

- 8      **Gesundheit – Arbeit – Soziale Sicherheit**  
      **Santé – Travail – Sécurité sociale**  
      **Sanità – Lavoro – Sicurezza sociale**

**21**

Auszug aus dem Urteil der Abteilung II  
i. S. A. gegen Medizinalberufekommission  
B-4484/2009 vom 23. März 2010

**Medizinalprüfungen. Beschwerdeverfahren betreffend Prüfungsleistungen. Ausschluss von Prüfungsfragen.**

**Art. 10 Abs. 3 der Verordnung vom 30. Juni 1983 über Einzelheiten des Verfahrens bei den eidgenössischen Medizinalprüfungen.**

1. **Korrekte Antworten müssen als richtig bewertet werden, auch wenn sie nicht den in der Vorlesung behandelten Lerninhalt betreffen (E. 5.4).**
2. **Ermessen der Prüfungskommission bei der Bestimmung des Punkteverteilungssystems und der Festlegung der Notenskala (E. 6).**
3. **Nachträglicher Ausschluss einzelner Prüfungsfragen von der Bewertung der Prüfung. Offensichtlicher inhaltlicher oder formeller Mangel einer Prüfungsfrage verneint. Eine statistische Auffälligkeit, die bei der Auswertung der Prüfungsergebnisse festgestellt wurde, ist kein Mangel. Die Vorinstanzen müssen das Vorliegen eines Mangels substantiiert darlegen (E. 7).**
4. **Reformatorischer Entscheid, wenn die Vorinstanzen sich im Rahmen der Instruktion mit den streitentscheidenden Fragen noch einmal umfassend auseinandergesetzt haben und die Neubewertung der Prüfung aufgrund der Erwägungen des vorliegenden Urteils weitgehend vorgezeichnet ist (E. 8.5).**

**Examens des professions médicales. Procédure de recours en matière de prestation d'examen. Non-prise en compte de questions d'examen.**

**Art. 10 al. 3 de l'ordonnance du 30 juin 1983 réglant les modalités du procédé des examens fédéraux des professions médicales.**

1. Les réponses correctes doivent être considérées comme valables, même si elles ne se rapportent pas à la matière traitée durant les cours (consid. 5.4).
2. Pouvoir d'appréciation de la commission d'examen dans la détermination du système de répartition des points et dans l'établissement de l'échelle des notes (consid. 6).
3. Non-prise en compte ex post de certaines questions dans l'évaluation de l'examen. En l'espèce, pas de lacune manifeste d'une question quant au fond ou à la forme. Un résultat statistique inattendu constaté lors de l'évaluation des prestations d'examen ne constitue pas une lacune. Les instances précédentes doivent établir concrètement l'existence d'une telle lacune (consid. 7).
4. La décision attaquée peut être réformée lorsque les instances précédentes, dans le cadre de l'instruction de la cause, ont examiné encore une fois en détail les questions litigieuses et que la nouvelle évaluation de l'examen est dans une large mesure déjà déterminée par les considérants du présent arrêt (consid. 8.5).

**Esami per le professioni mediche. Procedura di ricorso in materia di prestazioni d'esame. Mancata presa in considerazione di domande d'esame.**

**Art. 10 cpv. 3 dell'ordinanza del 30 giugno 1983 che regola le particolarità della procedura degli esami federali per le professioni mediche.**

1. Le risposte corrette devono essere considerate valide, anche se non si riferiscono ai programmi d'insegnamento trattati durante i corsi (consid. 5.4).
2. Potere d'apprezzamento della commissione d'esame nella determinazione del sistema di attribuzione dei punti e della scala di valutazione (consid. 6).
3. Mancata presa in considerazione ex post di determinate domande nella valutazione dell'esame. Negazione della nozione di domanda d'esame manifestamente lacunosa quanto al fondo o alla forma. Un risultato statistico inatteso, che è stato constatato

**nell'ambito della valutazione dei risultati d'esame, non costituisce una lacuna. Le autorità inferiori devono dimostrare in modo sostanziale l'esistenza di tale lacuna (consid. 7).**

- 4. Decisione riformatoria, in quanto le autorità inferiori, nell'ambito dell'istruzione della causa, si sono occupate ancora una volta a fondo della questione litigiosa e la nuova valutazione dell'esame è già indicata in gran parte dai considerandi della presente decisione (consid. 8.5).**

Die Beschwerdeführerin legte im Sommer 2008 die Prüfung des ersten Studienjahres des Studiengangs für Veterinärmedizin an der Universität Zürich ab. Mit Verfügung vom 16. Juli 2008 teilte ihr die Präsidentin der Prüfungskommission für Veterinärmedizin der Universität Zürich (Erstinstanz) mit, dass sie die Prüfung nicht bestanden habe. Sie werde von sämtlichen weiteren Prüfungen derselben Berufsart endgültig ausgeschlossen. Die Verfügung basierte massgeblich auf der Bewertung der Einzelprüfung Nr. IV, die als nicht bestanden gewertet wurde und für welche die Beschwerdeführerin in der Folge keine Kreditpunkte erhielt. In dieser Prüfung erhielt die Beschwerdeführerin 51 Punkte und hätte mindestens 53 Punkte erreichen müssen, um die Prüfung zu bestehen. Es konnten maximal 99 Punkte erreicht werden. Da nachträglich vier Fragen von der Bewertung ausgeschlossen wurden, reduzierte sich die höchstmögliche Punktzahl entsprechend. Die in der Einzelprüfung enthaltenen Fragen waren sowohl nach dem Wahlantwortverfahren als auch nach dem Kurzfragen-Kurzantworten-Verfahren konzipiert.

