der oberhalb in vier gleiche Intervalle aufgeteilt wird. Alle drei Möglichkeiten sind einander in Tabelle 1 gegenübergestellt.

Umrechnungsschlüssel in anderen Bereichen

Um festzustellen, ob Umrechnungsschlüssel in anderen Bereichen überhaupt verwendet werden und wie diese aussehen, wurden vorliegende Unterlagen gesichtet und unterschiedliche Institutionen telefonisch befragt, wobei diese Befragung nicht den Anspruch auf Repräsentativität erhebt. Die angeführten Verfahren sollen das vorhandene Methodenspektrum aufzeigen.

In **allgemeinbildenden Schulen** wird in der Primarstufe und Sekundarstufe I ebenso wie

in berufsbildenden Schulen normalerweise mit den bekannten 6 Notenstufen und deren verbalen Umschreibungen gearbeitet. Dem Lehrer ist es je nach Art der Lernerfolgskontrolle (Test, Klassenarbeit, Klausur, mündliche oder schriftliche Prüfungsarbeit) freigestellt, einen eigenen Punkte-Umrechnungsschlüssel zu erstellen, wobei gemäß der verbalen Umschreibungen und der Orientierung an der Gaußschen Normalverteilung eine größere Anzahl von Noten im Mittelbereich (3 und 4) als Noten im oberen und unteren Bereich angestrebt wird. Die Noten 1 und 6 stellen Ausnahmen dar. Es gibt hier allerdings in den einzelnen Bundesländern auch darüber hinausgehende Regelungen.

In der Sekundarstufe II (gymnasiale Oberstufe) wird mit einem linearen Punkteschlüssel mit 16 Stufen bewertet mit folgenden Zuordnungen: 0 Punkte = Note 6, 1–3 = Note 5, 4–6 = Note 4, 7–9 = Note 3, 10–12 = Note 2, 13–15 = Note 1. Die Bestehensgrenze liegt in der Regel bei der Note 4 bzw. bei 4 Punkten, also bei 31 Prozent der maximal möglichen Punktzahl.

Ansonsten gibt es offenbar in allgemeinbildenden und berufsbildenden Schulen sowie an Fachschulen keine allgemeinverbindlichen Regelungen über Punkteschlüssel, lediglich die verbal umschriebenen Notenstufen.

In den vom Sekretariat der Ständigen Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland herausgegebenen Allgemeinen Bestimmungen für Diplomprüfungsordnungen an Fachhochschulen, Universitäten und gleichgestellten Hochschulen sowie den Allgemeinen Bestimmungen für Magisterprüfungsordnungen in den aktuellen Fassungen finden sich keine Punkteschlüssel. Angegeben sind lediglich verbale Umschreibungen der 5 Notenstufen (Note 6 nicht vorgesehen). Die Note kann zur differenzierteren Bewertung um 0,3 erhöht oder herabgesetzt werden, wobei jedoch die Noten 0,7, 4,3 und 5,3 ausgeschlossen sind. Bestehenskriterium ist eine Note von mindestens 4,0.

Eine besondere Regelung besteht bei Juristen, bei denen ein Punktesystem vorgegeben ist.³ Hier gibt es 7 Noten und eine Zuordnung zu 0–18 Punkten. Bei rechnerisch ermittelten Gesamtnoten gilt ein anderer Punkteschlüssel als bei der Bewertung von Einzelleistungen. Dieser sieht breitere Punktintervalle für sehr gute und ungenügende Leistungen vor.

