

Gerfried Gebert*)

Wieviele wissen die Medizinstudenten in der Ärztlichen Vorprüfung?

Eine Ergebnisanalyse der schriftlichen Examina Herbst 1989 bis Herbst 1990

Eine Untersuchung der Prüfungsergebnisse im Fach Physiologie der Ärztlichen Vorprüfung vom Herbst 1989 (3) hatte ergeben, daß mehr als die Hälfte der Kandidaten damals nicht einmal ein Drittel der Physiologiefragen auf der Grundlage erlernten Prüfungswissens richtig beantworten konnte. Grundlage dieser Feststellung war eine Aufgliederung des Rohergebnisses (etwa 47 Prozent richtige Antworten) in den auf Wissen und den auf Raten entfallenden Anteil, wobei der letztere nach der Leistung der

insgesamt schwächsten Kandidaten abgeschätzt wurde. Die in dieser Vorprüfung deutlich besseren Ergebnisse der Fächer Anatomie und Biochemie sowie dann der sprunghafte Anstieg der Physiologie-Antwortquote auf etwa 54 Prozent richtige Antworten im Physikum danach legten die Vermutung nahe, daß das schlechtere Physiologie-Ergebnis vom Herbst 1989 durch eine ungewollt überhöhte Prüfungsanforderung, also durch einen „Ausrutscher“, zustande gekommen sein könnte.

Zur Klärung dieses Problems sind inzwischen alle Ergebnisse der drei klassischen medizinischen Vorprüfungsfächer in den drei Prüfungen vom Herbst 1989 bis Herbst 1990 analysiert worden.

Zur Ermittlung der Ratewahrscheinlichkeit wurde dieselbe Methodik wie in der ersten Studie (3) eingesetzt. Außerdem wurde die Analyse erweitert, da eine unmittelbare Beurteilung der im Mittel erworbenen Kenntnisse und Fähigkeiten selbst anhand der ratebereinigten Lösungshäufigkeit nur dann möglich ist, wenn

▷ die Multiple-choice (M. c.)-Fragen in der fachlichen Anforderung repräsentativ für den in der Ausbildung vermittelten Stoff sind,

▷ die M. c.-Fragen den Kandidaten nicht vor der Prüfung bekannt sind,

▷ die M. c.-Fragen gleichartige quantitative Anforderungen an die Prüfungsstoff-Kenntnis stellen.

Eliminierung der Ratewahrscheinlichkeit

Das Verfahren zur Abschätzung der Wahrscheinlichkeit, mit der die Lösung einer Frage nicht *geraten*, sondern aufgrund von in der Ausbildung erworbenen Kenntnissen und Fähigkeiten *erkannt* wird, soll hier

anhand von zwei im Herbst 1990 gestellten Physiologiefragen dargelegt werden.

(I) Die Belegzellen des Magens erzeugen über ihre luminale Membran einen H^+ -Ionen-Gradienten; dabei beträgt das H^+ -Konzentrationsverhältnis (intrazellulär: luminal) maximal etwa

- (A) 1:10
- (B) 1:10²
- (C) 1:10³
- (D) 1:10⁴
- (E) 1:>10⁵

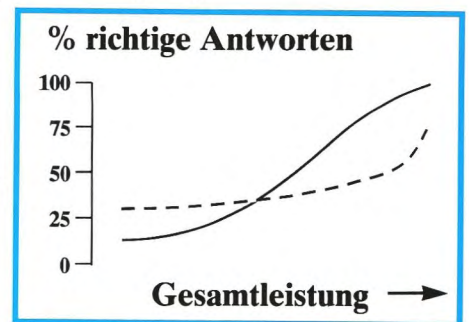
Lösung E; Anteil richtiger Antworten 40 Prozent

(II) Die efferente Projektion aus dem Globus pallidus

- (1) erreicht den Thalamus
- (2) ist inhibitorisch
- (3) benutzt Gammaaminobuttersäure (GABA) als Transmitter
- (A) nur 1 ist richtig
- (B) nur 1 und 2 sind richtig
- (C) nur 1 und 3 sind richtig
- (D) nur 2 und 3 sind richtig
- (E) 1-3 = alle sind richtig

Lösung E; Anteil richtiger Antworten 35 Prozent

In *Darstellung 1* ist die Lösungshäufigkeit der beiden Fragen in Relation zur Gesamtleistung der Kandidaten aufgetragen.