Am 2. August 2008 reichte die Beschwerdeführerin gegen diesen Entscheid Beschwerde bei der Medizinalberufekommission (MEBEKO, Vorinstanz) ein, welche die Beschwerde mit Entscheid vom 17. Juni 2009 abwies.

Die Beschwerdeführerin erhebt am 13. August 2009 Beschwerde beim Bundesverwaltungsgericht (BVGer). Sie rügt in ihrer Beschwerdeschrift und in ihrer Replik vom 13. November 2009, dass die Einzelprüfung Nr. IV unterbewertet sei. Des Weiteren bringt sie vor, bei korrekter Rechtsanwendung hätten die ausgeschlossenen Prüfungsfragen in die Bewertung einbezogen und umgekehrt eine Frage ausgeschlossen werden müssen, die unstreitig einen Druckfehler enthalte und von der Beschwerdeführerin falsch beantwortet worden sei. Ferner beanstandet sie in mehrfacher Hinsicht das bei der Prüfung verwendete Notengebungssystem.

In ihrer Vernehmlassung vom 16. September 2009 beantragt die Vorinstanz die Abweisung der Beschwerde.

Sie macht insbesondere geltend, es könne nur aufgrund einer Auswertung, die sich auf die Leistung aller Kandidatinnen und Kandidaten in einer Einzelprüfung beziehe, entschieden werden, ob eine Frage zu eliminieren sei. Namentlich sei die Tatsache, dass viele Studierende mit guten Prüfungsergebnissen eine Frage falsch beantworteten, ein sicheres Indiz dafür, dass eine Frage mit einem gravierenden Mangel behaftet sei. Sofern die Beschwerdeführerin das Punkteverteilungssystem bei Mehrfach-Wahl-Items beanstande, sei zu bedenken, dass die Leistungsbeurteilung grundsätzlich im Ermessen der Experten liege.

Mit Instruktionsverfügung vom 2. Dezember 2009 wurden die Vorinstanz und die Erstinstanz vom BVGer eingeladen, für jede der eliminierten Fragen im Einzelnen darzulegen, welche Gründe jeweils für den Ausschluss der Prüfungsfragen massgeblich waren und wie die Prüfung zu bewerten gewesen wäre, wenn die eliminierten Fragen bei der Bewertung berücksichtigt worden wären.

Mit Schreiben vom 13. Januar 2010 nahm die Vorinstanz Bezug auf die Stellungnahmen der Prüfungsexperten der Universitäten Bern und Zürich vom 17. Dezember 2009 beziehungsweise 21. Dezember 2009. Die Prüfungsexperten äusserten sich zu der Bewertung der oben genannten Prüfungsfrage und brachten im Übrigen vor, die eliminierten Fragen seien wegen ihrer statistisch auffälligen Werte ausgeschlossen worden. Sie zeigten eine geringe oder gar negative Trennschärfe, worunter die Fähigkeit einer Frage zu verstehen sei, die Kandidaten mit guter und schlechter Leistung in der Gesamtprüfung zu trennen. Wenn keine Frage eliminiert worden wäre, käme die Bestehensgrenze auf 54 Punkte zu liegen.

Das BVGer heisst die Beschwerde gut.

*Aus den Erwägungen:*

**5.** Die Beschwerdeführerin rügt weiterhin, die Beurteilung ihrer fachlichen Prüfungsleistung sei in mehreren Aufgaben der Einzelprüfung Nr. IV fehlerhaft erfolgt. Sie begehrt, die Bewertung der Prüfung entsprechend zu korrigieren, die Note in dieser Einzelprüfung anzuheben und die Leistungskontrollen der Prüfung des ersten Studienjahres der Vetsuisse-Fakultäten als bestanden zu bewerten.

**5.1** Wie dargelegt (...) überprüft das BVGer die Bewertung von Examenleistungen nur mit Zurückhaltung und weicht nicht ohne Not von der Beurteilung der erstinstanzlichen Prüfungsorgane ab. Den Examinatoren kommt bei der Beurteilung der Frage, ob ein Kandidat eine Prüfungsaufgabe richtig gelöst hat und welche Antworten als vertretbare Lösungen in Betracht kommen, ein grosser Beurteilungsspielraum zu. Es kann daher nicht Aufgabe der Beschwerdeinstanz sein, die gesamte Bewertung der Prüfung in den fraglichen Fächern gewissermassen zu wiederholen. Daraus folgt, dass die Rügen eines Beschwerdeführers, wonach die Bewertung seiner Prüfungsleistungen offensichtlich unangemessen gewesen sei, von objektiven Argumenten und Beweismitteln getragen sein müssen. Ergeben sich solche eindeutigen Anhaltspunkte nicht bereits aus den Akten, so muss der Beschwerdeführer selber substantiierte und überzeugende Anhaltspunkte dafür liefern, dass eindeutig zu hohe Anforderungen gestellt oder die Prüfungsleistung offensichtlich unterbewertet wurden. Er wird den Anforderungen an eine genügende Substantiierung seiner Rügen insbesondere dann nicht gerecht, wenn er sich einfach darauf beschränkt zu behaupten, seine Lösung sei vollständig und korrekt, ohne diese Behauptung näher zu begründen oder zu belegen. Sofern es ihm hingegen gelingt, eine Fehlbewertung seiner Prüfungsleistung in dieser Weise zu substantiieren, ist es wiederum Sache der Examinatoren, im Einzelnen und in nachvollziehbarer Weise darzulegen, warum eine Lösung des Beschwerdeführers falsch oder unvollständig ist und er daher nicht die Maximalpunktzahl erhalten hat (vgl. zu alledem das Urteil des BVGer B-4771/2008 vom 15. April 2009 E. 5.1).