Im öffentlichen Dienst sind einige hundert Prüfungsordnungen erlassen worden, die hinsichtlich der Bewertung sehr uneinheitlich sind. In einigen Ländern wird lediglich mit Noten, in einigen mit einem der Sekundarstufe II der allgemeinbildenden Schulen vergleichbaren Schlüssel gearbeitet. Das Bundesverwaltungsamt hat für Ausbildungen des Bundes in den Berufen Verwaltungsfachangestellte/r, Kartograph/-in und Ver- und Entsorger/-in Prüfungsordnungen erlassen, in denen der in Tabelle 1 dargestellte teil-lineare Umrechnungsschlüssel vorgegeben ist. Dieser Umstand hat schon mehrfach zu Auseinandersetzungen im Zusammenhang mit dem sowohl im öffentlichen Dienst als auch in der Wirtschaft ausgebildeten Beruf Kartograph geführt, da hier mitunter die Leistungen von Auszubildenden der gleichen Berufsschulklasse, die die gleiche Prüfung bearbeiteten, unterschiedlich benotet wurden.

Das Problem ist bis heute ungelöst. Das Bundesinstitut für Berufsbildung hat anlässlich einer Petition empfohlen, die Prüfungszeugnisse aus der Wirtschaft um eine Bewertung nach den Richtlinien des öffentlichen Dienstes zu ergänzen. Eine ähnliche Problematik besteht offenbar bei den Abschlüssen Techniker der IHK und Staatlich geprüfte Techniker.

Um weitere methodische Möglichkeiten aufzuzeigen, werden noch zwei weitere Beispiele dargestellt, ohne hier den Versuch eines internationalen Vergleichs zu unternehmen:

In der DDR wurde in der beruflichen Ausbildung ein Punkteschlüssel empfohlen, wobei in Abhängigkeit von der Art der erbrachten Leistung jeweils zu entscheiden ist, ob die Anwendung von Punkten bei der Notengebung sinnvoll ist. Das verbal umschriebene Zensurensystem ist immer ausschlaggebend. Folgende degressive Zuordnung wurde getroffen: 1 = mindestens 96 Prozent, 2 = mindestens 80 Prozent, 3 = mindestens 60 Prozent, 4 = mindestens 36 Prozent, 5 = unter 36 Prozent der maximal möglichen Punkte.

In Irland gibt es innerhalb der beruflichen Ausbildung Umrechnungsschlüssel. Hier wird als Beispiel die Regelung für Electricians dargestellt. Gearbeitet wird mit 7 Notenstufen (grades), die mit Buchstaben bezeichnet werden. Es erfolgen folgende Zuordnungen zu Punkten:

A = 85–100; B = 70–84; C = 55–69; D = 40–54; E = 25–39; F = 10–24; NG = 0–10.

Interessant ist dabei, daß die Bestehensgrenzen für die schriftliche und praktische Prüfung unterschiedlich angesetzt sind. Für das Bestehen einer schriftlichen Prüfung ist ein D ausreichend, für die praktische Prüfung ein C. Das Bestehen der gesamten Prüfung setzt das Bestehen aller verbindlichen Prüfungsteile voraus.

Es zeigt sich, daß recht unterschiedliche Umrechnungsschlüssel praktiziert werden, in

vielen Bereichen aber nur mit einem verbal umschriebenen Notensystem gearbeitet wird. Die durch die Musterprüfungsordnungen angestrebte Vereinheitlichung der Prüfungsordnungen ist hinsichtlich des Bewertungsverfahrens – abgesehen vom öffentlichen Dienst – weitgehend gelungen.

Empirischer Vergleich unterschiedlicher Umrechnungsschlüssel

Die Auswirkungen unterschiedlicher Punkteschlüssel sollen durch einen empirischen Vergleich anhand echter Daten einer schriftlichen Ausbildungsabschlußprüfung mit mehr als 3 000 Teilnehmern⁴ verdeutlicht werden. Die drei o. g. Umrechnungsschlüssel werden zunächst auf einen typischen Prüfungsteil angewandt mit einer durchschnittlichen Schwierigkeit von 0,70, also einen Satz von Aufgaben, die im Mittel von 70 Prozent der Teilnehmer richtig gelöst wurden (Beispiel A). Diese Höhe der Schwierigkeit wird nach Aussage von verschiedenen Aufgabenerstellungsinstitutionen bei der Entwicklung von Prüfungsaufgaben angestrebt und meist auch erreicht. Im zweiten Schritt erfolgt die Anwendung der drei Umrechnungsschlüssel auf einen schwierigeren Aufgabensatz, im dritten Schritt auf einen Aufgabensatz, in dem die Punktwerte der Prüfungsteilnehmer weniger stark streuen als üblich. Die Ergebnisse sind in Tabelle 2 zusammengefaßt.

