

EMS • Eignungstest für das Medizinstudium in der Schweiz • 2004

*Bericht 10 über die Durchführung und Ergebnisse 2004
Trends und Schlussfolgerungen aus sieben Jahren Testanwendung*

avec un résumé en français

K.-D. Hänsgen und B. Spicher

Bericht 10 (2004)



10Jahre ^{1994 - 2004}

Zentrum für Testentwicklung und Diagnostik
am Departement für Psychologie der Universität Freiburg - Schweiz



Hänsgen, Klaus-Dieter; Spicher, Benjamin (2004).

EMS Eignungstest für das Medizinstudium in der Schweiz 2004

Berichte des Zentrums für Testentwicklung, Nummer 10, im Auftrag der Schweizerischen Universitätskonferenz (SUK)

Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik, 2004

Unter Mitarbeit von Katharina Stress (Lektorat),

Roisin Acton, Mauro Bernasconi, Sébastien Simonet, Maurizio Strazzeri

Redaktion: Tanja Läng

Die Mitglieder des Beirates „Eignungstest“ der Schweizerischen Universitätskonferenz zum Zeitpunkt der Testabnahme 2004:

Dr. N. Ischi (Vorsitzender, Bern); Prof. Dr. P. Groscurth (Zürich); Dr. R. Hofer (Bern); Prof. Dr. M. Kleinmann (Zürich); Prof. Dr. J.-P. Montani (Freiburg); Prof. Dr. K. Opwis (Basel); Prof. Dr. M. Perrez (Freiburg); Prof. Dr. J.-L. Reymond (Bern); Prof. Dr. U. M. Spornitz (Basel); Prof. Dr. G. Trost (Bonn/Deutschland); K. Wechsler (Bern).

© ZTD Zentrum für Testentwicklung und Diagnostik 2004

am Departement für Psychologie der Universität Freiburg

Direktor: Prof. Dr. K.-D. Hänsgen

Route d'Englisberg 9, CH-1763 Granges-Paccot

E-Mail: ztd@unifr.ch

Internet: <http://www.unifr.ch/ztd>

Druck: Cric Print, Freiburg/Schweiz

Inhaltsverzeichnis

1	ZUSAMMENFASSUNG	5
2	RESUME	9
3	FAQ ZU EIGNUNGSTEST UND NUMERUS CLAUSUS	13
4	NUMERUS CLAUSUS (NC) UND MEDIZINSTUDIUM	28
5	ANMELDUNG ZU MEDIZINSTUDIUM UND TEST 2004.....	31
5.1	Anmeldungen, Testantritte und Zulassungen.....	31
5.2	Grösse der Testlokale	36
5.3	Testorte und Wunschuniversitäten.....	36
5.4	Wunschuniversitäten und Testorte nach Wohnkantonen	38
5.5	Testabsolvierung nach Alter und Geschlecht	42
5.6	Übernahme des Testergebnisses aus Vorjahren	44
5.7	Erneute Testteilnahme 2004	45
5.8	Sprachgruppen.....	46
5.9	Alter und Maturitätsjahr	49
6	BESCHREIBUNG DES VERWENDETEN EIGNUNGSTESTS.....	50
6.1	Aufbau des Tests	50
6.2	Berechnung der Werte	52
6.3	Mittlerer Rangplatz der Untertests.....	53
7	TESTANWENDUNG IN DER SCHWEIZ 2004.....	54
7.1	Organisation der Testabnahme.....	54
7.2	Verteilungsprüfung	54
7.3	Vergleich der Testfassungen 1998 bis 2004.....	57
7.4	Äquivalenz der Sprachversionen	59
7.4.1	Sprachvergleich für die Untertests	61
7.4.2	Darstellung des Korrekturverfahrens	64
7.4.3	Leistungen in identischen Testaufgaben 1998 – 2004	75
7.5	Vergleichbarkeit der Testlokale	77
7.6	Vergleich der Geschlechter	78
7.7	Vergleich der Kantone	84
7.7.1	Vergleich 2004 innerhalb der Deutschschweiz	84
7.8	Vergleiche für Altersgruppen.....	89
7.9	Vergleiche nach Wunschuniversitäten	91
8	ERGEBNISSE ZUR TESTGÜTE	94
8.1	Zuverlässigkeit	94
8.2	Binnenstruktur.....	96
8.3	Item-Trennschärfen.....	103

8.4	Konzentriertes und Sorgfältiges Arbeiten	107
9	ZUSAMMENFASSUNG KANTONE 1998 BIS 2004	110
10	LITERATUR	115
10.1	Originaltest zur Information und Vorbereitung.....	116
10.2	Frühere Berichte des ZTD.....	117
11	ANHANG ZUM BERICHT 2004	119
11.1	Beispielaufgaben für die Untertests.....	119
11.1.1	Untertest: Muster zuordnen.....	119
11.1.2	Untertest: Medizinisch-naturwissenschaftliches Grundverständnis	120
11.1.3	Untertest: Schlauchfiguren.....	121
11.1.4	Untertest: Quantitative und formale Probleme	121
11.1.5	Untertest: Konzentriertes und Sorgfältiges Arbeiten	122
11.1.6	Untertest: Figuren lernen	122
11.1.7	Untertest: Fakten lernen	123
11.1.8	Untertest: Textverständnis	123
11.1.9	Untertest: Diagramme und Tabellen	125
11.2	Weitere Detailergebnisse	126

1 Zusammenfassung

*Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2004 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Human-, Veterinär- und Zahnmedizin ab Wintersemester 2004/2005. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2003 herangezogen. **Die Ergebnisse sind nicht repräsentativ für Schlussfolgerungen hinsichtlich einer gesamthaften oder regionalen Evaluation von Bildungseinrichtungen und Bildungswegen.***

Im 10. Jahr des Bestehens des ZTD wurde der EMS zum siebten Mal durchgeführt. Die Anpassung des EMS an neue Studienbedingungen bei gleichzeitiger Erhaltung der hohen Prognosekraft für Studienerfolg hat begonnen und wird in den Folgejahren fortgesetzt.

Statistik

Im Jahre 2004 war für **Human-, Veterinär- und Zahnmedizin** ein Numerus Clausus für Personen notwendig, die sich an den Universitäten Basel, Bern, Freiburg oder Zürich angemeldet haben.

Bereits im **Februar 2004** waren mit den vorliegenden Anmeldungen die Kapazitäten dieser Universitäten in **Humanmedizin zu 249%**, in **Veterinärmedizin zu 182%**, in **Zahnmedizin zu 146%** ausgelastet. Dies ist ein Anstieg gegenüber dem Vorjahr um 50% (Humanmedizin) bzw. 20% (Veterinärmedizin). Zahnmedizin überschritt erstmals das Kriterium, welches einen NC auslöst.

Insgesamt 1453 Personen haben sich bis zur gesetzten Frist im **Mai 2004** bei der CRUS zum EMS angemeldet, 52 Personen übernahmen ihr Testergebnis aus dem Vorjahr. Die Zahl der Personen, die wiederholt zum EMS antraten, hat sich mit 96 ebenfalls deutlich erhöht. Diese Bereitschaft, sich nach einer Ablehnung erneut zu bewerben und noch einmal beim EMS anzutreten, deutet auf eine hohe Motivation hin.

Mit gültigem Ergebnis absolviert haben den EMS 1373 Personen. Vier Personen haben den Test vorzeitig abgebrochen, 97 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen. Diese Zahlen liegen im Bereich der Vorjahre.

Die verstärkte Nachfrage nach Studienplätzen wird nach jüngsten Prognosen des Bundesamtes für Statistik anhalten, dies gilt somit auch für das Medizinstudium und deckt sich mit internationalen Trends. Selbst eine realistische Erhöhung der Studienplätze würde einen Numerus Clausus nicht überflüssig machen. Die Qualität der Medizinausbildung und die mit Studienreformen verbundenen Verbesserungen wären bei einer Abschaffung des NC vermutlich nicht auf heutigem Niveau zu halten.

Organisation und Ablauf

Der Test fand am 2.7.2004 gleichzeitig an sieben Testorten in drei Sprachen (deutsch: Basel, Bern, Chur, St. Gallen, Zürich; französisch: Freiburg; italienisch: Bellinzona) statt. Zu beachten ist, dass die Wahl des Testortes **unabhängig vom zukünftigen Studienort** möglich ist. Es traten keine besonderen Vorkommnisse auf, welche die Auswertbarkeit

beeinträchtigen. In diesem Jahr gab es vier vorzeitige Testabbrüche und zwei Verwarnungen wegen unerlaubtem Verhalten im Test.

Testgüte

Eine mit den Vorjahren vergleichbar hohe Testgüte gewährleistet die Vergleichbarkeit der Ergebnisse, insbesondere zu 2003 (aus welchem Ergebnisse bekanntlich übertragen werden konnten).

Zuverlässigkeitskennwerte und Itemtrennschärfen bewegen sich auch 2004 im Bereich der bisherigen Jahre. Die Zuverlässigkeit des Testwertes liegt mit .91 im Bereich der Vorjahre und ist als ausreichend hoch einzuschätzen.

Der neu gestaltete Untertest „Konzentriertes und Sorgfältiges Arbeiten“ hat die Erwartungen erfüllt: Effekte eines exzessiven Trainings sind nicht mehr nachzuweisen, die Punkteverteilung ist deutlicher normalverteilt. Es wird weniger der Fleiss, sondern wieder mehr das eigentliche Testziel gemessen. Er bestimmt in der Faktorenanalyse deutlicher den Faktor „visuelle Fähigkeiten“ und die im Vergleich zum Vorjahr geringeren Interkorrelationen mit den anderen Untertests weisen auf mehr Eigenständigkeit dieser Leistung hin.

