

## Bewertung von Lehrerprüfungsakten

- Referat für den Arbeitskreis Archivische Bewertung innerhalb des VdA –

-

Den folgenden Ausführungen liegt meine im Jahr 2001 an der Archivschule Marburg entstandene Transferarbeit „Die Überlieferung des ‚Wissenschaftlichen‘ und ‚Staatlichen‘ Prüfungsamtes (für Erste Staatsprüfungen für das Lehramt an Schulen) im NW Staatsarchiv Münster“ zugrunde.

Mein Referat gliedert sich in drei Punkte: 1. stelle ich Ihnen kurz das Verfahren der Lehrerprüfung in Nordrhein Westfalen und das dabei entstehende Schriftgut vor und 2. das aus der Analyse des Verfahrens heraus entwickelte Bewertungsmodell. Schließlich werde ich 3. einen Blick auf das Problem der ‚besonderen‘ Einzelfälle werfen. Bei meinen Ausführungen handelt es sich um prinzipielle Überlegungen, da das Verfahren der Lehrerprüfung in anderen Bundesländern (z. B. in Baden-Württemberg) trotz geringer organisatorischer Abweichungen ähnlich verläuft.

### 1. Die Verfahrensanalyse

Gemäß der 'Ordnung der Ersten Staatsprüfung für Lehrämter an Schulen' (LPO) soll durch das Prüfungsverfahren "festgestellt werden, ob das Studium erfolgreich verlaufen ist und die erziehungswissenschaftlichen, die fachwissenschaftlichen und fachdidaktischen Kenntnisse und Fähigkeiten erworben worden sind, die in dem betreffenden Lehramt für die Ausübung des Lehrerberufs erforderlich sind." (§ 3 Abs. 2 LPO) Das zu diesem Zweck etablierte Verfahren wird in seinem äußeren *Ablauf* durch folgende Programmskizze bestimmt:

Der Lehramtskandidat stellt beim zuständigen Staatlichen Prüfungsamt (§ 10 Abs. 1 LPO) einen Antrag auf Zulassung zur Ersten Staatsprüfung (§ 14 LPO). Das Prüfungsamt entscheidet über die Zulassung zur Prüfung (§ 15 LPO). Es beauftragt einen vom Kandidaten vorgeschlagenen Hochschullehrer mit der Themenstellung für die schriftliche Hausarbeit (§ 17 Abs. 2 LPO). Der Hochschullehrer erstellt ein Gutachten und gibt eine Note (§ 17 Abs. 6 LPO), die von einem vom Prüfungsamt benannten Zweitgutachter entweder bestätigt oder (gegebenenfalls unter Einbeziehung eines Drittgutachters) revidiert wird (§ 17 Abs. 7 LPO). Der Prüfling schreibt mehrere schriftliche Arbeiten unter Aufsicht (§ 18 LPO), für die jeweils eine die Note begründende Beurteilung von einem Erstgutachter abgefaßt und von einem Zweitgutachter wiederum bestätigt oder korrigiert wird. Der Kandidat absolviert daraufhin mehrere mündliche Prüfungen (§ 20 LPO), die vor einem vom Prüfungsamt berufenen Ausschuß stattfinden und über die jeweils schriftliche Protokolle mit Angabe einer Endnote angefertigt werden. Das Prüfungsamt ermittelt aus den Beurteilungen der schriftlichen

Hausarbeit, der Klausuren und der mündlichen Prüfungen Einzelnoten für die in den jeweiligen Fächern erbrachten Prüfungsleistungen (§ 25 LPO) sowie eine Gesamtnote für die Erste Staatsprüfung (§ 26 LPO). Über die bestandene Erste Staatsprüfung wird ein Zeugnis, über die nicht bestandene Prüfung eine Bescheinigung ausgestellt (§ 29 Abs. 1 LPO). Darüber hinaus kann eine Erweiterungsprüfung (§ 28 LPO) in einem zusätzlichen Fach angeschlossen sowie bei Nichtbestehen die Prüfung wiederholt werden (§ 27 LPO).