Bei einem Prüfungssatz mit herkömmlicher Schwierigkeit von 0,70 (Beispiel A) ist der lineare Schlüssel wenig zweckmäßig, da er bei dieser Aufgabenschwierigkeit zu einer viel zu hohen Anzahl guter Noten und einer viel zu geringen Durchfallquote führt. Der teil-lineare Schlüssel führt zu einer akzeptablen Durchfallquote von sechs Prozent, ebenso wie der degressive Schlüssel, gleichzeitig jedoch zu einer außerordentlich hohen An-

Tabelle 2: **Quoten guter Noten und Durchfallquoten bei verschiedenen Umrechnungsschlüsseln und unterschiedlichen Schwierigkeiten und Punkteverteilungen in den Prüfungsdaten**

			Degressiver Schlüssel	Linearer Schlüssel	Teil-linearer Schlüssel
A	Übliche Prüfungsschwierigkeit	Anteil Note 1 u. 2	21	60	39
		Durchfallquote	6	0,2	6
B	Höherer Schwierigkeitsgrad	Anteil Note 1 u. 2	4	18	8
		Durchfallquote	49	20	49
C	Verringerte Punktstreuung	Anteil Note 1 u. 2	11	41	24
		Durchfallquote	14	1	14

zahl der Noten 1 und 2. Es erscheint fraglich, ob es tatsächlich wünschenswert ist, wenn dieser Anteil bei 39 Prozent liegt, solange mit diesen Noten eine besondere Auszeichnung verbunden sein soll.

Die bei Prüfungen im dualen System gewünschte Notenverteilung bei einem Schwierigkeitsgrad von 0,70 ergibt sich am ehesten bei Verwendung des degressiven Schlüssels. Die Notenverteilung entspricht weitgehend der Gaußschen Normalverteilung mit größter Häufung im Mittelbereich und einer niedrigen, aber durchaus nennenswerten Durchfallquote auf der einen und einer bei 21 Prozent liegenden Quote guter Noten auf der anderen Seite.

Eine völlig andere Situation ergibt sich bei Aufgaben mit höherer Schwierigkeit. Dies wird am Beispiel B in Tabelle 2 demonstriert, bei dem die Daten so manipuliert wurden, daß die durchschnittliche Schwierigkeit bei 0,50 liegt, also die Aufgaben im Durchschnitt nur von 50 Prozent der Prüflinge richtig bearbeitet wurden.

Man sieht hier sehr deutlich, daß sowohl der degressive als auch der teil-lineare Schlüssel zu inakzeptabel hohen Durchfallquoten von 49 Prozent führen, während diese Quote beim linearen Schlüssel in der eher akzeptablen Größenordnung von 20 Prozent liegt. Es ist klar, daß sich bei einer Bestehensgrenze

von 50 Prozent der Punkte und einer durchschnittlichen Schwierigkeit von 50 Prozent Aufgabenlösung eine Durchfallquote von ebenfalls etwa 50 Prozent ergeben muß. Auch im Bereich der besseren Noten käme für diese Daten nur der lineare Schlüssel in Betracht, da die Quote der Noten 1 und 2 vor allem bei Verwendung des degressiven, jedoch auch des teil-linearen Schlüssels für die Praxis zu gering sein dürfte.

Neben der Schwierigkeit spielt auch die Streubreite der vorkommenden Punkte eine wichtige Rolle. Streuen sie relativ stark, was bedingt durch unterschiedlich schwierige Teilaufgaben meist der Fall ist, erscheint ein anderer Schlüssel zweckmäßig als bei geringer Streuung.