Gleichbehandlung der Sprachgruppen

Wie in den Vorjahren treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Untertests gleichermassen auf. Auch dieses Jahr liegen die Punktzahlen der deutschsprachigen Teilnehmer durchwegs über jenen der französisch- und italienischsprachigen Kandidaten. Die Adaptation des EMS wurde wiederum so gestaltet, dass eine maximale Äquivalenz der Aufgaben erreicht wird. Mögliche noch vorhandene testbedingte Unterschiede werden mit dem DIF-Verfahren identifiziert und ausgeglichen. Dabei war ein Ausgleich nur bei 2 sprachabhängigen Untertests pro Sprachgruppe notwendig. Die Zahl der zu korrigierenden Items ist minimal (2 Items französisch- und 6 Items italienischsprachig), was für die gute Übersetzungsqualität spricht. Die Gewähr, dass testbedingte Unterschiede sicher ausgeschlossen sind, zeigt sich im Vergleich der sprachabhängigen und weniger sprachabhängigen Untertests: In den vier weniger sprachabhängigen Untertests sind die Unterschiede in allen Sprachgruppen deutlich grösser als in den fünf sprachabhängigen Untertests.

Differenzierung nach Alter

Es bestätigt sich auch in diesem Jahr der aus den Vorjahren bekannte Trend beim Testwert, betrachtet nach Alter und Zeitpunkt des Ablegens der Maturität: Jüngere Studienbewerber erreichen höhere Testwerte. Alle älteren erhalten etwas geringere Testwerte, wenn sie die Maturitätsprüfung früh abgelegt haben. Die Gruppe der Älteren mit später Maturitätsprüfung hat verglichen dazu noch etwas geringere Testwerte.

Wir erinnern daran, dass auch der Erfolg in den Vorprüfungen in etwa diesem Trend folgt und der Test dies hier nur korrekt vorhersagt.

Differenzierung nach Geschlecht

Bekanntlich ist die Erfolgswahrscheinlichkeit, die erste Vorprüfung zu bestehen, für Männer um 7% höher als für Frauen. Wenn der Test diese Erfolgswahrscheinlichkeit vorhersagen will, muss er auch diesen Unterschied feststellen. In den letzten Jahren war der Unterschied nur in der Tendenz feststellbar und statistisch nicht signifikant.

Bei einer Betrachtung nach Disziplin und Sprachgruppe ist in diesem Jahr ein signifikanter Unterschied zugunsten der Männer für Humanmedizin und für die deutsche Sprache feststellbar. Zum einen ist dafür die Statistik selbst verantwortlich, indem in grösseren Stichproben bereits geringere Unterschiede als signifikant nachgewiesen werden. Zum anderen ist der Frauenanteil in der Bewerberkohorte Humanmedizin wiederum gestiegen. Es ist eher unwahrscheinlich, dass eine unterschiedliche Repräsentativität der Bewerbergruppen für ein Medizinstudium für die gesamten Geschlechtergruppen gleiche Mittelwerte erzielt. Der festgestellte Unterschied bewegt sich allerdings im Rahmen der Vorjahre und summa summarum werden aufgrund der theoretischen Zulassungsrate 372 Frauen gegenüber 231 Männern zugelassen.

Die Chancengleichheit ist auch 2004 gegeben (die Anforderungen waren zumindest gleich), eine Benachteiligung liegt nicht vor.

Differenzierung nach Maturitätsquote und Kanton

Es muss aufgrund mehrerer Presseberichte über Rangreihenvergleiche der Kantone noch einmal auf die Einschränkungen der Repräsentativität hingewiesen werden: Es handelt sich „nur“ um Medizinbewerbungen, die unterschiedlich repräsentativ für alle Maturanden der Kantone sind. In zweisprachigen Kantonen kommt dazu, dass Französischsprachige und Zweisprachler ohne Numerus Clausus in Genf, Lausanne oder Neuenburg ein Studium aufnehmen können. Auch können „sozialpsychologische“ Phänomene in Kantonen unterschiedlich sein (auch im Vergleich der Jahrgänge innerhalb eines Kantons), indem z.B. einmal nur die besten Maturitätsabsolventen sich für ein Medizinstudium bewerben und zum anderen ein solcher „Gruppendruck“ nicht besteht.

Ein genereller Einfluss der Maturitätsquote auf den Testwert lässt sich wiederum nicht nachweisen (3 Gruppen der Kantone nach Maturitätsquote zusammengefasst unterscheiden sich nicht signifikant). Die Unterschiede 2004 für die Kantone mit den höchsten Bewerbungszahlen bestätigen allerdings die aus den Vorjahren bekannten Trends. Hier scheint es über die sieben Jahre drei Gruppen der Kantone zu geben, die ihre Positionen nur innerhalb der Gruppe tauschen.

Trotz der möglichen Missbrauchsgefahr haben wir uns entschlossen, die zusammenfassenden Statistiken für alle Kantone weiter zu veröffentlichen. Immerhin sind nunmehr 6720 Testabnahmen erfolgt und die vorhandenen Unterschiede werden statistisch immer besser abgesichert. Es liegt seitens der Kantone ein Interesse vor, das Abschneiden ihrer Personen im Zulassungsverfahren zu erfahren – aber nur bezogen auf das Abschneiden bei den Medizinbewerbungen. Wenn die Zahlen Anlass zu konstruktiven Verbesserungen bei der EMS-Vorbereitung oder auch bei der Gestaltung des Bildungssystems geben, so muss dies nicht falsch sein. Der Rückschluss „schlechtes Abschneiden ist gleich schlechtes Bildungssystem“ ist aber auf jeden Fall falsch.

Schlussfolgerungen

Auch die siebte Session des EMS konnte erfolgreich absolviert werden – trotz einer deutlichen Zunahme der Teilnehmerzahlen. Die Testergebnisse sind gleich zuverlässig wie in den Vorjahren.

Mit dem Numerus Clausus (und dem dazu verwendeten EMS) steht ein bewährtes Verfahren zur Verfügung, welches eine kapazitätsgenaue Zulassung aufgrund der Studieneignung ermöglicht und auch die Umleitungen nach einem objektiven Kriterium regelt.

Die Entscheidung der Notwendigkeit des EMS wird weiterhin jährlich getroffen. Aufgrund der statistischen Prognosen muss auch künftig ein qualitativ hoch stehender Test zur Verfügung stehen. Die Konstanz des EMS wird darin bestehen, bewährte Teile mit hoher Vorhersagekraft für den Studienerfolg beizubehalten. Einzelne Veränderungen und Erweiterungen wurden bereits vorgenommen und werden weiter folgen, um die Vorhersagekraft auch auf die Weiterentwicklung der Studienanforderungen auszurichten.

2 Résumé

Le présent rapport expose les résultats obtenus lors du test réglant l'admission en médecine humaine, vétérinaire et dentaire au semestre d'hiver 2004/05. Il fait également intervenir des données comparatives des éditions antérieures (1998-2003). Mais rappelons que ces résultats ne sont pas suffisamment représentatifs pour permettre de tirer des conclusions quant aux établissements de formation ou à l'éducation en général, que ce soit au plan suisse ou régional.

Le CTD, qui fête son 10^e anniversaire cette année, a organisé pour la 7^e fois le test d'aptitudes pour les études en médecine en Suisse (AMS). Il a déjà commencé à l'adapter aux nouvelles exigences tout en veillant à préserver sa haute valeur prédictive quant à la réussite des études. Il poursuivra ce travail au cours des prochaines années.

Statistiques

En 2004, il a été nécessaire d'imposer un **numerus clausus (NC)** aux personnes désirant commencer leurs études en **médecine humaine, vétérinaire et dentaire** à l'Université de Bâle, Berne, Fribourg ou Zurich.

En effet, les pré-inscriptions dépassaient déjà, en **février 2004**, la capacité d'accueil de ces universités de **249% en médecine humaine**, de **182% en médecine vétérinaire** et de **146% en médecine dentaire**, ce qui représente, par rapport à 2003, une augmentation de 50% (en médecine humaine) et de 20% (en médecine vétérinaire). Quant à la médecine dentaire, c'est la première fois que le nombre de pré-inscriptions est supérieur à la capacité d'accueil, raison pour laquelle le NC a été appliqué pour cette discipline également.

En **mai 2004**, date butoir pour l'inscription au test, 1453 personnes en tout s'étaient annoncées auprès de la CRUS, tandis que 52 personnes reprennent leur résultat de l'année précédente et que 96 personnes se portent candidat pour la 2^e fois – un chiffre en nette augmentation. Mais se représenter à l'AMS, après un échec, c'est faire preuve d'une grande motivation, ce qui est en soi tout à fait souhaitable.

Le jour J, 1373 personnes ont passé le test, 4 personnes l'ont interrompu prématurément tandis que 97 personnes s'étaient désistées ou ne sont simplement pas venues: ces valeurs correspondent à celles des années précédentes.

Selon les plus récentes prévisions de l'Office fédéral de la statistique, confirmées par l'évolution internationale, la médecine ne sera pas épargnée par l'engouement pour les études. Le **numerus clausus** reste par conséquent nécessaire, même si on augmente raisonnablement le nombre de places d'études. En effet, il ne serait probablement pas possible, sans **numerus clausus**, de maintenir la qualité de la formation médicale au niveau actuel, ni de lui faire bénéficier des améliorations issues de la réforme des études.

Organisation et déroulement

Le test, trilingue, s'est déroulé le 2 juillet 2004 au même moment dans sept villes différentes (français: Fribourg; allemand: Bâle, Berne, Coire, St-Gall, Zurich; italien: Bellinzone). A

noter qu'il est possible de choisir le lieu du test **indépendamment du futur lieu d'études**. Le test n'a été marqué par aucun incident susceptible d'en restreindre l'utilisation. Cette année, 4 personnes ont interrompu prématurément leur test et 2 ont dû être rappelées à l'ordre.

Qualité du test

Le test étant tout aussi bon que les années précédentes, il est possible d'en comparer les résultats, en particulier avec 2003 (dont les résultats ont pu être repris).

En 2004, la fiabilité et la qualité des items ont des valeurs semblables à celles des années précédentes. Avec 0.91, la fiabilité est estimée nettement suffisante.

Le sous-test "Travail avec soin et concentration" a été réaménagé et répond désormais aux attentes: la répartition des points est nettement plus normale, par conséquent un surentraînement n'a plus d'influence notable sur les résultats. Car, conformément à son but, le sous-test mesure davantage le soin et la concentration que l'assiduité. Il détermine plus clairement le facteur "capacités visuelles" dans l'analyse factorielle et sa corrélation avec les autres sous-tests, plus faible qu'en 2003, s'explique par la spécificité de ce travail.