**5.2** Die Beschwerdeführerin macht geltend, ihre Lösung der Aufgaben A 13, A 19, E 5, E 10, E 11, E 15 sei als falsch bewertet worden, obwohl ihre Antwort korrekt oder zumindest vertretbar gewesen sei. Es handelt sich hierbei um Aufgaben, die nach dem Kurzfragen-Kurzantworten-Verfahren konzipiert sind. Die Beschwerdeführerin begründet ihre Auffassung im Hinblick auf die einzelnen Aufgaben und stützt sich jeweils auf Belege aus der Fachliteratur oder aus Vorlesungsskripten. Die Rügen sind daher genügend substantiiert.

**5.3** Indessen haben die Experten, was die Aufgaben mit Ausnahme von Prüfungsfrage E 10 betrifft, ihrerseits detailliert begründet, weshalb die Lösungen der Beschwerdeführerin nicht den Anforderungen der Aufgaben entsprachen. Sie konnten nachvollziehbar darlegen, dass ihre Korrektur nicht im Widerspruch zu den von der Beschwerdeführerin angeführten Quellen steht. Die Beschwerdeführerin hat dagegen ihrerseits nichts vorgebracht, um diese Darlegung der Experten als offensichtlich unzutreffend zu widerlegen. Unter Berücksichtigung der eingeschränkten

Kognition des BVGer bei der Bewertung von fachlichen Prüfungsleistungen kann eine offensichtliche Unterbewertung der Aufgaben A 13, A 19, E 5, E 11 und E 15 folglich nicht festgestellt werden.

**5.4** Anders verhält es sich hingegen bei der Bewertung der Aufgabe E 10. In dieser Aufgabe lautete die Fragestellung: « Nennen Sie 2 Organe, in denen funktionelle Einheiten mit einem extrem flachen Plattenepithel ausgekleidet sind. » Es können in dieser Aufgabe maximal zwei Punkte erreicht werden. Die Beschwerdeführerin gab als Antwort « Ovar » und « Milz » an, wofür sie keine Punkte erhielt und ist der Auffassung, dass die Antwort « Ovar » korrekt sei, weshalb ihr mindestens ein Punkt hätte erteilt werden müssen. Sie beruft sich auf ein Vorlesungsskript, in dem es heisst: « Im Ovar sind also folgende Follikel-Typen zu finden: Primordialfollikel: [...] Follikelepithel ist flach und ebenfalls wenig differenziert [...] ». In der Stellungnahme der Prüfungsexperten vom 10. November 2008 wird vertreten, dass die richtige Antwort « Lunge und Niere » sei. Die Begründung lautet: « Diese beiden Organe sind, im Gegensatz zu Milz und Ovar, in der Vorlesung und im Kurs behandelt worden. Milz und Ovar können daher nicht Grundlage für die Prüfung im 1. JK sein. » Diese Begründung ist nicht nachvollziehbar und daher offensichtlich ungeeignet, den bei der Beurteilung der Prüfung erfolgten Punktabzug zu rechtfertigen. Es versteht sich von selbst, dass die Richtigkeit der Lösung einer Prüfungsaufgabe nicht davon abhängen kann, ob diese Frage vorher in einer Vorlesung thematisiert wurde. Angesichts dieser wenig aufschlussreichen Begründung des Prüfungsexperten ersuchte das BVGer die Vorinstanz und die Erstinstanz in ihrer Instruktionsverfügung vom 2. Dezember 2009 noch einmal dazu Stellung zu nehmen, ob Aufgabe E 10 richtig oder falsch gelöst worden sei. Die Erstinstanz räumt in ihrer Stellungnahme vom 21. Dezember 2009 ein, dass die Antwort der Beschwerdeführerin zumindest weit überwiegend eben doch der in der Fragestellung angegebenen Definition entspreche. Sie führt hierzu aus, die Antwort der Beschwerdeführerin sei insofern richtig, als es sich bei den Primordialfollikeln um funktionelle Einheiten des Ovars handle und dass diese ein dünnes Epithel besitzen. Die Erstinstanz meint jedoch, die Antwort der Beschwerdeführerin sei dennoch als falsch zu bewerten und argumentiert, das Epithel der Primordialfollikel sei nicht « extrem » flach im Sinne des Aufgabentextes, zudem sei die Dicke des Epithels in den funktionellen Einheiten von Niere und Lunge bedeutend für deren Funktion, worauf in der Vorlesung ausdrücklich hingewiesen worden sei. Dies ist jedoch wenig überzeugend. Wie die Erstinstanz selbst anführt, kann man darüber streiten, wie der recht unbestimmte Begriff « extrem » im Kontext der Aufgabenstellung zu verstehen ist. Es erscheint in der Tat