Dies ist insbesondere bei Summen- und Durchschnittswerten aus mehreren Prüfungsleistungen relevant. Werden die Punktwerte der Prüfungsteile aufsummiert, dann der Summenwert ins 100-Punkte-Schema transformiert und in eine Note umgewandelt, haben solche Summenwerte normalerweise eine deutlich geringere Streuung als die in den Summenwert eingehenden Einzelleistungen. In Abbildung 1 wird dies verdeutlicht durch Gegenüberstellung der Häufigkeitsverteilung eines Prüfungsteils und eines Summenwerts aus diesem und drei weiteren Prüfungsteilen. Es ist deutlich zu sehen, daß die Kurve des

Prüfungsteils sehr viel flacher verläuft als der Summenwert. Damit nimmt die Wahrscheinlichkeit besonders guter und schlechter Noten durch die Summenbildung ab.

Bei Umwandlung dieses Summenwerts nach den drei Umrechnungsschlüsseln in Noten ergeben sich die in Tabelle 2, Beispiel C dargestellten Quoten für gute und schlechte Prüfungsleistungen.

Beim Vergleich von Beispiel A und C wird deutlich, daß die Anzahl guter Leistungen für einen derartigen Summenwert beim degressiven Schlüssel mit nur noch elf Prozent möglicherweise etwas zu niedrig ist. Hier ist der teil-lineare Schlüssel eher geeignet, sofern etwa 20 Prozent der Prüflinge mit guten Noten ausgezeichnet werden sollen. Dieser generell geltende Umstand – geringere Streuung zusammengefaßter Leistungen – dürfte auch der Grund dafür sein, daß bei den Juristen (s. o.) ein anderer Schlüssel für die Bewertung der Gesamtleistung vorgesehen ist als für die Einzelleistungen.

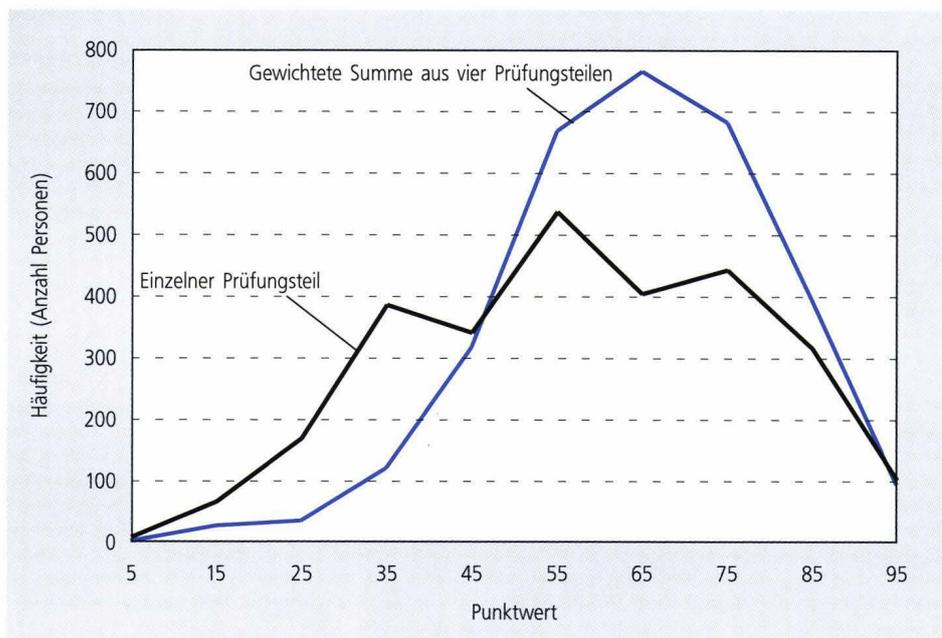
Umrechnungsschlüssel für Durchschnittsnoten

Mit dem letzten Beispiel wird ein sehr wichtiger Aspekt berührt: die Frage, ob man bei Durchschnittsnoten oder Gesamtnoten mit einem anderen Schlüssel arbeiten sollte als bei der Bewertung von Einzelleistungen.