Egalité de traitement entre groupes linguistiques

Comme par le passé, il existe des différences entre les groupes linguistiques dans les sous-tests à forte composante linguistique; différences qui se retrouvent dans les sous-tests à faible composante linguistique. Cette année également, les germanophones ont obtenu bien davantage de points que les candidats francophones ou italophones. L'AMS a été de nouveau adapté de manière à obtenir une équivalence maximale des exercices. Les différences qui seraient dues au test sont identifiées grâce à la procédure DIF et le cas échéant compensées: cela n'a été nécessaire dans chaque groupe linguistique que pour deux sous-tests à forte composante linguistique. Le nombre d'items à corriger est minime (2 en français et 6 en italien), ce qui témoigne de la qualité de la traduction. La comparaison entre sous-tests permet d'exclure avec certitude que les différences sont dues au test lui-même: les sous-tests à faible composante linguistique présentent des différences entre groupes linguistiques nettement plus grandes que les sous-tests à forte composante linguistique.

Différences entre classes d'âge

Cette année confirme les observations faites les années précédentes quant à l'âge et au moment de passer la maturité: les candidats les plus jeunes obtiennent de meilleurs résultats. Les candidats plus âgés arrivent à des résultats légèrement moins bons s'ils ont passé rapidement leur maturité, et encore un peu moins bons s'ils l'ont passé tardivement.

Pour mémoire, la réussite aux examens propédeutiques suit à peu près la même courbe, c'est donc bien ce que prédit le test.

Différences entre femmes et hommes

Il est notoire que la probabilité de réussite au premier examen est 7% plus élevée chez les hommes que chez les femmes. Il faut que l'on retrouve cette différence au niveau du test, s'il

est censé la prédire. Il n'avait été possible jusqu'ici de déceler qu'une tendance statistiquement non significative.

Cette année, en considérant chaque discipline et groupe linguistique, on constate une différence significative en faveur des hommes, inscrits en médecine humaine et germanophones. Cela s'explique d'une part, statistiquement, car une différence minimale est déjà significative pour de grands échantillons et, d'autre part, car la proportion des femmes a encore augmenté cette année en médecine humaine. On peut partir du principe que la différence proportionnelle entre femmes et hommes parmi les candidats aux études de médecine n'influe en rien leurs valeurs moyennes. La différence constatée se situe de toute manière dans les limites des années précédentes et, en raison du taux d'admission théorique, ce sont, dans l'ensemble, 372 femmes contre 231 hommes qui sont admis.

En 2004, le test respecte également l'égalité des chances (les exigences étant les mêmes), aucun genre n'ayant été discriminé.

Différences entre cantons et taux de maturité

Il est nécessaire de revenir sur les limites à apporter à la représentativité des résultats AMS, en raison d'une série d'articles de presse comportant un classement des cantons: le test ne concerne "que" des candidatures aux études de médecine, dont la représentativité par rapport aux autres bacheliers du canton est très variable. Dans les cantons bilingues, les francophones et les bilingues peuvent commencer leurs études de médecine à Genève, Lausanne ou Neuchâtel sans être soumis à un *numerus clausus*. De même, certains phénomènes "socio-psychologiques" peuvent varier d'un canton à l'autre (voire d'une année à l'autre dans un même canton), au sens où une année, par exemple, seuls les meilleurs bacheliers décident d'étudier la médecine, alors qu'une autre fois, il n'y a pas un tel "effet de groupe".

Il n'est pas possible d'établir un lien entre taux de maturité et résultats au test (les 3 groupes de cantons réunis en fonction du taux de maturité ne se différencient pas de manière significative). Il est vrai que 2004 confirme les tendances mises en évidence les années précédentes pour les cantons présentant le plus grand nombre de candidats: en sept ans, les cantons semblent former trois groupes, où ils changent simplement de position d'une année à l'autre.

Nous avons décidé de continuer à publier les statistiques globales pour tous les cantons en dépit du risque d'abus. 6720 résultats sont tout de même en jeu et les différences constatées seront statistiquement toujours mieux établies. Les cantons sont intéressés à savoir comment leurs candidats se distinguent dans la procédure d'admission aux études de médecine mais uniquement dans ce contexte-là. Si les chiffres donnent l'occasion d'améliorer la préparation à l'AMS ou l'aménagement du système de formation, il n'y a rien à redire. Mais il serait tout à fait erroné de conclure que les "derniers de classe" ont un mauvais système de formation.

Conclusions

La septième session de l'AMS s'est déroulée avec succès, malgré un accroissement net de la participation. Les résultats obtenus sont tout aussi fiables que ceux des années précédentes.

Le *numerus clausus* (et l'AMS appliqué à cet effet) fournit une procédure éprouvée afin de n'admettre aux études de médecine que les personnes les plus aptes en fonction des places disponibles et de procéder à des transferts au moyen d'un critère objectif.

La décision d'organiser un AMS continuera à être prise chaque année. Les prévisions statistiques nous contraignent à préparer chaque année un test de qualité. L'AMS restera constant si on maintient les sous-tests ayant un haut pouvoir prédictif quant au succès académique. Nous avons déjà procédé à quelques changements ou développements que nous allons poursuivre afin de garantir au test un haut pouvoir prédictif pour les nouvelles exigences en matière d'études.

3 FAQ zu Eignungstest und Numerus Clausus

Seit der Einführung des EMS 1998 ist einige Zeit vergangen. Nach der anfänglich intensiven Diskussion über Für und Wider des NC wird heute nur noch selten die Problematik in ihrer ganzen Komplexität diskutiert. Ein Risiko ist, dass die Intentionen der Einführung von NC und EMS dabei in Vergessenheit geraten. Einige der folgenden Fragen wurden von der Erziehungsdirektion des Kantons Bern, Amt für Hochschulen, gestellt.

1. Welche Länder kennen Zulassungsbeschränkungen zum Medizinstudium?

Die bisher einzige europäische Studie, Ebach & Trost (1997)¹, kommt zum Ergebnis: „Praktisch überall in Europa findet eine Auswahl unter den Bewerbern für medizinische Studiengänge statt. In den meisten Ländern wird die Auswahlentscheidung vor der Zulassung zum Studium getroffen.“

Dieser Stand ist bis heute so geblieben ist. Aktuell haben nach unseren Recherchen 33 europäische Länder eine Zulassung nach Eignung und Leistung (Belgien und die Schweiz sind seit 1996 dazugekommen). Die USA, Kanada, Japan und Australien haben ebenfalls einen NC und wenden Eignungskriterien an (USA, Australien und Japan haben auch einen Test ähnlich dem EMS).

In Deutschland kommt es zu einer Wiederbelebung des Interesses an Eignungstests. Die Zulassung soll zukünftig auf Bundesländerebene geregelt werden und einige Bundesländer bereiten den Einsatz von Tests wieder vor. Ursache ist die Unzufriedenheit mit den Abiturnoten, die offenbar die Eignung zum Medizinstudium nicht spezifisch genug ausdrücken. Auch in anderen Disziplinen mit Numerus Clausus steht die Einbeziehung anforderungsbezogener Eignungstests zur Diskussion.

2. Welche Modelle der Zulassungsbeschränkungen gibt es?

Die folgende Tabelle stellt Zulassungskriterien zusammen, die in verschiedener Kombination angewendet werden; Ihre Vor- und Nachteile sind hier aufgelistet. Nicht aufgeführt sind dabei **Schulwissensprüfungen**, die vor allem in Ländern mit heterogenen Bildungsvoraussetzungen angewendet werden (wodurch praktisch auch die Maturität als Fähigkeitsnachweis etwas entwertet wird). Da die Maturität in der Schweiz vorausgesetzt wird, ist dies hier nicht notwendig.

Der Eignungstest ist dabei das einzige Verfahren, welches alle wichtigen Anforderungen an ein NC-Kriterium erfüllt:

- Wissenschaftlich ist mehrfach nachgewiesen, dass die **Eignung** erfasst wird (erfüllt also die vorgegebenen gesetzlichen Grundlagen der Kantone);
- Es ist eine kapazitätsentsprechende Zulassung punktgenau möglich;
- Rechtsgleichheit und Willkürfreiheit (eine Forderung des Bundesgerichtes für den Numerus Clausus) werden durch die standardisierten und damit immer gleichen Bedingungen bei Durchführung und Bewertung der Ergebnisse garantiert;

¹ Literatur: Eignungskriterien bei der Zulassung zum Medizinstudium in Europa: Ergebnisse einer Erhebung in allen europäischen Ländern; G. Trost in: Eignungsprüfung für das Medizinstudium, Kriterien und Testverfahren, Bericht über das Internationale Symposium in Bern 8. November 1996, herausgegeben von K.-D. Hänsgen und N. Ischi, [Bericht 3 ZTD](#)