übermässig spitzfindig, die ansonsten nunmehr unbestritten korrekte Antwort allein aufgrund des Fehlens dieses wenig präzisen Definitionsmerkmals als falsch zu bewerten. Welche Bedeutung die Dicke des Epithels für die Funktion eines bestimmten Organs hat, war zudem nicht Gegenstand der Aufgabenstellung. Wie bereits erwähnt, kann es ferner für die Richtigkeit der Antwort nicht von Bedeutung sein, ob die betreffende Frage in der Vorlesung behandelt wurde. Zusammenfassend kann somit festgestellt werden, dass die Beschwerdeführerin substantiiert darlegen konnte, dass ihre Lösung der Aufgabe E 10 zu Unrecht als falsch gewertet wurde. Die Prüfungsexperten waren hingegen nicht in der Lage, in nachvollziehbarer Weise zu erklären, weshalb die Bewertung dieser Aufgabe fehlerfrei erfolgt ist. Ihr Vorbringen spricht im Gegenteil dafür, dass die Antwort der Beschwerdeführerin zumindest weitgehend korrekt war. Unter diesen Umständen muss – auch unter Berücksichtigung der Grundsätze über die Einschränkung der Kognition bei der Beurteilung von fachlichen Prüfungsleistungen – angenommen werden, dass insofern eine offensichtliche Fehlbewertung vorliegt. Die Rüge der Beschwerdeführerin erweist sich daher im Hinblick auf Aufgabe E 10 als begründet.

6. Die Beschwerdeführerin beanstandet ferner das Punkteverteilungssystem. Bei den Mehrfach-Wahl-Items finde keine faire, proportionale Abstufung bei der Punkteverteilung statt, in der für jede richtige Antwort Punkte vergeben würden. Es verstosse zudem gegen das Willkürverbot und gegen den Grundsatz der Rechtsgleichheit, wenn die Notenspannweite oberhalb der Note 4 sechs Punkte betrage und unterhalb der Note 4 sieben Punkte. Diesbezüglich geht die Vorinstanz zutreffend davon aus, dass es Aufgabe der Prüfungskommission ist, die Modalitäten des Punkteverteilungssystems zu bestimmen und insbesondere die Notenskala festzulegen (vgl. BVerGE 2010/10 E. 5.2). Die Prüfungskommission hat hierbei einen gewissen Entscheidungsspielraum. Da nicht erkennbar ist, dass sich die Prüfungskommission bei der Festlegung der Notengrenzen und der Festlegung der Punkteverteilung von sachfremden Erwägungen leiten liess, kann nicht davon ausgegangen werden, dass sie willkürlich handelte. Die Rüge der Beschwerdeführerin ist daher unbegründet.

7. Die Beschwerdeführerin rügt des Weiteren, es seien vier Aufgaben, von denen sie zwei unbestritten korrekt gelöst habe, ohne ersichtlichen Grund von der Bewertung ausgeschlossen worden. Demgegenüber berufen sich Vorinstanz und Erstinstanz darauf, dass die eliminierten Fragen wegen ihrer statistisch auffälligen Werte ausgeschlossen worden seien. Bei der Auswertung der Prüfungsergebnisse aller Kandidaten sei aufgefallen, dass diese Aufgaben eine geringe oder gar negative Trennschärfe hätten. Darunter sei die Fähigkeit einer Frage zu verstehen, die Kandida-

ten mit guter und schlechter Leistung in der Gesamtprüfung zu trennen. Diese statistische Auffälligkeit stelle ein starkes Indiz für das Vorliegen eines Mangels dar, der den Ausschluss der jeweiligen Frage rechtfertige.

**7.1** Die Rüge der Beschwerdeführerin betrifft nicht die Bewertung ihrer Prüfungsleistung, sondern die Bewertungsmaßstäbe, die für alle Kandidaten gelten, die an der betreffenden Prüfungssession teilgenommen haben. Sie ist daher mit voller Kognition zu prüfen.

**7.2** Gemäss Art. 10 Abs. 3 der Verordnung vom 30. Juni 1983 über Einzelheiten des Verfahrens bei den eidgenössischen Medizinalprüfungen (SR 811.112.18, nachfolgend: Prüfungsverfahrensverordnung) werden Fragen oder Antworten bei der Bewertung nicht berücksichtigt, wenn sie einen *offensichtlichen inhaltlichen oder formalen Mangel* erkennen lassen. Dieser Vorschrift liegt die Überlegung zugrunde, dass es einerseits dem Zweck der Prüfung, den Kenntnisstand der Kandidaten mit hinreichender Genauigkeit zu überprüfen, zuwider laufen würde, wenn solche Fragen in die Bewertung einbezogen würden. Zudem kann der Ausschluss von erkennbar fehlerhaften Prüfungsfragen dazu dienen, dass die Prüfungen verschiedener Prüfungssessionen einen weitgehend einheitlichen Schwierigkeitsgrad haben. Indessen kann der Ausschluss von Prüfungsfragen zu Ungleichbehandlungen führen, weil einerseits Kandidaten, die diese Fragen korrekt beantwortet haben, durch den Ausschluss einen Verlust an Punkten erleiden und andererseits sich die Gesamtleistung von Kandidaten möglicherweise verbessert, wenn eine Frage eliminiert wird, die sie falsch beantwortet haben. Der Ausschluss von Prüfungsfragen darf daher nicht willkürlich erfolgen, sondern muss auf einem sachlichen Grund beruhen. Art. 10 Abs. 3 Prüfungsverfahrensverordnung bestimmt daher, dass der nachträgliche Ausschluss einzelner Prüfungsfragen zwar nicht von vornherein unzulässig ist, allerdings nur im Ausnahmefall und unter bestimmten Voraussetzungen erfolgen darf. Fraglich ist, ob diese Voraussetzungen vorliegend erfüllt sind. Es ist zu prüfen, ob die ausgeschlossenen Prüfungsfragen einen Mangel erkennen lassen und ob dieser Mangel als offensichtlich anzusehen ist.