Die Musterprüfungsordnungen für Ausbildungen und Umschulungen regeln nicht, wie ein Gesamtergebnis zu ermitteln ist. Die Prüfung wird, sofern dies durch die Ausbildungsordnung nicht anders bestimmt ist, als bestanden bewertet, wenn in den einzelnen Prüfungsteilen ausreichende Leistungen erbracht werden. Anders ist dies in Fortbildungsprüfungen und Meisterprüfungen geregelt, wo der Durchschnitt aller Leistungen zu errechnen ist, der dann mindestens ausreichend sein muß.

Hier liegt nun ein generelles Dilemma: die mathematische Unzulässigkeit einer Durch-

Abbildung 1: Häufigkeitsverteilung der Punktwerte eines Prüfungsteils und der Summe aus vier Prüfungsteilen (jeweils innerhalb des 100-Punkte-Systems)



schnittsbildung von Noten auf der einen Seite und die Unzweckmäßigkeit des degressiven Umrechnungsschlüssels für Durchschnittspunktzahlen auf der anderen. Der Durchschnitt darf nur bei metrischen Skalen berechnet werden, also bei Skalen mit konstanten Abständen zwischen benachbarten Skalenpunkten. Die Notenskala erfüllt diese Bedingung generell nicht, insbesondere nicht nach Anwendung des degressiven Umrechnungsschlüssels, da die einer Note zugeordneten Punktabstände ungleich sind. Die Struktur der Punkteskala kann als metrisch angesehen werden, da die Abstände zwischen benachbarten Skalenwerten hier eher als konstant angesehen werden können.

Damit dürfte die Bildung von durchschnittlichen Leistungen nur auf der Ebene der Punkte erfolgen, wobei dann der Durchschnittswert in eine Note transformiert wird. Dies führt jedoch beim degressiven Umrechnungsschlüssel, wie Ecker⁵ ausführt, zu einer für den Prüfling schwer einsehbaren Ungerechtigkeit. Werden in der schriftlichen Prüfung 92 Punkte, also Note 1, erreicht, im praktischen Teil 67 Punkte, Note 3, ergibt

sich nicht die erwartete Gesamtnote 2, sondern eine 3, denn die Durchschnittspunktzahl 79,5 fällt nach dem Umrechnungsschlüssel in das Intervall für die 3. Diese mathematisch zulässige Vorgehensweise hat also in Kombination mit dem degressiven Schlüssel die Konsequenz, daß die Gesamtnote schlechter ausfällt als bei Durchschnittsbildung von Noten bzw. bei Zugrundelegung eines linearen Schlüssels. Es muß also ernsthaft diskutiert werden, ob es nicht notwendig wäre, das Verfahren der Durchschnittsbildung und der Umrechnung in Noten neu zu regeln. Eine solche Durchschnittsbildung findet übrigens auch bei der Einbeziehung von mündlichen Ergänzungsprüfungen in das Prüfungsergebnis statt, wo sich durch die Gewichtung zusätzliche Probleme ergeben. Auf diesen Aspekt kann an dieser Stelle nicht näher eingegangen werden.

Welcher Schlüssel ist für welche Situation geeignet?

Aus den Ausführungen sollte deutlich geworden sein, daß die Zweckmäßigkeit eines Um-

rechnungsschlüssels von den Rahmenbedingungen und Anforderungen der Prüfung abhängt. Es gibt keinen richtigen oder falschen Schlüssel, sondern lediglich für verschiedene Bedingungen zweckmäßigere und weniger zweckmäßige Verfahren. Relevant sind vor allem folgende Faktoren:

- Schwierigkeitsgrad der Aufgaben: Werden die Aufgaben im Durchschnitt von 50 Prozent der Prüflinge gelöst oder sind dies erheblich mehr (Schwierigkeit 0,50 oder größer 0,50)?
- Streuung der Punktwerte: Gibt es viele Prüfungsteilnehmer mit Punktzahlen nahe dem Mittelwert oder verteilen sich die Punktzahlen über einen breiten Bereich, ist die Streuung also klein oder groß?
- Bestehenskriterium: Wird die Erreichung der Hälfte der maximal möglichen Punkte verlangt oder deutlich weniger?
- Auszeichnung durch gute Leistungen: Sollen besonders gute Noten (1 und 2) die Ausnahme sein (ca. zehn Prozent der Prüflinge) oder von einem größeren Anteil der Prüflinge (ca. 20 Prozent) erreicht werden?