Das gesamte Verfahren findet seinen schriftlichen Niederschlag in der *Prüfungsakte*, die für jeden Lehramtskandidaten angelegt wird und im wesentlichen folgende Bestandteile enthält:

1. *vor Beginn des Verfahrens entstandene Unterlagen*: einen (mehreseitigen) Anmeldebogen mit persönlichen Daten, einen Lebenslauf (meistens fortlaufend, z. T. auch tabellarisch), dazu den sogenannten 'akademischen' Lebenslauf (Übersicht über besuchte Hochschulen und Zusammenstellung der Lehrveranstaltungen), Leistungsnachweise und eine Kopie des Abiturzeugnisses; 2. *während des Verfahrens entstehende Unterlagen*: die schriftliche Hausarbeit mit Gutachten, mehrere Klausuren mit Gutachten, Protokolle der mündlichen Prüfungen, Kopie des Zeugnisses über die bestandene Staatsprüfung bzw. der Bescheinigung über das Nichtbestehen; 3. (gegebenenfalls) *Sonstiges*: Korrespondenz zwischen dem Kandidaten und dem Prüfungsamt (betrifft zumeist Terminverschiebungen wegen Krankheit, Einreichung ärztlicher Atteste, Bitte um Verlängerung der Bearbeitungszeit für die Hausarbeit usw.) bzw. zwischen dem Prüfungsamt und einzelnen Prüfern (meistens im Zusammenhang mit Terminen und mit der Erstellung der verschiedenen Gutachten).

Betrachtet man das Prüfungsverfahren in seiner Gesamtheit, so enthält es Bestimmungen und Begriffe, die in unterschiedlich hohem Grade formalisiert sind und nach denen sich die Bestandteile der Prüfungsakte ebenfalls untergliedern lassen: Hochformalisiert sind die im Zulassungsantrag erhobenen Daten (Name, Adresse, Geburtsdatum usw.). Alles, was vom Prüfling hier anzugeben ist, wird im Vorhinein genau bezeichnet und es besteht aufgrund der Rasterung des Formulars auch keine Interpretationsmöglichkeit. Einzige Ausnahme bildet der Inhalt des Lebenslaufs, der in der Regel eine gewisse individuelle Prägung aufweist, wengleich auch hier bestimmte Elemente (z. B. Bildungsgang) immer wiederkehren und so zur Ausbildung von festen Konventionen der Lebensdarstellung führen. Ungleich detaillierter freilich als durch solche Konventionen ist im Verfahren festgelegt, welche Unterlagen dem Antrag beizufügen sind: Das Abiturzeugnis kann nicht durch den Führerschein ersetzt werden, die Leistungsnachweise nicht durch einfache Teilnahmebescheinigung; und auch die Zeugnisse und Bescheinigungen selbst enthalten ihrerseits wiederum fest umrissene Angaben (Namen, Noten, Daten usw.). Alles in allem stellt sich somit gerade derjenige Teil des

Prüfungsverfahrens, der sich mit den Zulassungsvoraussetzungen befasst, als ein weitgehend formalisierter Verfahrensabschnitt dar, der die Vielfalt der Situationen von vornherein strikt auf bestimmte Typen beschränkt und auf diese Weise in der Gesamtheit auch hohe Redundanzen produziert.