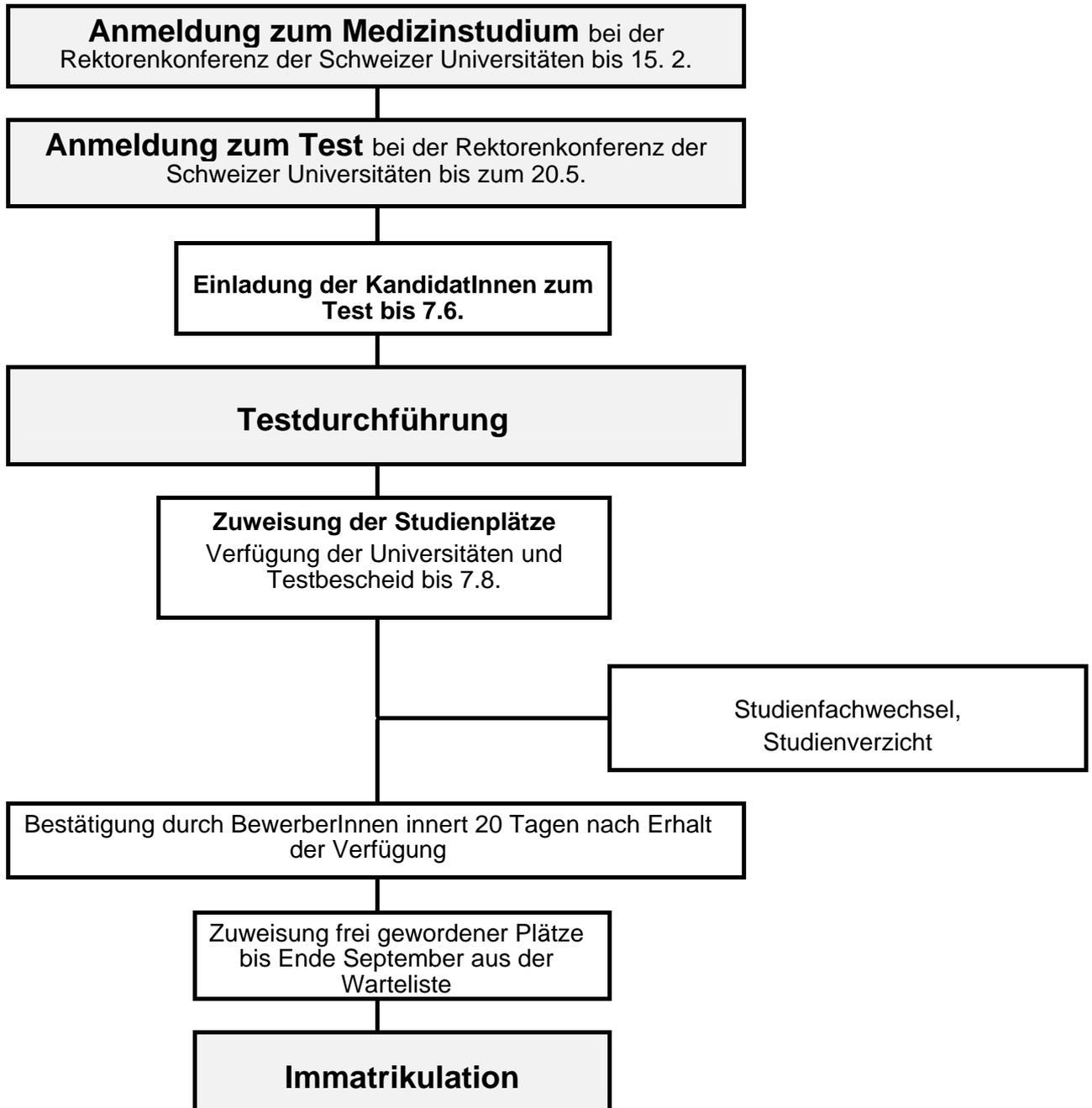
- Der Test ist zugleich ökonomisch vertretbar.

Auch bei allen Veränderungen im Rahmen der Bologna-Deklaration (BA für Medizin) sollte man bedenken, dass ein Medizinstudium ohne Klinik und Praxis nicht möglich ist. Es scheint eine Errungenschaft der Studienreformen, Praxisbezug so früh als möglich herzustellen und bereits im ersten Jahr mit der Entwicklung der oft geforderten Berufseignung zu beginnen. Da die Klinikkapazitäten nicht beliebig erweiterbar sind, kann man die Notwendigkeit eines Numerus Clausus vielleicht hinauszögern (um den Preis eines geringeren Praxisbezuges), nicht aber verhindern. Ein „versteckter NC“ innerhalb des Studiums wäre notwendig. Das Medizinstudium ist, im Unterschied zu vielen anderen Studiengängen, letztendlich auch eine hochspezialisierte Berufsausbildung, die man vom ersten Tage an auf dieses Ziel ausrichten sollte. „Umwege“ würden vielleicht auch das Abschlussalter der Absolventen erhöhen und die Zeitdauer der möglichen Berufstätigkeit verringern.

	Eignungstest	Maturitätsnote	Aufnahme- gespräche	Sozialpraktika	Intrauniversi- täre Selektion
Erfasst Eignung?	Ja, wissenschaftlich nachgewiesene Vorhersage von Studiendauer und Studienleistungen	Möglich, nicht nachgewiesen Schweiz: schwerer vergleichbar	Möglich, wenn standardisierte Beurteilung, mehrere Beurteiler und Training dieser Beurteiler	Nein, dient mehr dem Erkennen der Neigung. Betreuer würden durch Notwendigkeit der Eignungseinschätzung überfordert – Anforderung wie Aufnahmegespräche!	Ist dem Test nicht überlegen!
Kapazitätsentsprechende Zulassung?	Ja	Nein, zu wenig Differenzierung der Werte, um ausreichend zu differenzieren	Ja	Nein, alle erfolgreichen Praktikanten wollen dann auch studieren	Weiter Überlastung der Kapazität im 1. Jahr, weiter grössere Abbrecherzahl nach 1 Jahr; Prüfung ist versteckter NC
Rechtsgleichheit	Ja	Nein kantonale, schul- und lehrerbedingte Benotungsunterschiede	Nur bei Training der Beurteiler und Standardisierung der Bedingungen (aufwändig)	Nur wenn wirklich gleiche Praktikums- und Beurteilungsbedingungen bestehen würden (Illusion)	Ja
Willkürfreiheit	Ja				Wenn Prüfungen vergleichbar sind
Ökonomie	Ja	Ja	Nicht für alle realisierbar (dann zu teuer)	Nein (in Zürich erwies es sich als nicht durchführbar wg. fehlender Plätze)	Hohe Kosten im 1. Jahr
Negative Folgen bekannt?	Nein	Rückwirkung auf Benotungssystem der Maturität	„Trainingsmarkt“ würde wachsen, weil sehr gut trainierbar	Geschultes Betreuungspersonal nötig (Kosten)	Versteckter NC

3. Wie ist das Vorgehen bei der Durchführung von Zulassungsbeschränkungen in der Schweiz?

Nachfolgend ein einfaches Schema für 2004: Bis 15.2.2004 erfolgte die Anmeldung bei der Hochschulrektorenkonferenz (CRUS). Erst danach wurde durch die Universitätskonferenz (SUK) entschieden, ob und wo der Numerus Clausus und damit der Eignungstest notwendig sind.



4. Warum wurde dieses Vorgehen (und kein anderes) gewählt?

Dies war ein Entscheid der damaligen Schweizerischen Hochschulkonferenz (SHK) nach Abwägung aller Möglichkeiten im Jahre 1995 (Vorbereitung) bzw. 1998 (erster Einsatz) als optimale Variante für die Schweiz.

- Durch die preisgünstige Wiederverwertbarkeit des überprüften und bewährten deutschen TMS ergab sich ein besonderer Vorteil. Eine Eigenentwicklung wäre wesentlich teurer gewesen.
- Eignungsgespräche und Sozialpraktika erwiesen sich als zu teuer bzw. nicht umsetzbar (Betreuungsaufwand, Schulungsaufwand). Maturitätsnoten sind nicht vergleichbar und man befürchtete auch Rückwirkungen auf die Notengebung, wenn diese einen selektiven Charakter bekämen.
- Eine **intrauniversitäre Selektion nach dem ersten Jahr** überlastet die Universitäten mindestens während dieser Zeit und es wäre in dieser Zeit noch kein Reformstudium umsetzbar. In der Westschweiz, wo man so verfährt, ist die Kapazitätsüberschreitung nicht so drastisch wie in ZH und BE, sowie neuerdings auch in BS. Es würde wesentlich mehr Personen geben, welche die Prüfung nicht bestehen „dürfen“ (weil die Bestehensquote als „versteckter NC“ wirken müsste), um in den folgenden Studienjahren ein vertretbares Betreuungsverhältnis zu finden. Die Zahl der Patienten und Betten kann in Universitätskliniken nicht unbegrenzt gesteigert und an einen zu hohen Ausbildungsbedarf angepasst werden. Nach dem Nichtbestehen der Prüfung muss es die Möglichkeit der Prüfungswiederholung geben. Es würden dadurch wesentlich mehr **Repetenten** vorhanden sein, welche die Ausbildungskapazitäten zusätzlich belasten.

5. Welche Auswirkungen hätte eine unbeschränkte Zulassung zum Medizinstudium allgemein?

Es bestünde Zulassungsnotwendigkeit **aller im Februar Angemeldeten**. Selbst bei Rückzügen von 20% bis zur Platzvergabe bedeutet dies, dass in Humanmedizin **für NC-Universitäten die Kapazität neuerdings um ca. 100%** überschritten würde (Personen, denen nach der Anmeldung auch ein Studienplatz zugewiesen werden müsste – 2004 waren es 1360 Anmeldungen auf 546 Plätze).

Diese Situation ist nicht mehr vergleichbar mit der Lage Mitte der 90er Jahre, als es „nur“ Überschreitungen **um 30%** gab und der NC als „Notmassnahme“ eingeführt wurde. Wenn man damals bereits einen dringenden Handlungsbedarf erlebte, weil die Studienbedingungen zum Teil chaotisch waren, so würde heute noch wesentlich mehr „Entscheidungsdruck“ entstehen.

Auch Umleitungen wären heute ohne EMS praktisch kaum noch möglich, da die Nachfrage in Basel, Bern und Zürich die Kapazitäten so deutlich überschreitet, dass sie von Freiburg nicht „aufgefangen“ werden kann. Bei gesamtschweizerischer Betrachtung würde auch Neuenburg als weitere nicht ausgelastete Universität dafür nicht ausreichen.

Eine gewisse Gefahr besteht darin, dass durch den seit sieben Jahren wirkungsvollen NC dieser „Druck“ etwas in Vergessenheit gerät, weil die Lage stabil ist. Hier genügt es allerdings, sich in den Nachbarfakultäten umzusehen (z.B. Psychologie), wo weiterhin die chaotischen Bedingungen vorherrschen.

Ohne NC wären vor allem **Studienreformen nicht möglich**, die mit mehr Betreuungsaufwand verbunden sind. Die Rückkehr zum „Massenstudium“ wäre zumindest im ersten Jahr zwingend. Dies wäre wohl der grösste Einschnitt.