Aus der Kombination dieser Faktoren ergeben sich 16 unterschiedliche Situationen, für die sich jeweils andere Umrechnungsschlüssel anbieten (vgl. Tabelle 3). Einige Kombinationen sind nicht sinnvoll, wie z. B. die durchschnittliche Aufgabenschwierigkeit von 0,50 bei einem Bestehenskriterium von 50 Prozent der möglichen Punkte. Diese Kombination muß zu einer wohl kaum wünschenswerten Durchfallquote von 50 Prozent führen.

Greifen wir drei Beispiele heraus:

- Derzeitige Prüfungen im dualen System sind durch durchschnittliche Aufgabenschwierigkeiten von 0,70 gekennzeichnet, also deutlich leichter als 0,50. Die Streuung der Punkte ist groß, das Bestehenskriterium mit 50 Punkten festgelegt und ein Anteil von 20 Prozent guter Noten wünschenswert. Für diese Situation eignet sich gemäß Tabelle 3 also am ehesten der degressive Umrechnungsschlüssel.

Tabelle 3: **Entscheidungshilfe für unterschiedliche Umrechnungsschlüssel und die Rahmenbedingungen der Prüfungserstellung**

Aufgabenschwierigkeit	0,5								deutlich größer als 0,5							
	klein				groß				klein				groß			
Bestehensgrenze	50	< 50			50	< 50			50	< 50			50	< 50		
% gute Leistungen ca.	20	10	20	10	20	10	20	10	20	10	20	10	20	10	20	10
zweckmäßiger Schlüssel	-	-	S	L	-	-	L	S	T	D	-	-	D	D*	-	-

D: Derzeit praktizierter degressiver Schlüssel
D*: Stärker degressiver Schlüssel mit im oberen Bereich noch engerer Streubreite
L: Linearer Schlüssel
S: Symmetrischer degressiver Punkteschlüssel (jeweils gleiche Punktebereiche für 3 und 4, 2 und 5, 1 und 6, groß bei 3 und 4, kleiner bei 2 und 5, klein bei 1 und 6), Orientierung an der Gaußschen Normalverteilung.
T: Teil-linearer Punkteschlüssel
-: Kombination ist innerhalb des Ausbildungswesens nicht sinnvoll

- Für Gesamtnoten, die durch Durchschnittsbildung der Einzelleistungen auf der Ebene von Punkten vorgenommen wurden, gelten ebenfalls Schwierigkeit größer 0,50, Bestehensgrenze 50 und gewünschter Anteil von 20 Prozent guter Noten, jedoch ist die Punktstreuung kleiner. Daher eignet sich hier am ehesten der teil-lineare Schlüssel.
- In der gymnasialen Oberstufe sind die Aufgaben vergleichsweise schwierig, die durchschnittliche Schwierigkeit wäre mit 0,50 zu beziffern, wobei die Punktwerte wenig streuen. Das Bestehenskriterium liegt deutlich unter 50 Prozent der maximal möglichen Punkte (4 von 15 Punkten). Mit besonders guten Noten wird nur ein kleiner Prozentsatz der Schüler ausgezeichnet. Nach Tabelle 3 wäre also der lineare Schlüssel heranzuziehen.

Fazit

In diesem Beitrag sollte deutlich gemacht werden, daß es nicht sinnvoll ist, aufgrund unbefriedigender Notenverteilungen bei Prüfungen den Umrechnungsschlüssel in das Zentrum der Diskussion zu stellen. Er ist lediglich ein technisches Hilfsmittel, das je nach Rahmenbedingungen der Aufgabensätze, angestrebter Durchfallquote und gewünschtem Anteil guter Leistungsbeurteilungen anders konstruiert werden muß.