Anders verhält es sich mit den schwach formalisierten Komponenten des Prüfungsverfahrens: Die Varianzen beginnen hier schon mit der Bildung von *Prüfungsausschüssen*, die für jede einzelne Prüfung personell unterschiedlich zusammengesetzt sind (§ 11 LPO). Vielfalt besteht dann – trotz vorgegebener, aber weit gefaßter Rahmenbestimmungen – auch bei der Auswahl der *Themen* sowohl für die schriftliche Hausarbeit und die Klausuren, als auch für die mündliche Prüfung. Fast unbegrenzt ist schließlich die Gestaltungsfreiheit bei der inhaltlichen Ausfüllung dieser Themen, der jeweiligen Akzentsetzung, der zur Anwendung gelangenden Methoden und Interpretationen. So wird z. B. ein Vertreter der ‚politischen Geschichte‘ eine Prüfung anders akzentuieren als ein Anhänger der ‚Wirtschafts- und Sozialgeschichte‘. Die Prüfungsordnung selbst macht in diesen Punkten kaum Vorgaben. Trotzdem gibt es auch hier gewisse Regelmäßigkeiten, bestimmte Muster der Selektivität. Für diese Muster verantwortlich ist vor allem der jeweilige Prüfer als der aktiv Gestaltende des Verfahrens. Durch den Prüfer und seine Arbeitsschwerpunkte ist die Zahl der Themen, die der Prüfling wählen kann, von Anfang an mehr oder weniger stark eingeschränkt; viele Prüfer wirken schon vor der Anmeldung zur Prüfung darauf hin, dass der Kandidat mit seinen Vorschlägen innerhalb eines vorgegebenen thematischen Rahmens bleibt. Eingeschränkt ist auch der Prüfer, der im Rahmen der schriftlichen und vor allem der mündlichen Prüfung die Fragen zu den Themen vorstrukturiert und dabei methodisch nuanciert. Diese Feststellung bedeutet im Umkehrschluss natürlich, dass ein und derselbe Prüfer, bedingt durch seine jeweils individuellen fachlichen Präferenzen jede Prüfung – natürlich auf einer abstrakten Ebene des Vergleichs – *ähnlich* gestaltet, und zwar weitgehend unabhängig von der Persönlichkeit des jeweiligen Prüflings. Diese Ähnlichkeit verschiedener Prüfungen wird noch größer, je enger der zeitliche Abstand zwischen ihnen ist. So wird vermutlich ein und derselbe Prüfer, der an einem Tag mehrere Prüfungen (womöglich zu gleichen Themenschwerpunkten) abhält, seine Fragen 'ähnlicher' formulieren, als er dies bei thematisch vergleichbaren Prüfungen im Abstand von mehreren Jahren tun würde.

Individuelle, zeitbedingte Gewohnheiten und Standards betreffen nicht zuletzt die Form der Benotung, deren Kriterien auf der Ebene der Prüfungsordnung ebenso wie die Ausgestaltung der einzelnen Themen nur grob festgelegt sind. Auf diese Weise, d. h. durch die Interpretationsbedürftigkeit unbestimmter Wert- bzw. Notenbegriffe (eine "gute" Leistung

z. B. ist in der Prüfungsordnung ziemlich dehnbar definiert als eine solche, die "*erheblich* über den *durchschnittlichen* Anforderungen" liegt), ergeben sich im Kern des Verfahrens z. T. beträchtliche Subsumtionsprobleme. Um sie zu lösen, steht den einzelnen Prüfern ein recht großer Beurteilungsspielraum zur Verfügung; der Prüfer, der, wie bereits erläutert, die Themen, die methodische Ausrichtung der Fragen usw. selbst bestimmt, kann also ebenso individuell, d. h. in je spezifischer Art und Weise, über die Maßstäbe der Benotung entscheiden. Greifbar werden solche Unterschiede unter anderem in der Beobachtung, dass manche Prüfer bei den Studenten in dem Ruf stehen, 'streng' oder 'hart' zu benoten, während andere als 'mildere' oder weniger anspruchsvolle Prüfer gelten. Die Benotungsweise, die zwar nach individuellen, bei ein und demselben Prüfer jedoch meist nach immer ähnlichen Kriterien erfolgt, trägt also nicht unerheblich dazu bei, dass es auch in den Bereichen, in denen die LPO selbst auf Formalisierungen verzichtet, auf der Ebene der *informellen* Struktur zu einer gewissen *partiellen* Gleichförmigkeit kommt. Das bedeutet, dass es in der vom Programm zunächst offen gehaltenen Informationsverarbeitung trotzdem eine gewisse Regelmäßigkeit und folglich auch Redundanzen gibt. Aus dieser Feststellung heraus lässt sich das im folgenden vorgestellte Bewertungsmodell entwickeln.