**7.3** Ob ein inhaltlicher oder formaler Mangel im Sinne von Art. 10 Abs. 3 Prüfungsverfahrensverordnung vorliegt, kann nur in Anbetracht der jeweiligen Aufgabenstellung beurteilt werden. Ein solcher Mangel ist beispielsweise anzunehmen, wenn die Aufgabenstellung missverständlich formuliert, unzumutbar schwierig oder gar unlösbar ist, des Weiteren, wenn sie ausserhalb des Curriculums liegt (so auch die Erstinstanz [...]) und nachfolgend E. 7.3.1).

**7.3.1** Im bisherigen Verfahren äusserten sich die Prüfungsexperten zur Frage des Ausschlusses von Prüfungsfragen vor der Vorinstanz im Einzelnen wie folgt:

« Bei der Auswertung der Prüfung werden sämtliche Prüfungsfragen [...] analysiert. Dabei werden die aus psychometrischer Sicht auffälligen Fragen jeweils den Dozenten zur Beurteilung zugestellt. Dabei wird überprüft, ob ein Schlüsselfehler vorliegt, oder ob die Frage aus Sicht des Dozenten eliminiert werden muss [...] » (...);

« Entsprechend machen wir [die Abteilung für Assessment und Evaluation der Erstinstanz] in einem ersten Auswertungsschritt die Examinatoren aufgrund der statistischen Fragenanalyse auf mögliche Problemfragen aufmerksam. Sie entscheiden aufgrund der fachlichen Überprüfung, welche Fragen vor der Ermittlung und Benotung der Kandidatenleistung zu eliminieren sind. » (...);

« Die Entscheidung, ob eine aus psychometrischer Sicht auffällige Frage eliminiert werden soll, wird wiederum durch die Dozierenden persönlich (schriftlich) vorgenommen » (...).

Sodann äusserten sie sich im Verfahren vor dem BVGer wie folgt:

« Frage 4 [...]: Diese Frage war aufgrund der fehlenden Trennschärfe aufgefallen und sollte vom Dozenten vor einer weiteren Verwendung überarbeitet werden. Der Dozent entschied sich dann aber, diese Frage auch schon in der laufenden Prüfung zu eliminieren (Kommentar [...]: « Die Frage ist offensichtlich doch etwas schwierig, obwohl der Unterschied zwischen Blastozyste und Blastula erklärt wurde. ») *Eine Elimination der Frage war aus messtechnischen Gründen nicht zwingend nötig*. Da aber die Trennschärfe schwach und der p-Wert sehr tief ist [...], ist eine Elimination trotzdem gerechtfertigt.[...]

Frage 19 [...]: [...] Diese Frage wurde aufgrund der schlechten Trennschärfe eliminiert. Die Vorlesung wurde in diesem Jahr von einem anderen Dozenten gehalten, als vom Fragenautor; infolgedessen *kann es sein, dass der diesjährige Dozent nicht so detailliert auf diesen Lehrinhalt eingegangen ist wie der Dozent im Vorjahr*. [...]

Frage 40 [...]: Diese Frage muss aus messtechnischer Sicht entfernt werden (Kommentar [...]: « Schwieriges Item mit negativer Trennschärfe. 19,8 und 16,7 % der KandidatInnen haben die Antworten C und D gewählt. Diese KandidatInnen haben im Mittel eine bessere Prüfungsleistung als die Richtigantwortenden »). Aus diesem Grund ist eine Elimination gerechtfertigt).[...]

Frage 43 [...]: Diese Frage muss aufgrund der fehlenden Trennschärfe eliminiert werden (Kommentar [...]: « Sehr schwieriges Item ohne Trennschärfe. 32,3 % der KandidatInnen haben Antwort C gewählt. Diese haben im Mittel eine bessere Prüfungsleistung erzielt als die

*Richtigantwortenden.*) Aufgrund des geringen p-Werts und der fehlenden Trennschärfe muss angenommen werden, dass die Studierenden geraten haben. [...] › (...). »

Es ist augenfällig, dass weder der Umstand, dass ein Dozent eine gewisse Thematik *möglicherweise* weniger einlässlich behandelt haben könnte, als ein anderer Dozent im Vorjahr, kein hinreichend bestimmtes, formales oder inhaltliches Ausschlusskriterium im Sinne der genannten Verordnungsbestimmung darstellt (betrifft Frage 19). Gleich verhält es sich mit dem Umstand, dass 16,7 bzw. 19,8 % (aber auch 32 %) der Kandidaten eine gestellte Frage falsch beantworteten, « *im Übrigen aber im Mittel eine bessere Prüfungsleistung [haben] als die Richtigantwortenden* ». Auch hier fehlt es an einem hinreichend bestimmten, formalen oder inhaltlichen Ausschlusskriterium (betrifft Fragen 40 und 43). Ein rechtsgenügender Ausschlussgrund ist auch dort klar zu verneinen, wo die Prüfungsexaminatoren selber ausführen, eine *Elimination* der Frage sei aus messtechnischen Gründen *nicht zwingend erforderlich* (betrifft Frage 4). Hierauf ist nachfolgend zurückzukommen.