Niemand wird behaupten, daß diese Aspekte quasi naturgegeben sind. Sie sind Setzungen oder gewachsene bzw. gewohnte Rahmenbedingungen einer Leistungsbeurteilung. Der Schwierigkeitsgrad von Aufgabensätzen hat sich bei einer bestimmten Größenordnung eingependelt, bedingt durch die langjährige Erfahrung der Aufgabenersteller und Rückmeldungen aus den jeweils vorangegangenen Prüfungen. Unter den heutigen Bedingungen bei Prüfungen im dualen System ist also der praktizierte degressive Umrechnungsschlüssel am geeignetsten.

Unbefriedigend ist in diesem Zusammenhang, daß auch die Durchfallquote als Setzung betrachtet wird, also definiert wird, wie viele Prüflinge durchfallen sollen. Sehr viel besser wäre, wie oben erwähnt, eine echte kriteriumsorientierte Prüfung, bei der derjenige durchfällt, der objektiv das Lernziel nicht erreicht hat. Vor dem Hintergrund der derzeit praktizierten eher vagen Festlegung von Prüfungsanforderungen und dem Fehlen klarer Kriterien ist dieser Ansatz aber utopisch.

Bei einer Veränderung des Verfahrens, also beispielsweise der Einführung eines linearen Schlüssels mit einem Bestehensgrenzwert von 33 Prozent, müßten künftig Aufgaben

mit einem durchschnittlichen Schwierigkeitsgrad von 0,50 entwickelt werden. Die Aufgabenersteller müßten langjährige Erfahrungen über Bord werfen und sich darauf einstellen, erheblich schwierigere Aufgaben zu entwickeln. Der Nutzen dürfte mit dem Aufwand für diesen Schritt kaum in einem sinnvollen Verhältnis stehen.

Ein Aspekt sollte weiter im Auge behalten werden: die bei Durchschnitts- oder Gesamtnoten verringerte Punktstreuung. Sollte hier tatsächlich die Quote guter Leistungen unerwünscht niedrig liegen oder als ungerecht angesehen werden, müßte über ein anderes Umrechnungsverfahren für Durchschnittsnoten nachgedacht werden. Der degressive Schlüssel wäre also nur bei der Bewertung der Prüfungsteile heranzuziehen, während bei der Berechnung der Gesamtnote der teil-lineare Schlüssel verwendet wird. Ehe man hier aber zu Entscheidungen kommt, wäre eine genaue Analyse des Problems auf der Basis systematischer empirischer Vergleiche und Simulationen zu empfehlen.

Anmerkungen:

¹ Vgl. u. a. Reisse, W.: *Integrierte Prüfung – Prüfungsform der Zukunft?* In: *Gewerkschaftliche Bildungspolitik*, Heft 11/12 (1996), S. 8–12; Reisse, W.: *Integrierte Prüfungen*. In: *Cramer, G.; Schmidt, H.*

Wittwer, W.: *Ausbilderhandbuch*. Köln 1996, Kap. 5.4.7; Schmidt, J. U. (Hrsg.): *Kaufmännische Prüfungsaufgaben – handlungsorientiert und komplex!?* Bielefeld 1997. Seyfried, B.: *Als Gruppe geprüft und beurteilt?* In: *Informationen für die Beratungs- und Vermittlungsdienste der Bundesanstalt für Arbeit*, Nr. 3723 (1995), S. 2985–2991

² In den §§ 41, 46, 47, 77, 81 und 95 Berufsbildungsgesetz/§§ 38, 42, 42 a und 50 Handwerksordnung ist festgelegt, daß bei Berufsausbildung, Fortbildung, Umschulung und Meisterprüfung von der zuständigen Stelle eine Prüfungsordnung zu erlassen ist. Zur Vereinheitlichung wurden vom Bundesausschuß für Berufsbildung Richtlinien für Prüfungsordnungen erlassen (Abschluß- und Gesellenprüfungen: Richtlinien vom 9. 6. 71, Umschulungsprüfungen: Richtlinien vom 2. 11. 71, Meisterprüfung: Richtlinien vom 18. 2. 72, Fortbildungsprüfungen: Richtlinien vom 18. 4. 73). Diese Richtlinien enthalten jeweils eine Musterprüfungsordnung, in der für die Bewertung von Prüfungsleistungen ein einheitliches Punktesystem für die Umrechnung von Punkten in Noten vorgegeben ist.