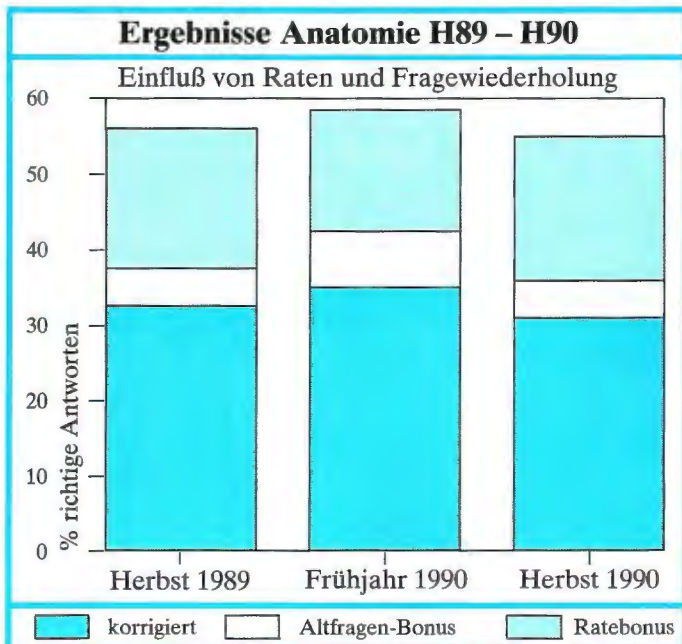


Darstellung 1: Anteil richtiger Antworten auf die Frage I (durchgezogene Linie) und auf die Frage II (unterbrochene Linie) in Relation zur Gesamtleistung der Kandidaten

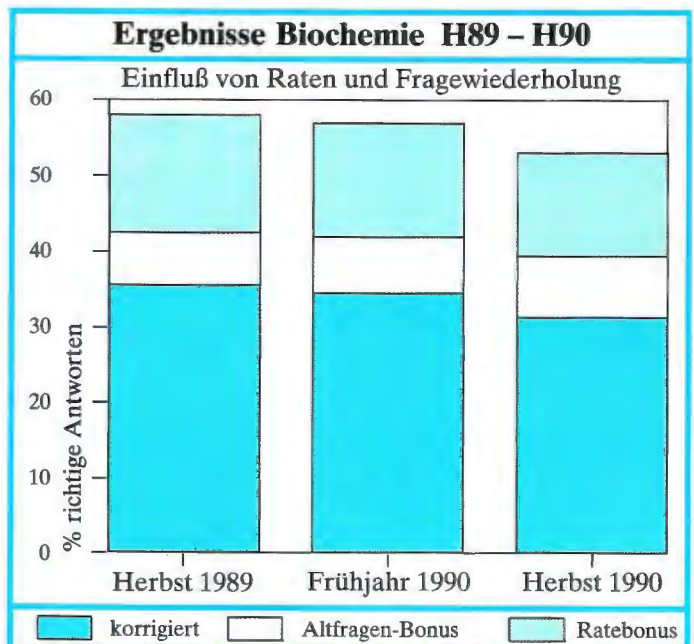
Von den leistungsschwächsten Prüflingen beantworteten die Frage (I) etwa 15 Prozent und die Frage (II) etwa 30 Prozent richtig. Am oberen Ende der Leistungsskala liegt der Anteil richtiger Antworten bei 95 Prozent beziehungsweise bei 70 Prozent.

Verwendet man die Leistung der insgesamt schlechtesten Kandidaten als Ratewahrscheinlichkeitsindikator (3), so ergeben sich als Prozentsätze ausbildungsbedingt gewußter Lösungen für die Frage (I) 29

*) Institut für Medizinische und Pharmazeutische Prüfungsfragen (IMPP, Mainz). Die vorliegende Arbeit gibt die persönliche Auffassung des Autors wieder.



Darstellung 2: Einfluß der empirisch ermittelten Ratewahrscheinlichkeit und wiederholter Fragestellungen auf das Ergebnis der Anatomie-Prüfungen vom Herbst 1989 bis Herbst 1990



Darstellung 3: Einfluß der empirisch ermittelten Ratewahrscheinlichkeit und wiederholter Fragestellungen auf das Ergebnis der Biochemie-Prüfungen vom Herbst 1989 bis Herbst 1990

Prozent und für die Frage (II) nur sieben Prozent, obwohl sich die Lösungshäufigkeiten für die beiden Aufgaben (40 Prozent und 35 Prozent) nicht so wesentlich voneinander unterscheiden.