Man würde zu intrauniversitärer Selektion gezwungen sein, was zu den unter Punkt 4 geschilderten Nachteilen führen würde.

Auf alle Kantone kämen **massive finanzielle Belastungen** zu. Wenn ein vorklinischer Medizinstudienplatz allein nur mit 23.000 SFR pro Jahr nach den Pauschalbeträgen der IUV gerechnet wird, dann entsprächen die 2004 zusätzlich notwendigen 500 Studienplätze einer Summe von **11.5 Millionen SFR**, die von den Kantonen aufzubringen wären.

6. Welche Auswirkungen hätte es, wenn eine Universität im Alleingang beschliessen würde, die Zulassungsbeschränkungen nicht mehr weiterzuführen?

Wenn dies in der Deutschschweiz eine Universität täte, würden sich mit hoher Wahrscheinlichkeit viel mehr Personen dort anmelden. Das würden alle tun, die ein Studium in der Schweiz in deutscher Sprache absolvieren und ihren Studienplatz sicher haben wollen. Es gäbe dann keine Möglichkeit, diese zurückzuweisen. Wenn 2004 etwa 500 Personen keinen Platz finden, so kann man annehmen, dass dies mindestens die Zunahme wäre.

7. Wie viele Personen wurden in der Vergangenheit aufgrund der Testergebnisse gesamtschweizerisch pro Jahr an der Aufnahme des Medizinstudiums effektiv gehindert?

Es wurde niemand wirklich daran gehindert, Medizin in der Schweiz zu studieren. In jedem Jahr bestand die Möglichkeit, ein Studium in Genf, Lausanne oder Neuenburg aufzunehmen, wo keine Zulassungsbeschränkungen gelten. Das Problem besteht in der Überlastung der deutschsprachigen Universitäten, die daher den NC einführen mussten.

Wie zeigt sich die **dissuasive Wirkung des Tests**? Der Test ist eine Herausforderung, der sich zu stellen ohne Zweifel ein Mindestmass an Selbstsicherheit und Leistungsbereitschaft voraussetzt. Diese dissuasive Wirkung ist dabei gewollt – sich dem Test als Herausforderung zu stellen, ist selber bereits ein Eignungskriterium für spätere berufliche Anforderungen. Dissuasive Wirkungen zeigen sich dabei in

- Nichtanmeldungen zum Studium wegen dem EMS-Test (eine nicht kalkulierbare Zahl von Personen);
- Nichtanmeldung zum Test, Rückzug vor der Testanmeldung (siehe statistischer Bericht);
- Nichtantritt zum Test trotz Anmeldung (siehe statistischer Bericht);
- Nichtakzeptieren der sich möglicherweise ergebenden Umleitungen, weil eine Zulassung an der Wunschuniversität nicht möglich ist (hier sind besonders die Umleitungen nach Freiburg zu erwähnen, die weniger häufig akzeptiert werden).

Theoretische Zahlen für Ablehnung (Zahl der Bewerbungen mit gültigem EMS-Ergebnis minus Kapazitäten, Stand Anfang Juli jeden Jahres):

Disziplin	1998	1999	2000	2001	2002	2003	2004
Humanmedizin	130	132	124	153	157	309	505
Veterinärmedizin		32	19	11	20	58	59
Zahnmedizin							10

Durch Rückzüge (nicht akzeptierte Umleitungen etc.) werden nach der ersten Studienplatzzuweisung noch einmal Plätze frei, die aus der Warteliste wiederum nach der Testleistung vergeben werden. Aufgrund der jahrelangen Erfahrungen wird bereits in der ersten Zuteilungsrunde die Kapazität spezifisch für die Universitäten so überschritten, dass bereits danach die Universitäten ausgelastet sind und in einer zweiten Zuteilungsrunde nur wenige Personen noch nachrücken.

Zählt man alle Personen als „nicht abgewiesen“, denen innerhalb dieses Verfahrens einmal ein Studienplatz angeboten werden kann (auch wenn dieser nicht zusagt), dann verringert sich die Zahl der Abgewiesenen: 1998 z.B. auf 81, 1999 (ab da mit Veterinärmedizin) auf 78, 2002 auf nur 49, um 2003 auf 236 zu steigen. Die Zahl der „**tatsächlich Abgelehnten**“ liegt also unter den Zahlen der Tabelle (Rücktritte von Personen, die einen Studienplatz erhalten hätten, der aber nicht ihren Wünschen entsprechend war, zählen nicht als Ablehnung).

8. Wie lauten die diesbezüglichen Prognosen für den Zeitraum 2005 bis 2008?

Seit 4 Jahren verzeichnen wir einen kontinuierlichen Anstieg der Bewerberzahlen (siehe statistischer Bericht). Zunächst führte man dies auf Doppelmaturitäten zurück – der erneute Anstieg 2004 in Bern und die Zunahme in Basel sind dadurch aber nicht mehr erklärbar.

Die neueste Prognose des BFS¹ bestätigt nun, dass die Studienanfängerzahlen nach der Zunahme wegen Doppelmaturität **nicht mehr auf das Ausgangsniveau zurückgehen**, sondern dass es **in den nächsten 10 Jahren zu einem weiteren Anstieg der Studierendenzahlen** kommt. Der massive Anstieg 2004 geht danach nicht mehr zurück, sondern es wird wahrscheinlich noch mehr Bewerbungen geben. Mit einer Entspannung „von selbst“ ist also nicht zu rechnen.

Selbst wenn man die Kapazitäten in einem realistischen Rahmen erhöhen würde, um dem steigenden Medizinerbedarf zu begegnen, ist mit einer Überlastung der Universitäten zu rechnen und das Problem wird sich nicht „von allein“ lösen. Dabei scheint der Medizinerberuf trotz der beschränkten Niederlassungsmöglichkeiten in freier Praxis an Attraktivität zu gewinnen. Der erhöhte Bedarf an Krankenhausärzten vermittelt dem Arztberuf vermutlich eine gesicherte(re) Zukunftsperspektive.

9. Gegenwärtig wird ein höherer Medizinerbedarf vor allem in Krankenhäusern erwartet. Muss man deshalb den NC nicht abschaffen?

Beide Fragen sollten voneinander unabhängig diskutiert werden. Selbst wenn die Kapazitäten in realistischer Dimension erhöht würden, bliebe eine Überlastung der Universitäten voll erhalten, welche 1998 zum Entscheid eines NC führte.

In den Zeiten der Doppelmaturitäten wurden die Kapazitäten in der Humanmedizin beispielsweise maximal ausgeschöpft (**648 statt gegenwärtig 546**). Wenn man zu diesen Maxima zurückkehren würde, wären 2004 immer noch **403 Personen** nicht mit einem Studienplatz versorgt – also mehr als in Bern und Zürich zusammen als Kapazität überhaupt vorhanden ist. 1998 waren 130 Personen nicht mit einem Studienplatz versorgt und diese Überlastung war so gross, dass ein NC eingeführt werden musste. Da mit einer Entspannung der Bewerberzahlen nicht zu rechnen ist, wird sich daran wohl nichts ändern und man wird auf ein regulierendes Zulassungskriterium angewiesen sein.

10. Gibt es Untersuchungen darüber, in welche Studienbereiche diese Personen ausweichen?

¹ Neue Hochschulprognosen. Zunahme der Anzahl Studierender an den Hochschulen bis etwa 2013, Nr. 0350-0406-30 Neuchâtel, 29. Juni 2004, Bundesamt für Statistik

In diesem Zusammenhang belegt die Entwicklung bei den Zahlen der Anmeldungen für die Universitäten Genf, Lausanne und Neuenburg, dass die Personen NICHT in die Westschweiz ausweichen, um dort Medizin zu studieren. Systematische Untersuchungen sind uns nicht bekannt.

11. Welche Möglichkeiten stehen diesen Personen offen, um doch noch Medizin studieren zu können?

Eine erneute Bewerbung (mit ggf. erneuter Testabsolvierung) oder die Bewerbung in Genf, Lausanne oder Neuenburg (ohne NC) und ein Studium auf Französisch stehen offen. Letzteres findet weniger Anklang (siehe statistischer Bericht). **Theoretisch kann aber JEDE/R Medizin studieren, die/der es will, muss es aber in einer anderen Sprache tun.** Auf die Erfüllung des Zulassungskriteriums wird im Übrigen streng geachtet – das Verfahren ist fair, es gibt kein „Überlisten“ (wichtig für die Fairness). Auch bei einem Disziplinwechsel von einer medizinischen Disziplin ohne NC in eine solche mit NC oder für einen Übertritt von der Universität Neuenburg an eine Universität mit NC muss der EMS absolviert werden.

12. Wären die Anmeldezahlen höher, wenn Zulassungsbeschränkungen nicht bestehen würden?

Sehr wahrscheinlich wären die Zahlen deutlich höher, der EMS hat einen gewollt dissuasiven Charakter (siehe oben). Vergleiche dazu auch den letzten Punkt im statistischen Bericht 2004.

- (1) war der Zuwachs der Bewerbungszahlen in Medizin bisher im Trend geringer als der Zuwachs in anderen Studien-Disziplinen und demgegenüber auch verzögert.
- (2) hat z.B. die Abschaffung des TMS-Tests in Deutschland (wohlgemerkt galt auch danach immer noch ein NC!) schnell zu einer Verdoppelung der Bewerbungszahlen geführt.
- (3) wächst auch international die Nachfrage nach einem Medizinstudium.
- (4) sind die Anstiege 2003 und 2004 in der Schweiz eventuell auch darauf zurückzuführen, dass nunmehr die dissuasive Wirkung etwas nachlässt, da genügend Erfahrungen mit dem EMS vorliegen.

13. Wie erfolgt die Finanzierung der Zulassungsbeschränkungen?

Das ZTD wird durch die IUV (Interkantonale Universitätsvereinbarung) aus Zinseinnahmen (Beiträge der Kantone) finanziert. Testkosten (einschliesslich der Weiterentwicklung) und Durchführungskosten werden aus der Kostenbeteiligung der Teilnehmenden (200 SFR pro Person) gedeckt. Das zuständige Personal der CRUS für das Anmelde- und Zulassungsverfahren wird durch den Bund und die Universitätskantone finanziert.