³ 4. Verordnung über eine Noten- und Punkteskala für die erste und zweite juristische Prüfung vom 3. 12. 81 (BGBl. I S.1243)

⁴ Die Daten einer Abschlußprüfung von Industriemechanikern wurden uns freundlicherweise von der PAL (Prüfungsaufgaben-Entwicklungsstelle bei der IHK Stuttgart) zur Verfügung gestellt.

⁵ Ecker: Teilnehmer von Kammerprüfungen benachteiligen. Unveröff. Manuskript. Homburg/Saar: Technisch-gewerbliches und Sozialpflegerisches Bildungszentrum. November 1995



Jens U. Schmidt (Hrsg.)

KAUFMÄNNISCHE PRÜFUNGS- AUFGABEN – HANDLUNGS- ORIENTIERT UND KOMPLEX!?

**KONZEPTE UND AUFGABENBEISPIELE FÜR
DIE NEUGESTALTUNG KAUFMÄNNISCHER
PRÜFUNGEN**

1997, 218 Seiten,
Bestell.- Nr. 102.204,
Preis 29.00 DM

Prüfungen im kaufmännischen Bereich sollen handlungsorientierter und komplexer werden. Es gilt, anstelle punktueller Wissensabfragen Prüfungsaufgaben zu entwickeln, die Arbeitsaufträge aus der beruflichen Praxis abbilden. Gleichzeitig sollen verstärkt Schlüsselqualifikationen berücksichtigt werden.

In dem vorliegenden Band stellen Experten aus unterschiedlichen Bereichen ihre Konzepte vor und erläutern diese anhand konkreter Aufgabenbeispiele. Dabei werden auch weitere Fragen angeschnitten, die für das Prüfungswesen von grundlegender Bedeutung sind, so die zentrale bzw. dezentrale Aufgabenerstellung, die Prüfungskosten, die Sicherstellung von Bezügen zu den Ausbildungsordnungen und Rahmenlehrplänen bei komplexen Aufgaben sowie die Gestaltung der praktischen Prüfung.



Ernst Ross (Hrsg.)

BERUFSBEZOGENES FREMDSPRACHENLERNEN

**NEUE KONZEPTIONEN, INHALTE,
METHODEN UND MEDIEN**

1997, 224 Seiten,
Bestell. Nr. 102.205,
Preis 29.00 DM

Der vorliegende Sammelband befaßt sich mit dem berufsbezogenen Fremdsprachenlernen. Er beinhaltet die Darstellung, Einschätzung und Diskussion von neuen Konzeptionen, Inhalten, methodischen und medialen Ansätzen für das Fremdsprachenlernen bei Auszubildenden, Facharbeitern und Fachangestellten. Dabei werden unterschiedliche Aspekte des Problemfeldes und der bildungspolitischen Diskussion um dieses Thema aufgegriffen. Im Mittelpunkt stehen neben einer allgemeinen Bestandsaufnahme Erfahrungen aus Pilotprojekten und Modellversuchen, neue Ansätze zum Fremdsprachenlernen bei Kammern und Firmen, Möglichkeiten und Grenzen des computer- und telekommunikationsgestützten Fremdsprachenlernens sowie Reaktionen und Erfahrungen von Schulen, Verlagen und Bildungsträgern.

► Sie erhalten diese Veröffentlichungen beim
W. Bertelsmann Verlag GmbH & Co. KG
Postfach 10 06 33
33506 Bielefeld
Telefon (0521) 911 01-0
Telefax (0521) 911 01-79