Eliminierung des Einflusses bereits bekannter Fragen

Die Auswirkung wiederholt gestellter Fragen, die bekanntlich beim Zweiteinsatz in der Regel erheblich besser beantwortet werden als beim ersten Mal, wurde durch eine zweite Auswertung unter Ausschluß solcher Fragen erfaßt.

Als „Altfragen“ eingestuft wurden identisch wiederholte Aufgaben und weitere solche Fragen, deren Beantwortung sich unmittelbar aus der Kenntnis der Lösung von bereits in der Vorprüfung gestellten Fragen ergab.

Die durchschnittliche Lösungshäufigkeit wortidentisch und lösungsidentisch wiederholter Aufgaben lag in allen untersuchten Fächern ähnlich hoch (> zehn Prozentpunkte) über diejenigen noch nicht bekannter Aufgaben.

Umfang des pro Aufgabe geprüften Wissens

Um den Einfluß des Anforderungsumfanges (Zahl der mit einer einzelnen Aufgabe geprüften Sachverhalte) auf die Lösungsquote zu untersuchen, wurden bei der Analyse die Aufgaben in solche eingeteilt, zu deren Lösung die Kenntnis eines Gegenstands ausreichte, und solche, die nur über die Kenntnis mehrerer Gegenstände richtig beantwortet werden konnten.

Aufgaben, bei denen die richtige Kombination mehrerer Aussagen gefunden werden mußte, konnten je nach Lösungsschlüssel in eine der beiden Kategorien fallen. Aufgaben mit einer richtigen und vier falschen Antworten wurden in die höhere Kategorie eingeordnet, wenn die Lösung mehrere zu prüfende Aussagen enthielt.

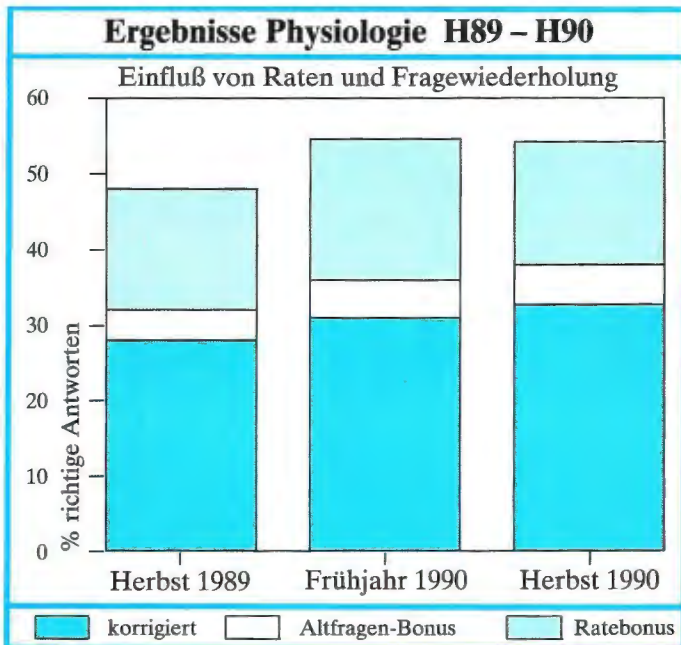
Das entsprechende Vorgehen kann anhand der beiden zitierten Prüfungsaufgaben erläutert werden. Zur Beantwortung der Aufgabe (I) mußte der Kandidat wissen, daß der pH-Wert des Magensafts unter zwei fallen kann und daß der intrazelluläre pH-Wert bei sieben liegt. Außer-

dem mußte er fähig sein, den resultierenden pH-Unterschied von über fünf in den angegebenen H^+ -Gradienten umzurechnen.