**7.3.2** Die Vorinstanz macht geltend, ein Mangel sei allein aufgrund der (unbestrittenen) Tatsache anzunehmen, dass bei der Auswertung der Prüfungsergebnisse aller Kandidaten im Hinblick auf die betreffenden Aufgaben eine statistische Abweichung, namentlich eine geringe Trennschärfe festgestellt worden sei. Im angefochtenen Beschwerdeentscheid (E. 11.6, 2. Lemma) führt sie aus:

« Für die MEBEKO, Ressort Ausbildung, sind die Gründe, die im Einzelfälle zu einer Elimination führen, nicht entscheidungswesentlich. Ausschlaggebend für die Elimination ist alleine schon die Tatsache, dass die Auswertung der einzelnen Fragen ein auffälliges Ergebnis zeitigt. Als statistisch auffällig und daher potentiell zu eliminieren sind insbesondere Fragen, die ungenügend diskriminieren, das heisst zu wenig zwischen Studierenden mit guten und weniger guten Kenntnissen unterscheiden » (...).

Sie geht zutreffend davon aus, dass diese Auffälligkeit ein Indiz für das Vorliegen eines Mangels sein kann. Es kann nicht ausgeschlossen werden, dass der Grund für die geringe Trennschärfe einer Aufgabe in der Fehlerhaftigkeit der betreffenden Prüfungsfrage besteht. Eine derartige statistische Auffälligkeit sagt jedoch unmittelbar nichts darüber aus, von welchem Mangel die jeweilige Aufgabe betroffen und wie dieser beschaffen ist. Selbst wenn sie das Vorliegen eines Mangels indiziert, ist es Sache der Vorinstanzen, im Rahmen ihrer Pflicht zur Sachverhaltsaufklärung konkret zu benennen, welcher Mangel ihres Erachtens gegeben ist und inwiefern sich dieser auf die betreffende Auffälligkeit bei der Aus-

wertung der Prüfungsergebnisse ausgewirkt hat. Die Vorinstanzen sind indessen vom Vorliegen eines Mangels ausgegangen, ohne Bezug auf die einzelnen Aufgabenstellungen zu nehmen. Wie oben (E. 7.3.1) dargelegt, führen die Prüfungsexperten in ihren Stellungnahmen hinsichtlich der meisten Prüfungsfragen lediglich aus, die Fragen seien « aus messtechnischer Sicht », insbesondere aufgrund fehlender oder geringer Trennschärfe auszuschliessen gewesen. Man könne davon ausgehen, dass « die Studierenden geraten haben ». Ferner könne es sein, dass der diesjährige Dozent nicht so detailliert auf den betreffenden Lehrinhalt eingegangen sei, wie der Dozent im Vorjahr (...). In keinem Fall gehen die Experten jedoch auf den konkreten Inhalt der eliminierten Aufgaben ein. Sie können nicht substantiiert darlegen, weshalb die ausgeschlossenen Prüfungsfragen ihrer Ansicht nach inhaltlich oder formal mangelhaft sind und wie dies im Zusammenhang mit der festgestellten statistischen Auffälligkeit steht. Solches ist auch für das BVGer nicht ersichtlich. Eine mehr oder minder ausgeprägte statistische Auffälligkeit für sich alleine, welche auch aus reinem Zufall aufgetreten sein könnte, stellt indessen keinen « Mangel » dar, der gemäss Art. 10 Abs. 3 Prüfungsverfahrensverordnung den nachträglichen Ausschluss einer Prüfungsfrage gestattet. Vielmehr ist zur Feststellung eines Mangels eine Einzelfallbetrachtung, die auf die Aufgabenstellung der zu eliminierenden Prüfungsfrage eingeht und in der die für den Ausschluss der Frage massgeblichen Gründe nachvollziehbar dargelegt werden, zwingend erforderlich. Eine solche wurde – nachdem aus den Vorakten kein hinreichender « Mangel » im Sinne der Prüfungsverfahrensverordnung ersichtlich geworden war – vom BVGer mit Verfügung vom 2. Dezember 2009 einverlangt, konnte aber – wie dargelegt – auch in diesem Verfahrensstadium nicht in genügend substantiierter Weise erbracht werden. Damit ist mit der Beschwerdeführerin davon auszugehen, dass es vorliegend an hinreichenden Gründen fehlt, die den nachträglichen Ausschluss der streitbezogenen Fragen gerechtfertigt hätten. Dies umso weniger, als Art. 10 Abs. 3 Prüfungsverfahrensverordnung das Vorliegen eines *offensichtlichen* Mangels verlangt, wovon nach dem Gesagten auch nicht ansatzweise ausgegangen werden kann.

**7.3.3** Anzumerken bleibt in diesem Zusammenhang, dass es für den Umfang einer solchen Einzelfallprüfung und die Kriterien, die bei dieser Prüfung zu beachten sind (vgl. insofern die Ausführungen unter E. 7.3), auch nicht massgebend sein kann, um welchen Aufgabentypus es sich jeweils handelt. Es ist nicht ersichtlich und ergibt sich auch nicht aus den massgebenden Verfahrensvorschriften, weshalb etwa für die Fehlerhaftigkeit einer Frage vom Typ Wahlantwortverfahren wesentlich andere

Massstäbe gelten sollten, als bei einer Frage, die nach Kurzfragen-Kurzantworten-Verfahren konzipiert ist. In beiden Fällen geht es darum, eine knappe Antwort auf eine präzise Frage zu finden, wobei im Fall des Wahlantwortverfahrens lediglich die Besonderheit besteht, dass mehrere Antworten im Aufgabentext vorgegeben werden und der Kandidat aus diesen die Korrekten auswählen muss. Dies wird indessen von keiner Seite in Abrede gestellt, so dass sich an dieser Stelle Weiterungen hierzu erübrigen.