14. Wie hoch sind die Kosteneinsparungen aufgrund der Zulassungsbeschränkungen grob einzuschätzen? Inwiefern trägt der EMS zu Kosteneinsparungen bei?

Nachgewiesene Prämisse: Die Evaluation (s.u.) zeigt, dass Personen mit schlechteren Testwerten länger studieren und schlechtere Abschlussnoten erreichen. Die Zulassung erfolgt tatsächlich nach der Eignung. Wir können sicher sein, dass die Chancen der Abgelehnten drastisch schlechter sind, als die der Zugelassenen (siehe Analyse der über die Warteliste Zugelassenen).

Das Ziel des NC besteht darin, die vorhandenen Ausbildungsressourcen für Medizin optimal zu nutzen. Diese sind nicht beliebig erweiterbar. Es sollen diejenigen Personen zugelassen werden, die (1) tatsächlich das Studium abschliessen, (2) dies mit guten Leistungen tun und (3) dies in einer angemessenen Zeit tun (und die Kapazitäten wieder freigeben). Diese drei Kriterien bestimmen die Eignung.

Sehr streng gerechnet: Wenn jeder Abbrecher nur ein Semester studieren würde, wären bei Semesterkosten um 50.000 SFR schon bei ca. 15 bezüglich des Erfolges richtig vorhergesagten Personen die Kosten des gesamten EMS überschritten (Kosten für Test und ZTD-Aufwand).

Ohne NC müsste man alle Personen erst einmal zulassen. Wenn dies wie 2004 ca. 500 Personen mehr wären und allein wegen der beschränkten Kapazitäten in den späteren klinischen Fächern ein „versteckter“ Numerus Clausus während des Studiums zur Anwendung kommen müsste, würden die Kosten für die dann zu erwartende grössere Abbrecherzahl sehr schnell in astronomische Bereiche steigen.

Ins Gewicht fällt auch das Ausmass der Selektion, das sich aus dem Verhältnis von Bewerbungen zu vorhandenen Studienplätzen ergibt. Die „Verschärfung“ des NC der letzten beiden Jahre wird sich noch weiter positiv auf die Kosteneinsparung auswirken: Der zulassungsrelevante Testwert steigt an und es werden dann deutlich mehr Personen zugelassen, die noch bessere Chancen haben, das Studium in angemessener Zeit und mit guten Leistungen abzuschliessen. Bei einem „schwächeren“ NC werden Personen mit geringerer Eignung zugelassen bzw. gelangen über die Warteliste zu einem Studienplatz (siehe Evaluation).

15. Wurde der EMS immer gleich gestaltet?

Von 1998 bis 2003 wurde immer die gleiche Teststruktur verwendet. Die Aufgaben sind aber jedes Jahr andere (äquivalente Testformen). Im Jahre 2003 fand ein Workshop „Weiterentwicklung des EMS“ in Freiburg statt, um zukünftige Anpassungen zu besprechen. Jedes Testkonzept muss laufend auf die Übereinstimmung mit den Anforderungen überprüft werden und nicht zuletzt dank der Einführung des Numerus Clausus konnten Studienreformen umgesetzt werden. Bewährtes soll im Test dabei erhalten bleiben. Im Jahr 2004 gab es eine erste Veränderung des Untertests „Konzentriertes und Sorgfältiges Arbeiten“, er wurde modernisiert und ist jetzt auch weniger anfällig für Übungseffekte.

Geplant sind weitere neue Untertests, die planerische und organisatorische Kompetenzen erfassen. Die Entwicklung erfolgt gemeinsam mit dem Institut für Test- und Begabungsforschung in Bonn.

16. Welche Evaluationen der Tests wurden durchgeführt?

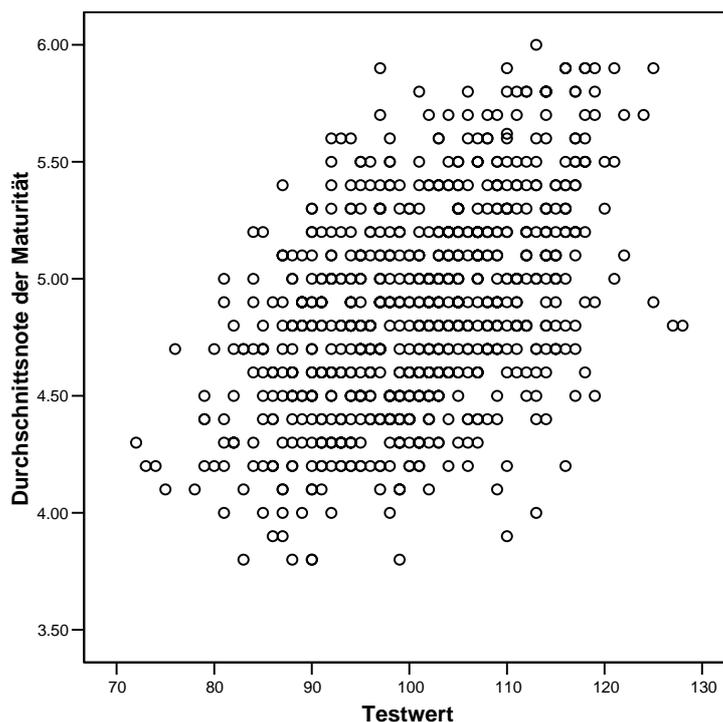
Der EMS dürfte mit zu den am besten evaluierten Tests auf diesem Gebiet überhaupt gehören: Zahlreiche Untersuchungen liegen aus Deutschland zu allen wichtigen Fragen vor und fanden Bestätigung durch Befunde aus der Schweiz. Es erfolgen umfassende jährliche Evaluationen der Durchführung, die in den Berichten 4 bis 10 des ZTD veröffentlicht worden sind. Ausserdem gibt es gesonderte Berichte für verschiedene Fragestellungen: Übersicht: http://www.unifr.ch/ztd/ems/berichte/welcome_b.htm

Zudem gibt es Studien zu Sonderfragestellungen:

- Erste Evaluationsstudie zum Zusammenhang Testwert und Studienerfolg (in Bericht 6 des ZTD, 2000), sowie [Bericht über die Evaluation des EMS 2000](#)
- Zweite Evaluation zum Zusammenhang Testwert und Studienerfolg (in Bericht 7 des ZTD 2001), sowie
- Hänsgen K.-D., Spicher B. [Numerus Clausus: Finden wir mit dem «Eignungstest für das Medizinstudium» die Geeigneten? \[1653\] Schweizerische Ärztezeitung 2002](#)
- Hänsgen K.-D., Spicher B. [Numerus Clausus: le «test d'aptitudes pour les études de médecine» \(AMS\) permet-il de trouver les personnes les plus aptes? Bulletin des médecins suisses 2002](#)
- Erste Befragung der Kandidaten zur Vorbereitung auf den EMS <http://www.unifr.ch/ztd/ems/vorbereitungsreport2003.pdf>

Die Evaluation wird kontinuierlich weitergeführt, es läuft die dritte Evaluation zum Zusammenhang Testwert und Studienerfolg. Summa summarum kann mittlerweile auf sieben Jahre erfolgreicher Durchführung ohne Probleme zurückgeblickt werden.

17. Gibt es Untersuchungen über einen Zusammenhang zwischen Testerfolg und Maturitätszeugnis?



Dies wurde nicht explizit untersucht, weil die Maturitätsleistung nicht als NC-Kriterium zur Diskussion steht.

In der Befragung zur EMS-Vorbereitung 2003 wurden die Maturitätsleistungen erstmals mit erfragt, um Leistungsgruppen nach Schulleistungen zu bilden. Mit den Ergebnissen konnte indirekt bestätigt werden, dass gute Maturitätsleistungen auch mit guten Testleistungen korrelieren ($r=0.48$). Sie sind also auf einen gemeinsamen Fähigkeitsfaktor zurückzuführen, sind aber nicht das Gleiche (siehe die „Punktwolke“ in der nebenstehenden Abbildung). Dabei wurde vor allem die mangelnde

Vergleichbarkeit der Ergebnisse deutlich, es konnten nur drei Leistungsgruppen verlässlich gebildet werden (Unterschiede in den Notenskalen und Masstäben). Für eine kapazitätsgenaue Zuteilung der Studienplätze mit ausreichender Zuverlässigkeit für jede einzelne Person reicht die Differenzierung der Leistungen nicht aus. Ein weiteres Argument gegen die Maturitätsnoten war, dass es Rückwirkungen auf das Benotungssystem geben könnte, indem Noten nicht mehr „unvoreingenommen“ vergeben würden, wenn daran die berufliche Zukunft festgemacht wäre.

18. Gibt es Untersuchungen über einen Zusammenhang zwischen Testerfolg und Schichtzugehörigkeit?

Die Schichtzugehörigkeit wurde bisher nicht erfasst. Aufgrund der Literaturbefunde wäre zu erwarten, dass der EMS „schichtenneutraler“ ist als die Maturitätsnote, da er eher die reinen Fähigkeiten misst und sich Fleisseeffekte wie der reine Wissenserwerb weniger stark auswirken. Es hätten diejenigen Personen mit vergleichbaren Fähigkeiten die gleichen Chancen, auch wenn sie sich durch unterschiedliche „Intensität“ der Wissensaneignung und z.B. Förderung im Elternhaus im vorhandenen Wissen unterscheiden. Dies wäre ein Argument für die Fairness des EMS.

19. Gibt es Untersuchungen über einen Zusammenhang zwischen Testerfolg und Geschlecht?

Von 1998 bis 2003 erreichte der Unterschied beim Testwert zwischen Männern und Frauen nie statistische Signifikanz, sodass entgegen früheren Befürchtungen hier von einer vollen Chancengleichheit beider Geschlechter ausgegangen werden kann. Durch die grösseren Stichproben ist 2004 der Unterschied zwischen den Geschlechtern in Humanmedizin, deutsche Sprachgruppe, erstmals signifikant - aber in Veterinär- und Zahnmedizin sind die Frauen tendenziell sogar besser als Männer, ohne dass der Unterschied signifikant wird.