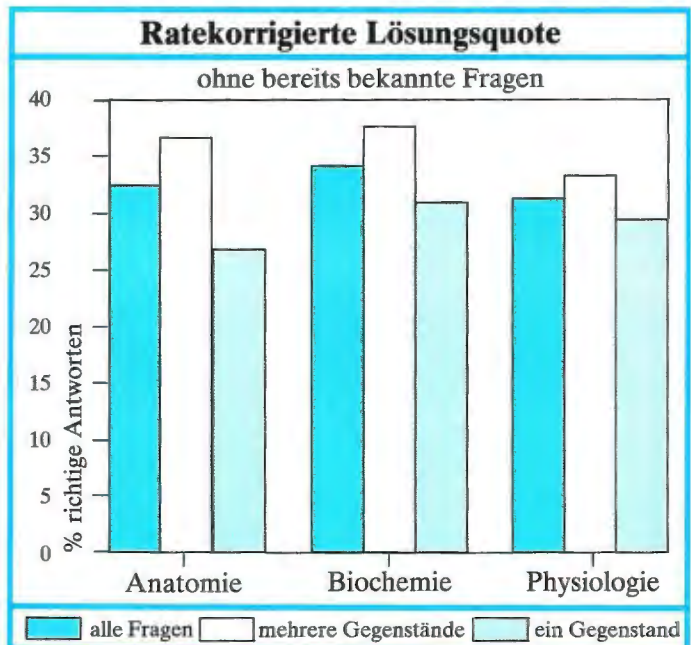
Die Aufgabe prüft daher neben der Kenntnis zweier biologischer Fakten die der Definition des pH-Werts und die Fähigkeit zur Umwandlung von logarithmischen Angaben in Zehnerpotenzen. Ihre Beantwortungsquote darf auf keinen Fall gleichgesetzt werden mit der einer Frage, die lediglich die Kenntnis des minimalen pH-Werts im Magensaft verlangt. Bei der Beispielfrage (II) ist offensichtlich die Kenntnis von drei verschiedenen Eigenschaften des zu den Basalganglien gehörenden Globus pallidus zum Erkennen der richtigen Antwortkombination notwendig. Damit gehören beide Fragen zur Mehrgegenstände-Kategorie.

Längsschnitt der Prüfungsergebnisse

Der Einfluß der wie beschrieben abgeschätzten Ratewahrscheinlichkeit und des Einsatzes wiederholter Fragen auf die Fachergebnisse der schriftlichen Vorprüfungen vom



Darstellung 4: Einfluß der empirisch ermittelten Ratewahrscheinlichkeit und wiederholter Fragestellungen auf das Ergebnis der Physiologie-Prüfungen vom Herbst 1989, vom Frühjahr 1990 und vom Herbst vergangenen Jahres



Darstellung 5: Ratekorrigierte Beantwortung der noch nicht bekannten Prüfungsfragen der Fächer Anatomie, Biochemie, Physiologie der Vorprüfungen vom Herbst 1989 bis Herbst 1990, aufgegliedert nach Fragen zu einem oder zu mehreren Prüfungsgegenständen

Herbst 1989 bis Herbst 1990 ist aus den *Darstellungen 2 bis 4* (auf der Vorderseite und oben) ersichtlich.

Die Gesamthöhe der Säulen entspricht dem globalen Anteil richtiger Lösungen. Durch rechnerische Ratekorrektur erhält man das um den „Ratebonus“ bereinigte Ergebnis, und nach zusätzlichem Abzug des „Altfragenbonus“ durch Beschränkung der ratekorrigierten Auswertung auf erstmals gestellte Aufgaben ergibt sich der korrigierte „echte“ Anteil richtiger Antworten. Dieser liegt in allen drei Vorprüfungen und in allen analysierten Fächern bei etwa einem Drittel.

Daß die offiziellen Prüfungsleistungen meist deutlich über 50 Prozent richtiger Antworten liegen, ist auf die Ergebnisaufbesserung durch Raten und Vorkenntnis von Prüfungsfragen zurückzuführen. Besonders deutlich wird dies beim Sprung der Physiologieergebnisse von etwa 47 Prozent (Herbst 1989) auf etwa 55 Prozent (Frühjahr 1990). Das korrigierte Mittel richtiger Lösungen, das von der Relation Prüfungsstoffkenntnis zu inhaltlicher Prüfungsanforderung bestimmt wird, nahm nur

von 29,1 Prozent auf 30,9 Prozent zu. Gleichzeitig stiegen aber die mittlere empirisch ermittelte Ratewahrscheinlichkeit (von zirka 22 Prozent auf zirka 30 Prozent) und der Altfragenbonus (ratekorrigiert von 2,6 auf 4,8 Prozentpunkte) stark an, so daß die Verbesserung des Ergebnisses um acht Prozentpunkte in erster Linie durch *Unterschiede in der Prüfungszusammenstellung* bedingt ist.