Was schliesslich den Schwierigkeitsgrad der einzelnen Prüfungsfragen betrifft, liegt es in der Natur einer Prüfung, dass sie sowohl leichtere als auch schwierigere Aufgaben enthält. Von einem « offensichtlichen » Mangel aufgrund eines hohen Schwierigkeitsgrades sollte daher nur ausgegangen werden, wenn die Schwierigkeit der Aufgabe so unzumutbar hoch ist, dass von einem durchschnittlichen Kandidaten nicht erwartet werden kann, dass er sie richtig löst. Dies ist vorliegend soweit ersichtlich nicht der Fall und wurde denn auch von keiner Seite geltend gemacht.

**7.4** Es ist somit davon auszugehen, dass die Voraussetzungen für den Ausschluss der Prüfungsfragen nicht vorliegen. Die Vorinstanz hätte die Fragen deshalb bei der Bewertung berücksichtigen müssen. Die Rüge der Beschwerdeführerin erweist sich daher insofern als begründet.

**7.5** Die Beschwerdeführerin rügt schliesslich, dass eine Prüfungsfrage nicht von der Bewertung ausgeschlossen worden sei, obwohl die Aufgabenstellung fehlerhaft formuliert sei. Die Prüfungsfrage 13 enthält unbestritten einen Schreibfehler (« p2 »; korrekt gewesen wäre « p<sup>2</sup> »). Das Vorbringen der Vorinstanz, die Fragestellung könne trotz des Schreibfehlers ohne Weiteres richtig interpretiert werden, erscheint nicht von vornherein fernliegend. Es wird durch den Umstand gestützt, dass kein auffälliger Anteil der Kandidaten die Aufgabe falsch gelöst hat. Da zudem der Ausschluss einer Prüfungsfrage nur in eng begrenzten Ausnahmefällen erfolgen darf, erscheint es als vertretbar, dass die Prüfungskommission die Aufgabe nicht von der Bewertung ausgeschlossen hat.

**8.** Zusammenfassend kann festgehalten werden, dass die Rügen der Beschwerdeführerin insofern begründet sind, als sie die Bewertung ihrer Prüfungsleistung in Aufgabe E 10 und den Ausschluss von Prüfungsfragen beanstandet. Der Beschwerdeführerin fehlen zwei Punkte, um die Prüfung zu bestehen. Aus den unbestrittenen Darlegungen der Prüfungsexperten ergibt sich, dass sie, falls die Prüfungsfragen nicht ausgeschlossen worden wären, zwei zusätzliche Punkte erhalten hätte und die Bestehensgrenze gleichzeitig um einen Punkt nach oben verschoben worden

wäre. Daraus folgt, dass die Beschwerdeführerin die Prüfung besteht, wenn man ihr einen zusätzlichen Punkt für Aufgabe E 10 erteilt und zudem die ausgeschlossenen Fragen berücksichtigt. Die Rügen der Beschwerdeführerin sind somit in einem Umfang begründet, der geeignet ist, ein genügendes Prüfungsergebnis zu rechtfertigen. Der Prüfungsentscheid erweist sich somit als rechtsfehlerhaft und ist aufzuheben. Es stellt sich jedoch die Frage, welche weiteren rechtlichen Konsequenzen an die Begründetheit der genannten Rügen zu knüpfen sind.

**8.1** Sofern der angefochtene Prüfungsentscheid lediglich Verfahrensfehlern unterliegt, kann dies, selbst wenn sie unzweifelhaft nachgewiesen sind, grundsätzlich nur dazu führen, dass ein Beschwerdeführer den betroffenen Prüfungsteil gebührenfrei wiederholen darf, nicht aber zur Erteilung des Prüfungsausweises. Der Grund hierfür besteht darin, dass für die Erteilung eines Diploms in jedem Fall ein gültiges und genügendes Prüfungsergebnis Voraussetzung ist. Es besteht ein gewichtiges öffentliches Interesse daran, dass nur Kandidaten den entsprechenden Ausweis erhalten, welche den damit verbundenen hohen Erwartungen auch nachgewiesenermaßen entsprechen. Nach ständiger Praxis des BVGer und seiner Vorgängerorganisationen ist deshalb ein gültiges und nachweislich genügendes Prüfungsergebnis grundsätzliche Voraussetzung für die Erteilung eines Prüfungsausweises. Liegt wegen Verfahrensfehlern kein gültiges Prüfungsergebnis vor, so ist diese Voraussetzung nicht erfüllt und es bleibt in der Regel keine andere Lösung, als die betreffende Prüfung durch den Betroffenen wiederholen zu lassen (vgl. das Urteil des BVGer B-7894/2007 vom 19. Juli 2008 E. 4.1 sowie Verwaltungspraxis der Bundesbehörden [VPB] 64.106 E. 6.6.2, VPB 61.31 E. 8.1).