## 2. Das Bewertungsmodell

Das Grundprinzip dieses Bewertungsmodells basiert (mit dem Ziel einer Abkehr von den vielfach auf die Übernahme ganzer Jahrgänge bzw. ausgewählter Buchstaben konzentrierten Auswahlverfahren) auf einer Orientierung am einzelnen Prüfer. Aufgrund der *pro Prüfer* entstehenden *partiellen* Redundanzen – jeder Prüfer vertritt *ein bestimmtes* Fach, prüft zudem *ausgewählte* Themen und Schwerpunkte *immer ähnlich* ab, hat *ein bestimmtes* System der individuellen Benotung, das er auf jeden Kandidaten *in etwa gleich* anwenden wird – ist es sinnvoll, Akten eines *jeden Prüfers* ins Archiv zu übernehmen. Bewahrt man auch nur eine Akte eines jeden Prüfers auf, so kann man schon von dieser einen Akte aus mit mehr als zufälliger Wahrscheinlichkeit auf die Beschaffenheit der anderen Akten desselben Prüfers zurückschließen. Anders gesagt: Ergebnisse, die später von einem Historiker oder einem anderen Forscher aus der Analyse der *einen* Akte des Prüfers gewonnen werden, lassen sich (wenn auch nur mit einem relativen Grad an Sicherheit) verallgemeinern. Allerdings erweist sich ein solcher Rückschluss von einer auf die anderen Akten unter zwei Bedingungen als problematisch: *a)* wenn es sich um Prüfungen für unterschiedliche Schulformen bzw. -stufen handelt, d. h. wenn andere Bestimmungen der LPO zur Anwendung gelangen. Daraus folgt konsequenterweise, dass *jeweils eine* (zufällig ausgewählte) Akte *eines* bestimmten Prüfers

für *jede* Schulform bzw. -stufe archiviert werden muss; *b*) wenn der zeitliche Abstand zwischen zwei übernommenen Akten zu groß wird, so dass sich zwischenzeitlich auf der Ebene des einzelnen Prüfers individuelle oder fachlich bedingte Verschiebungen (z. B. Schwerpunktverlagerungen oder methodische Weiterentwicklungen) vollzogen haben könnten. Um die gerade daraus sich ergebende Gefahr einer Verzerrung der Überlieferung so gut es geht zu kontrollieren, müssen die zeitlichen Schnitte zwischen den zu übernehmenden Akten möglichst dicht ausfallen. Bei dem konkret von mir bewerteten Bestand von Lehrerprüfungsakten im Staatsarchiv Münster erschien es (nicht zuletzt angesichts knapper Magazinressourcen) sinnvoll, für die Archivierung nur *jedes 3. Jahr* zu veranschlagen, dabei aber zur Optimierung des Verfahrens das 3. Jahr jeweils *pro Prüfer* zu rechnen, immer beginnend mit der Ernennung des jeweiligen Hochschullehrers zum Mitglied der Prüfungskommission. Dieses System der Übernahme führt zu einer zeitlich 'versetzten' Archivierung von Prüfungsakten, durch die im Regelfall (pro Fach) jedes Jahr mindestens einmal vertreten ist. Die dabei produzierte Archivierungsquote hält sich in einem relativ konstanten Rahmen (im Fall Münster lag sie bei knapp 15% im Mittelwert), die ich auch für vertretbar halte.

Die Auswahl durch das hier vorgestellte Bewertungsmodell – die empirische Sozialforschung spricht von einer 'disproportional geschichteten' Stichprobe – kann und will keine Repräsentativität im statistischen Sinne anstreben. Dennoch war sie im speziellen Fall die sinnvollste aller Möglichkeiten, zumal Registraturhilfsmittel (aus älterer Zeit eine Kartei der Prüflinge und seit 1994 eine elektronisch geführte Kandidatenverwaltung) vorhanden sind, die als Ergänzungsdokumentation verwendet werden können.