Der Einfluß der Zahl der pro Aufgabe geprüften Gegenstände auf die ratekorrigierte Lösungshäufigkeit ist für die erstmals gestellten Fragen in der *Darstellung 5* wiedergegeben, wobei die Fragen der drei Prüfungstermine zusammengefaßt wurden.

Vor allem im Fach Anatomie hängt die Quote richtiger Antworten erheblich von der Zahl der Prüfungsgegenstände ab. Aufgaben, deren Beantwortung nur die Kenntnis *eines* Sachverhalts erforderte, liegen in der ratekorrigierten Lösungshäufigkeit um 10,2 Prozentpunkte über solchen, die nur aus der Kenntnis *mehrerer* Prüfungsgegenstände her gelöst werden konnten.

Hierzu sei angemerkt, daß das Bundesverfassungsgericht den (ohne zwingende Notwendigkeit) erfolgen-

den Einsatz von Fragen verschiedener Typen und damit von qualitativ verschiedenem Anforderungsumfang (2) zur Begründung seiner Entscheidung, eine absolute Bewertung der Leistung in den schriftlichen Prüfungen nach der Ärztlichen Approbationsordnung als Verstoß gegen das Übermaßverbot der Verfassung zu untersagen (1), herangezogen hat.

Vorprüfungsergebnisse und die Prüfungsstoffkenntnis

In der vorangehend erläuterten Untersuchung des Einflusses der Ratewahrscheinlichkeit auf die Prüfungsergebnisse des Faches Physiologie im Herbst 1989 (3) wurde nachgewiesen, daß die Physikumskandidaten im Mittel nicht einmal ein Drittel der Aufgaben von dem in der Ausbildung erworbenen *Wissen* her lösen konnten. Diese Feststellung trifft nach der hier vorgelegten Analyse auch zu auf die Vorprüfungen des Jahres 1990. Dies gilt auch für die anderen beiden „klassischen“ Physikumsfächer Anatomie und Bio-

chemie, was erkennbar wird, wenn man nur die Aufgaben berücksichtigt, die noch nicht durch Einsatz in vorhergehenden Examina bekannt sein konnten. Auffällige Schwankungen im Globalergebnis zwischen einzelnen Prüfungsterminen erwiesen sich als großteils durch unterschiedliche Ratewahrscheinlichkeit und differierende Altfragenanteile bedingt.

Die eigentliche Prüfungsstoffkenntnis kann auf der Basis richtig gewußter Lösungen der in der Vorprüfung gestellten M. c.-Fragen nicht exakt beurteilt werden. Dies liegt nicht daran, daß M. c.-Fragen ungeeignet wären, Wissen und Fähigkeiten zuverlässig und valide zu messen, sondern an der konstruktiven Heterogenität der eingesetzten Aufgaben.

Der von Fach zu Fach und Prüfung zu Prüfung unterschiedliche Anteil von Aufgaben, zu deren Lösung die Kenntnis von mehr als einem Sachverhalt erforderlich ist, senkt die Quote richtiger Antworten in nicht ausreichend erfaßbarem Ausmaß unter den Anteil gewußten Prüfungsstoffs am gesamten Prüfungsstoff.

Allerdings hält sich dieser Fehler in Grenzen, wie sich aus dem Vergleich der Prüfungsleistungen bei Aufgaben mit einem beziehungsweise mit mehreren geprüften Gegenständen ergibt. Die Aussage, daß die durchschnittliche Prüfungsstoffkenntnis der Kandidaten in den drei Fächern Anatomie, Biochemie und Physiologie stabil unter 40 Prozent liegt, wird dadurch nicht in Frage gestellt.

Vorprüfung und Sicherung des Ausbildungserfolgs

„Ist der Studiosus noch so dumm, so schafft er doch sein Physikum“, lautet ein aus lange vergangenen (vor der Einführung der zentralen Prüfungen liegenden) Zeiten überlieferter Spruch. Es wäre billig, die Ergebnisse der hier vorgelegten Untersuchungen in diesem Sinne zu genereller Herabwürdigung der von den Kandidaten der Ärztlichen Vor-

prüfung erbrachten Leistungen zu mißbrauchen.