Es gibt mehrere Argumente, die eigentlich einen Unterschied beim Testwert zugunsten der Männer erwarten liessen (deutlich höhere Bewerberquote der Frauen und Auswirkungen auf die Repräsentativität für die Gesamtpopulation, geringfügig schlechtere Prüfungsleistungen der Frauen gegenüber Männern). Dass dieser bisher nicht vorhanden war, belegt, dass Frauen keinesfalls benachteiligt werden.

Mehr dazu finden Sie im Kapitel „Vergleich der Geschlechter“ in den jährlichen Berichten des ZTD.

20. Gibt es Untersuchungen über einen Zusammenhang zwischen Testerfolg und Sprachzugehörigkeit?

Dieser Frage wurde besondere Aufmerksamkeit gewidmet. Im ersten EMS-Jahr 1998 wurden alle vorhandenen Unterschiede (tatsächliche Leistungsunterschiede, unterschiedliche Repräsentativität der Bewerber für die Gesamtgruppe in den drei Sprachen) noch nivelliert, was Anlass zu berechtigter Kritik gab.

Seit 1999 wird nun das international eingeführte so genannte DIF-Verfahren eingesetzt, welches nur mögliche testbedingte Unterschiede ausgleicht, andere Unterschiede aber in den Daten belässt. Ab Bericht 5 des ZTD wird jährlich ausführlich darüber berichtet; wegen der Komplexität muss hier darauf verwiesen werden.

Ein spezielles Evaluationsverfahren (Vergleich von Tests mit hohen und niedrigen Sprachanforderungen) zeigt, dass die Korrekturen in jedem Falle die testbedingten Unterschiede sicher ausschliessen: Der Unterschied in den vier weniger sprachabhängigen Untertests ist in jedem Falle grösser als der in den fünf sprachabhängigen Untertests.

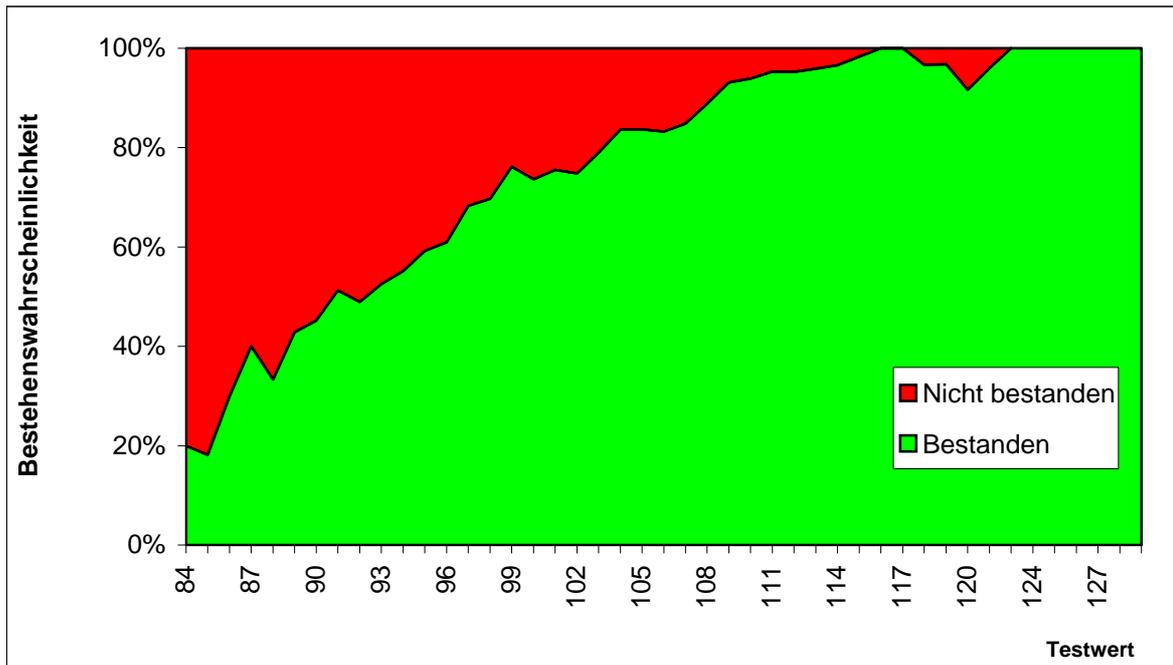
Die Zahl der Aufgaben, die überhaupt ausgeglichen werden mussten, hat über die Jahre deutlich abgenommen, was für die Optimierung der Übersetzungsprozesse spricht.

21. Gibt es Untersuchungen über die Eignung des Tests zur Voraussage der Studieneignung?

Dazu gibt es die speziellen Evaluationsberichte.

Die Ergebnisse der vorliegenden Evaluationen lassen es zu, für den Zeitraum bis zur zweiten ärztlichen Vorprüfung gesicherte Erkenntnisse zum Zusammenhang zwischen Zulassung aufgrund des Eignungstests und der Eignung zu gewinnen.

Schlüsselabbildung: Mit sinkendem Testwert sinken gleichzeitig auch die Chancen, die erste Vorprüfung im ersten Anlauf zu bestehen (was immer mit einer Studienverlängerung verbunden ist).



Der Eignungstest für das Medizinstudium in der Schweiz (EMS) erfasst nachweislich Studieneignung, indem (1) Personen mit höheren Testwerten die Prüfungen mit höherer Wahrscheinlichkeit und zeitlich eher bestehen als die mit niedrigeren und (2) Personen mit höheren Testwerten auch eine höhere Studienqualität erreichen, die sich in besseren Prüfungsnoten ausdrückt.

Schlüsseltabelle: Die Personen der Warteliste, die zur „Kapazitätsauffüllung“ nachträglich zugelassen worden sind (d.h. die eigentlich unter dem Kriterium lagen), haben deutlich schlechtere Chancen, die Prüfung zu bestehen.

Erfolg 1. Vorprüfung				
		Nicht bestanden	Bestanden	n
Testwert unter Kriterium	Anzahl	38	25	63
	Prozentverhältnis	60.3%	39.7%	
Testwert über Kriterium	Anzahl	238	849	1087
	Prozentverhältnis	21.9%	78.1%	
Erfolg 2. Vorprüfung				
		Nicht bestanden	Bestanden	n
Testwert unter Kriterium	Anzahl	1	1	2 (!)
	Prozentverhältnis	50.0%	50.0%	
Testwert über Kriterium	Anzahl	51	293	344
	Prozentverhältnis	14.8%	85.2%	

Der Zusammenhang zwischen Testergebnis und Studieneignung kann auf einem statistisch bedeutsamen Niveau und in jedem Wertebereich des Testwertes festgestellt werden, was die Verwendung des Testwertes als Zulassungskriterium rechtfertigt.

Die Ergebnisse sind für die erste und zweite Vorprüfung gesichert. Die Zulassung von weniger Geeigneten, nur um Kapazitäten auszunutzen, hat nachweislich wenig Erfolgsaussicht: Diese Personen haben um die Hälfte verringerte Erfolgschancen bei der ersten Vorprüfung, bei der zweiten Vorprüfung waren sogar nur zwei Personen aus der „Warteliste“ angelangt.

Diese Aussagen gelten für Human- und Veterinärmedizin, für alle Universitäten und für beide Geschlechter.

22. Sind die Testergebnisse über die Jahre vergleichbar?

Es gehört zu den grossen Vorzügen des EMS, dass er über die Jahre immer eine vergleichbare Struktur hat. Die Faktorenstruktur, die Zuverlässigkeitswerte sind nahezu identisch. Das liegt daran, dass neue Aufgaben jeweils an den älteren geeicht werden. Die Übertragbarkeit der Testwerte in das Folgejahr ist damit gerechtfertigt.

23. Ist der EMS trainierbar?

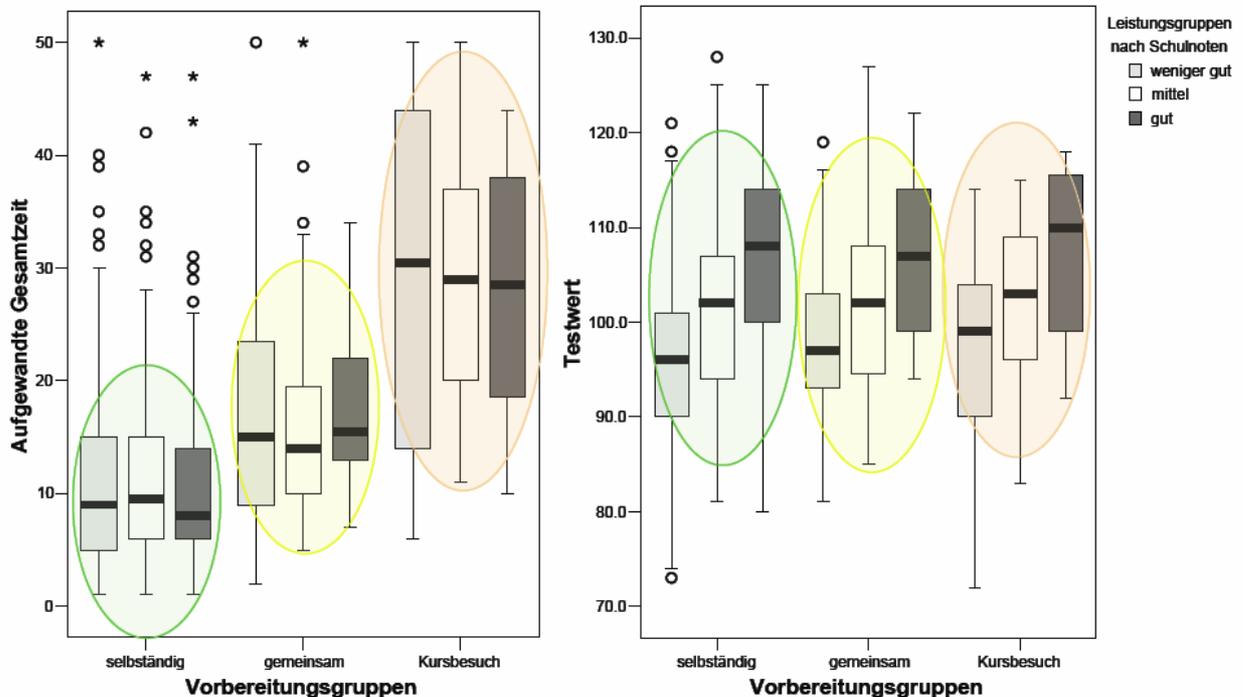
Wir machen einen Unterschied zwischen den Begriffen **Vorbereitung/Übung** und **„Training“**.