Bei aller Praktikabilität dieses Sampling-Verfahrens gilt es allerdings einen Aspekt nochmals besonders zu betonen: Bei den Ähnlichkeiten, die zwischen Akten ein und desselben Prüfers bestehen und die sich aus dessen individuellem 'Prüfungsstil' herleiten, handelt es sich immer nur um sozusagen *partielle* Redundanzen; die Regelmäßigkeiten des individuellen Prüfungsstils besitzen nicht die gleiche Rigidität und Selektivität wie die vollständige Programmierung hochformalisierter Verfahren (wie man sie z. B. aus dem Bereich der Steuerverwaltung kennt): bestimmte Themen und Methoden werden von einem bestimmten Prüfer lediglich *bevorzugt*, bestimmte Kriterien in der Bewertung nur *generell* stärker gewichtet als andere, ohne daß diese Konventionen den Charakter von 'Gesetzmäßigkeiten' annehmen. Genau wegen dieser, trotz der partiell entstehenden Redundanzen doch zugleich vorhandenen Individualität und Andersartigkeit eines jedes Falles, stellt sich bei den

Prüfungsakten – wie auch bei anderen Akten schwach formalisierter Verfahren – stets das Problem der 'besonderen', d. h. 'individuellen' Einzelfälle.

### 3. Das Problem der ‚besonderen‘ Einzelfälle

Die Relevanz und damit die Archivwürdigkeit von besonderen Einzelfällen lässt sich nicht über *formale* Regeln ermitteln, wie ich sie bislang mit Blick auf die Gesamtheit eines solchen Bestandes unter Berücksichtigung informationstheoretischer Überlegungen zum Verhältnis von Redundanz und Varietät erarbeitet und dem Bewertungsmodell einer disproportional geschichteten Stichprobe zugrunde gelegt habe. Die Wichtigkeit von besonderen Einzelfällen bemisst sich vielmehr nach *inhaltlichen* Kriterien, die abhängig vom jeweiligen Beobachterstandpunkt bzw. Erkenntnisinteresse variieren. Damit stellt sich für die Bewertung ein Perspektivismus-Problem, das objektiv letztlich unlösbar ist und einzig durch die Integration diskursiver Elemente in das Bewertungsverfahren, also gewissermaßen intersubjektiv überwunden werden kann. Nur ein (relativ offener) Meinungs-austausch der potentiellen Nutzer ist geeignet, um im Hinblick auf zu bewahrende Einzelfälle die Verallgemeinerbarkeit von Bewertungskriterien oder die Akzeptabilität von Partikularinteressen zu prüfen.

Unter dem Aspekt der Verallgemeinerbarkeit lassen sich vor allem folgende Kriterien zur Auswahl von Einzelfällen diskutieren, die deshalb so oder ähnlich wiederholt auch bereits bei der Archivierung von Lehrerprüfungsakten zur Anwendung gelangt sind:

1. *Landesgeschichtliche bzw. -kundliche* Arbeiten: Ein breiter Konsens ist offenbar in bezug auf eine zusätzliche Archivierung von landesgeschichtlichen Arbeiten zu erzielen. Nicht nur die Archivare selbst, sondern beispielsweise auch Repräsentanten von Politik und Öffentlichkeit sehen eine wichtige Aufgabe der Archive darin, landesgeschichtliche Forschung zu fördern. Allein die *Übernahme* entsprechender Staatsexamensarbeiten reicht freilich zu diesem Zweck nicht aus. Es muss zugleich auch die *Nutzung* solcher Arbeiten für die interessierte Forschung neu geregelt werden.
2. *Sonderfälle*: Unter dieser Bezeichnung sind alle Fälle zu fassen, in denen es Probleme bzw. Abweichungen im Programmablauf gegeben hat und diese Probleme möglicherweise sogar mit ausschlaggebend waren für etwaige verwaltungsinterne Programmänderungen. Akten, die solche Fälle dokumentieren, ließen sich wahrscheinlich am besten durch die Behörde selbst benennen, wie dies z. B. bei Justizakten bereits versucht wird. Ihre Archivierung wäre im Sinne eines 'Gedächtnisses' der Behörde zu verstehen, als ein