**8.2** Die Rüge, es seien zu Unrecht Prüfungsfragen von der Bewertung ausgeschlossen worden, enthält nicht nur materielle rechtliche, sondern auch verfahrensrechtliche Elemente. Die Beschwerdeführerin beanstandet nicht speziell die Einschätzung ihrer eigenen Prüfungsleistung, sondern ein Verhalten der Prüfungskommission, das sich auf die Rahmenbedingungen der Prüfungsbewertung und somit auf die Bewertung der Arbeiten aller Kandidaten bezog, die an der betreffenden Prüfungssession teilnahmen. Die Rüge betrifft jedoch, anders als die typischen Verfahrensrügen, keine Unregelmässigkeiten im Prüfungsablauf, sondern Fehler bei der Auswertung der Prüfungen. Vorliegend steht zudem fest, in welchem Umfang sich das Prüfungsergebnis der Kandidaten verbessern würde, wenn die Prüfungsfragen nicht ausgeschlossen worden wären. Dies ist, wie bereits erwähnt, bei Verfahrensfehlern normalerweise nicht der Fall (E. 8.1). Die im Hinblick auf die rechtlichen Folgen des Vorliegens von Verfahrensfehlern angeführten Kriterien (E. 8.1) können

daher insofern nicht zur Anwendung kommen. Es ist vielmehr im vorliegenden Zusammenhang sachgerecht, an einen zu Unrecht erfolgten Ausschluss von Prüfungsfragen die gleichen Rechtsfolgen zu knüpfen, wie an eine offensichtliche Fehlbewertung der Prüfung.

**8.3** Die Beschwerdeführerin ist somit nicht zu einer kostenfreien Wiederholung der Prüfung zuzulassen. Die Prüfungsbewertung muss vielmehr unter Berücksichtigung der vorangegangenen Erwägungen korrigiert werden. Fraglich ist aber, ob insofern kassatorisch oder reformatorisch zu entscheiden ist.

**8.4** Hebt das BVGer eine rechtsfehlerhafte Verfügung auf, entscheidet es in der Sache selbst oder weist diese ausnahmsweise mit verbindlichen Weisungen an die Vorinstanz zurück (Art. 61 Abs. 1 des Bundesgesetzes vom 20. Dezember 1968 über das Verwaltungsverfahren [VwVG, SR 172.021]). Als reformatorisches Rechtsmittel gestattet die Beschwerde der Rechtsmittelinstanz, über die Kassation hinaus in der Sache selbst abschliessend zu entscheiden, also das streitige Rechtsverhältnis zu regeln. Damit wird prozessökonomisch das Verfahren abgekürzt, indem sich nicht nochmals die Vorinstanz und allenfalls erneut die Rechtsmittelinstanz mit der Sache befassen muss. Wenn es um Fragen geht, die besondere Sachkenntnis bedingen oder wenn weitere Tatsachen festzustellen sind, ist es jedoch in der Regel nicht Sache des BVGer, als erste Instanz in einem Fachbereich zu entscheiden, in dem ein erheblicher Beurteilungsspielraum der fachkundigeren Vorinstanz zu respektieren ist (vgl. zu allem PHILIPPE WEISSENBARGER in: Bernhard Waldmann/Philippe Weissenberger [Hrsg.] Praxiskommentar VwVG, Zürich/Basel/Genf 2009, Art. 61 N 15 ff.; FRITZ GYGI, Bundesverwaltungsrechtspflege, 2. Aufl., Bern 1983, S. 232 f. mit weiteren Hinweisen und statt vieler BGE 131 V 407 E. 2.1.1). Namentlich in Prüfungsfällen ist beim Entscheid über die Frage, ob kassatorisch oder reformatorisch zu entscheiden ist, auch die Dauer des Verfahrens mitzuberücksichtigen.

**8.5** Im vorliegenden Fall ist zu beachten, dass das BVGer im Rahmen der Instruktion sowohl der Vorinstanz als auch der Erstinstanz die Möglichkeit gegeben hat, sich mit den streitentscheidenden Fragen noch einmal gesondert und umfassend zu befassen. Es liegen insofern detaillierte und ausführliche Stellungnahmen der Prüfungskommissionen vor. Zudem hat die Vorinstanz im Hinblick auf die Frage, ob eine Aufgabe ausgeschlossen werden kann, keinen grossen Ermessens- oder Beurteilungsspielraum. Art. 10 Abs. 3 Prüfungsverfahrensverordnung erlaubt den Ausschluss einer Prüfungsfrage nur in eng begrenzten Ausnahmefällen und ist zudem nicht als Ermessensvorschrift formuliert. Was die Beur-

teilung der Frage E 10 betrifft, wurde eine offensichtliche Fehlbewertung festgestellt, was die Vorinstanz zu respektieren hat. Was die Gewichtung der für die in Aufgabe E 10 als richtig zu wertenden Antwort zu vergebenden Punkte betrifft, ist das Ermessen der Vorinstanz ebenfalls stark eingeschränkt, da die genaue Punkteverteilung für die jeweiligen Antworten in einem verbindlichen Bewertungsraster exakt festgesetzt ist (vgl. hierzu das Urteil des BVGer B-3450/2007 vom 20. November 2008 E. 7.3). Wie eine Neubewertung der Prüfung im Fall der Rückweisung zu erfolgen hätte, ist daher vorliegend durch die vorangehenden Erwägungen weitestgehend vorgezeichnet. Es ist daher aus prozessökonomischen Gründen gerechtfertigt, reformatorisch zu entscheiden.