● Bei dem enormen Umfang des in zwei Jahren Vorklinik zu erwerbenden Wissens (nach entsprechenden Analysen [2] liegt allein im Fach Anatomie die Zahl der prüfungswürdigen Sachverhalte über 5000) ist eine ratebereinigte Antwortquote von 50 Prozent nicht als Halbwissen, sondern als solider und guter Wissensquerschnitt zu werten. Die geltende Bestehensregel erlaubt jedoch, daß auch Kandidaten mit weitaus geringeren Leistungen in den klassischen Vorprüfungsfächern das Physikum bestehen.

● Besonders problematisch ist die Möglichkeit, ein oder zwei der „klassischen“ Grundlagenfächer der Vorprüfung „abzuwählen“ und die zum Bestehen notwendigen Punkte in Fächern mit in Relation zur Prüfungsfragenzahl weitaus geringerem Stoffumfang (zum Beispiel Psychologie/Soziologie oder Biologie) zu holen.

Daß es sich hier nicht um hypothetische Spekulationen handelt, kann aus den detaillierten Ergebnisberichten des Instituts für Medizinische und Pharmazeutische Prüfungsfragen (Mainz) (4) entnommen werden. Einer solchen Taktik steht (etwas Losglück vorausgesetzt) auch der mündliche Teil der Ärztlichen Vorprüfung nicht im Wege, denn diese Prüfung muß nur in einem und höchstens in zwei der „klassischen“ Grundlagenfächer abgelegt werden.

Literatur

- (1) Bundesverfassungsgericht: Begründung des Urteils (Az. 1 BvR 1033/82) vom 14. 03. 1989, S. 36
- (2) Gebert, G.: Die schriftliche Prüfung nach der ÄAppO – Probleme der Leistungsmessung und Leistungsbewertung; Medizinische Ausbildung 2: 12–17 (1985)
- (3) Gebert, G., Voigtmann, K.: Entscheidend ist, was die (M. c.-)Prüfungen leisten. DÄBl 87: B 1679–1684 (1990)
- (4) IMPP: Ergebnisse der Prüfungen nach der Approbationsordnung für Ärzte, Herbst 1989 und Frühjahr 1990; IMPP-Berichte 1990

Anschrift des Verfassers:

Prof. Dr. med.
Gerfried Gebert
Rüsselsheimer Allee 22
W-6500 Mainz 43

Gerda Hasselfeldt für Indikationsregelung

Bundesgesundheitsministerin Gerda Hasselfeldt (CSU) lehnt es ab, den Ärztinnen und Ärzten „die alleinige Entscheidung über den Schwangerschaftsabbruch aufzudrängen“. Dies wäre mehr als verhängnisvoll. Sie würden in ihrer Verantwortung und Verpflichtung überfordert und entgegen ihrem christlichen und medizinischen Ethos mißbraucht, betonte die Ministerin auf einer Tagung der Katholischen Ärztarbeit Deutschlands zum Thema „Gerechtigkeit oder das Recht des Stärkeren in der Medizin“ in Mainz. Sie sprach sich damit gegen den von Bundesfamilienministerin Hannelore Rönsch (CDU) auf dem Deutschen Ärztetag vorgestellten Vorschlag aus, wonach der Arzt über die Notlage der Patientin und damit über die Möglichkeit des Abbruchs einer Schwangerschaft mitentscheiden müsse.

Gerda Hasselfeldt befürwortet eine Indikationsregelung, die sich „an den Wertvorstellungen des Grundgesetzes ausrichtet“. Dementsprechend dürfe der Staat auf das Mittel des Strafrechts zur Wahrung des Lebensrechts eines ungeborenen Kindes allenfalls dann verzichten, wenn die Schwangerschaft eine persönliche Belastung darstelle, die unmittelbar in das Recht der Frau auf persönliche Lebensgestaltung eingreife.

Dabei habe die obligatorische medizinische und soziale Beratung vor der Indikationsfeststellung eine zentrale Funktion für den Schutz des ungeborenen Lebens. Eine „wie auch immer geartete Fristenlösung“ lehnt Gerda Hasselfeldt entschieden ab. Das gelte auch für diejenigen Lösungsansätze, die lediglich eine Beratung vorsehen, die allein das Ziel verfolge, die formalen Voraussetzungen für einen Schwangerschaftsabbruch zu erfüllen. Die Bewältigung der Konfliktsituation gehe selbstverständlich auch den Mann an. Er habe nicht weniger Verantwortung als die Frau. „Im Gegenteil, auch Männer, die ihre Frauen zur Abtreibung drängen, sollten bestraft werden.“ ▷