Hilfsmittel der kontinuierlichen 'Selbstbeobachtung'. Ein weniger geeignetes Kriterium der Sonderfallarchivierung scheint dagegen die Note des Examens zu sein. Im baden-württembergischen Modell beispielsweise werden in diesem Sinne "Prüfungsakten und Zulassungs- bzw. Prüfungsarbeiten der Kandidaten, die alle Prüfungsteile (wissenschaftliche Arbeit und Endnote der beiden Hauptfächer, ohne Erweiterungsprüfung) mit der Note 1,0 abgelegt haben", als archivierungswürdig eingeschätzt. Abgesehen von der Tatsache, dass solche Examina extrem selten abgelegt werden, so dass die Vermutung nahe liegt, ihre Archivierung sei vorrangig als Maßnahme gegen hohe Übernahmequoten gedacht, stellt sich hier vor allem die Frage, für *wen* letztlich die im Sinne einer 'positiven Wertauslese' betriebene Archivierung dieser Akten überhaupt von Interesse sein könnte. Zwar würde aus Sicht z. B. der Einzeldisziplinen die Archivierung herausragender Arbeiten für eine spätere wissenschaftliche Auswertung durchaus Sinn machen; die allgemeine Bildungsforschung hingegen, die ebenfalls Prüfungsakten als Quelle benutzt, bräuchte für ihre Forschungen sicherlich zusätzlich auch die Akten von besonders schlechten Prüfungen. Was als archivisch wertvoll gelten kann, ist also eine Frage der Perspektive; die sehr gute Note allein gibt kein sicheres Kriterium.

3. *Prominente Persönlichkeiten*: Ein wirklich operationalisierbares Kriterium für 'Prominenz' wird generell schwer festzulegen sein, zumal wenn der zeitliche Abstand zwischen dem Ablegen der Prüfung und der Übernahme der Akte ins Archiv nicht allzu groß ist. Kandidaten, die später bekannte Wissenschaftler, Politiker, Künstler o. ä. geworden sind, haben diese Karriere wohl in den meisten Fällen erst Jahre nach dem Bestehen der Lehramtsprüfung gemacht; oftmals wird auch die Bedeutung einer Persönlichkeit erst aus größerer zeitlicher Distanz heraus offensichtlich. Wenn sich also ein Diskurs über die Kriterien von 'Prominenz' nur schwer führen lässt und objektive Möglichkeiten einer Standardisierung nicht existieren, wäre es immerhin wünschenswert, wenn der Archivar sich wenigstens im Rückgriff auf eine Form von 'kondensierter' Öffentlichkeit (z. B. durch Heranziehung von Medienberichten oder biographischen Nachschlagewerken) versichern würde, dass nicht nur er allein eine Person als bedeutend einstuft. Falls dennoch der Archivar allein – was häufig der Praxis entspricht – Prominente ohne ausweisbare Kriterien in einem Bestand ausmacht, so sollte er diese Auswahl namentlich als seine persönliche kennzeichnen. Dies ist allerdings kaum mehr als eine Notlösung.

### Zusammenfassung:

Das Modell, das ich hier zur Bewertung von Lehrerprüfungsakten vorschlage, ist ein zweistufiges: die *erste* Stufe ist eine Verfahrensanalyse, die nach dem Verhältnis von Redundanzen und Varietäten fragt und die Möglichkeiten einer größtmöglichen Komprimierung prüft.

Die *zweite* Stufe des Bewertungsmodells sieht vor, dass dort, wo innerhalb des Verfahrens große Varietäten ausgemacht werden, wo also eine neutrale Komprimierung nicht möglich ist, nach den inhaltlichen Kriterien der Bewertung gefragt wird. Jedoch sollte sich die Diskussion inhaltlicher Kriterien immer an den Spezifika des jeweiligen Verfahrens festmachen.

Folgendes würde ich gern diskutieren: Wird das Verfahren von den Anwesenden anders gesehen (etwa weniger formalisiert)? Wie beurteilen Sie die Übertragbarkeit auf andere Prüfungsverfahren (Juristen, Mediziner, 2. Staatsexamen der Lehrer etc)? Falls es eine generelle Übertragbarkeit gibt, kennen Sie andere Verfahren, die den Charakter von offenen Prüfverfahren haben?