Vorbereitung/Übung ist notwendig und wichtig. Dieser Aufwand scheint vertretbar, es handelt sich dabei um keine übermässige Belastung. Man muss die Instruktionen und Aufgabentypen kennen, damit man während der Bearbeitung sofort sein optimales Leistungsniveau findet. Hat man dieses erreicht, gibt es keine weitere Verbesserung.

Der Begriff „Training“ wird in Analogie zum Sport gern von den professionellen Trainingsanbietern verwendet. **Traditionell wird damit suggeriert, dass jede Leistung von jeder Person erreicht werden kann.** Dies wurde eindeutig widerlegt:

Im Jahr 2003 wurde eine Befragung zur EMS-Vorbereitung durchgeführt (vgl. dazu den Vorbereitungsreport <http://www.unifr.ch/ztd/ems/vorbereitungsreport2003.pdf>).

Es wurde deutlich, dass sich bis zu einem bestimmten Zeitaufwand die Testleistung verbessert – danach aber nicht mehr. Die Fähigkeiten (geschätzt durch die Schulleistungen) bestimmen den Testwert – im Verhältnis dazu hat unterschiedlicher Vorbereitungsanfang praktisch einen vernachlässigbaren Einfluss. Dies entspricht der Erwartung: Es wird mit dem EMS kein Wissen geprüft, welches man „pauken“ kann.



Die Auseinandersetzung wird vom ZTD offensiv geführt und scheint auch erfolgreich, damit nicht fälschlich der Eindruck im Raum steht, man könne nur mit teuren Vorbereitungskursen den EMS erfolgreich absolvieren (http://www.unifr.ch/ztd/ems/info_teilnehmer.htm#gering)

24. Welche Vorbereitungsmöglichkeiten gibt es?

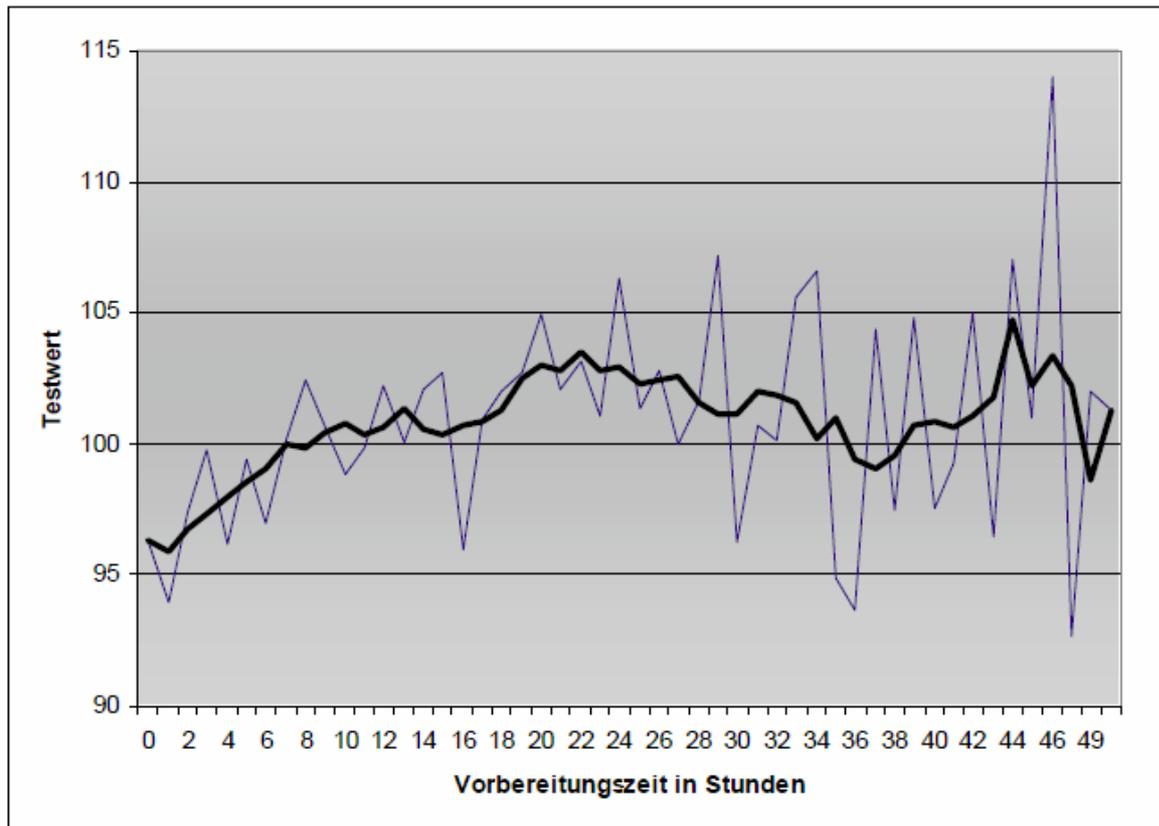
Empfohlene Vorbereitung:

- Durcharbeiten der Test Info (wird jährlich an die Kandidaten versendet)
- Bearbeiten einer veröffentlichten Originalversion unter realen Zeitbedingungen

Zusammenfassung der Ergebnisse der Studie zur Vorbereitung 2003:

- Man muss sich auf den EMS vorbereiten. Ohne Vorbereitung oder mit zu wenig Vorbereitung erreicht man nachweislich geringere Testleistungen und hat geringere Zulassungschancen.
- „Viel hilft dabei nicht viel“: Die Dauer der Beschäftigung allein ist nicht so sehr ausschlaggebend. Es kommt darauf an, wie man sich vorbereitet.
- Die gemeinsame Vorbereitung scheint einer nur selbständigen Vorbereitung etwas überlegen zu sein. Dies gilt vor allem für die Durchführung der Originalversion.
- Ein sehr wichtiges Übungselement ist die **Durchführung der veröffentlichten Originalversion** unter echten Zeitbedingungen – dies bringt den meisten Leistungszuwachs.

- Kommerzielle Trainingskurse sind dem selbständigen oder gemeinsamen Üben keinesfalls überlegen. Es gibt sogar Hinweise, dass sie unter bestimmten Bedingungen kontraproduktiv sein können. Der deutlich höhere zeitliche (und finanzielle) Aufwand wird nicht durch verbesserte Leistungen gerechtfertigt.



Der notwendige Aufwand für die Vorbereitung beträgt ca. 20 Stunden, mehr Zeit bringt nachweislich keine deutlichen Effekte.

Nicht empfohlen:

- Durcharbeiten von dem EMS nachempfundenen Materialien der Trainingsanbieter (die Aufgabenqualität ist zweifelhaft)
- Exzessiver Besuch professioneller Trainingskurse, wenn sie deutlich über die Bearbeitung der Originalversion hinausgehen

25. Gibt es professionelle Anbieter von Trainings? Schneiden die Nutzniesser solcher Angebote im Test besser ab?

Es gibt drei uns bekannte Anbieter, die Trainingskurse professionell durchführen. Diese sind aus Deutschland und haben dort bereits Kurse für den TMS angeboten, welche dort toleriert wurden. Bereits zwei deutsche Studien zeigten aber, dass die gezielte individuelle Vorbereitung dem Besuch professioneller Kurse ebenbürtig ist. Die Befragung 2003 der Schweizer EMS-Teilnehmer bestätigte dieses Resultat: Es gibt keine Leistungsunterschiede, nur Unterschiede im Aufwand (siehe oben).

Der EMS ist sozial verträglich, indem keine hohen Kosten für die Vorbereitung aufzuwenden sind.

Es ist eine wichtige Aufgabe, darüber zu informieren: Das ZTD klärt auf seiner Homepage <http://www.unifr.ch/ztd/ems/> auch über diese Fragen auf.

26. Mit welchen Mitteln könnte die Berufseignung getestet werden?

Die Studieneignung bleibt die wichtigste Voraussetzung der Berufseignung, beides sind keine Gegensätze. Studieneignung muss vorliegen, Berufseignung kann und soll sich v.a. im Studium weiter entwickeln und ist zu Studienbeginn noch nicht ausreichend differenziert (es ginge allenfalls um **Voraussetzungen für die Ausprägung einer Berufseignung**).

Für **sozial-kommunikative Fähigkeiten** gibt es nach wie vor nur das Interview bzw. Aufnahmegespräch, welches annähernd dieses Merkmal erfasst. Minimal dauert ein solches Gespräch mindestens 30 Minuten mit 3 Beurteilern, welche ein ausreichendes Training machen müssten (auch Vergleichbarkeit zwischen den Teams ist herzustellen). Dieses für alle Bewerber anzuwenden wäre praktisch unbezahlbar.

Es gab einen Vorstoss in Belgien mittels Video, um Kommunikation einzuschätzen. Trost stellte 2003 die Ergebnisse vor mit dem Resultat: Der Test ist extrem aufwändig, bisher ist keine Beziehung zur Studieneignung feststellbar. Der Zusammenhang zur Berufseignung wurde noch nicht untersucht. Deshalb wird dies vorerst nicht weiter verfolgt.

Ab 2005 wird es aber aufgrund der Rückmeldung der Fakultäten einen neuen Untertest geben, der die Fähigkeit „Planen und Organisieren“ als eine der wichtigsten neuen, bisher unberücksichtigten Anforderungen realisiert. Dieser dürfte auch für Berufseignung wichtig sein. Weitere Erweiterungen und Anpassungen sind in Vorbereitung.

4 Numerus Clausus (NC) und Medizinstudium

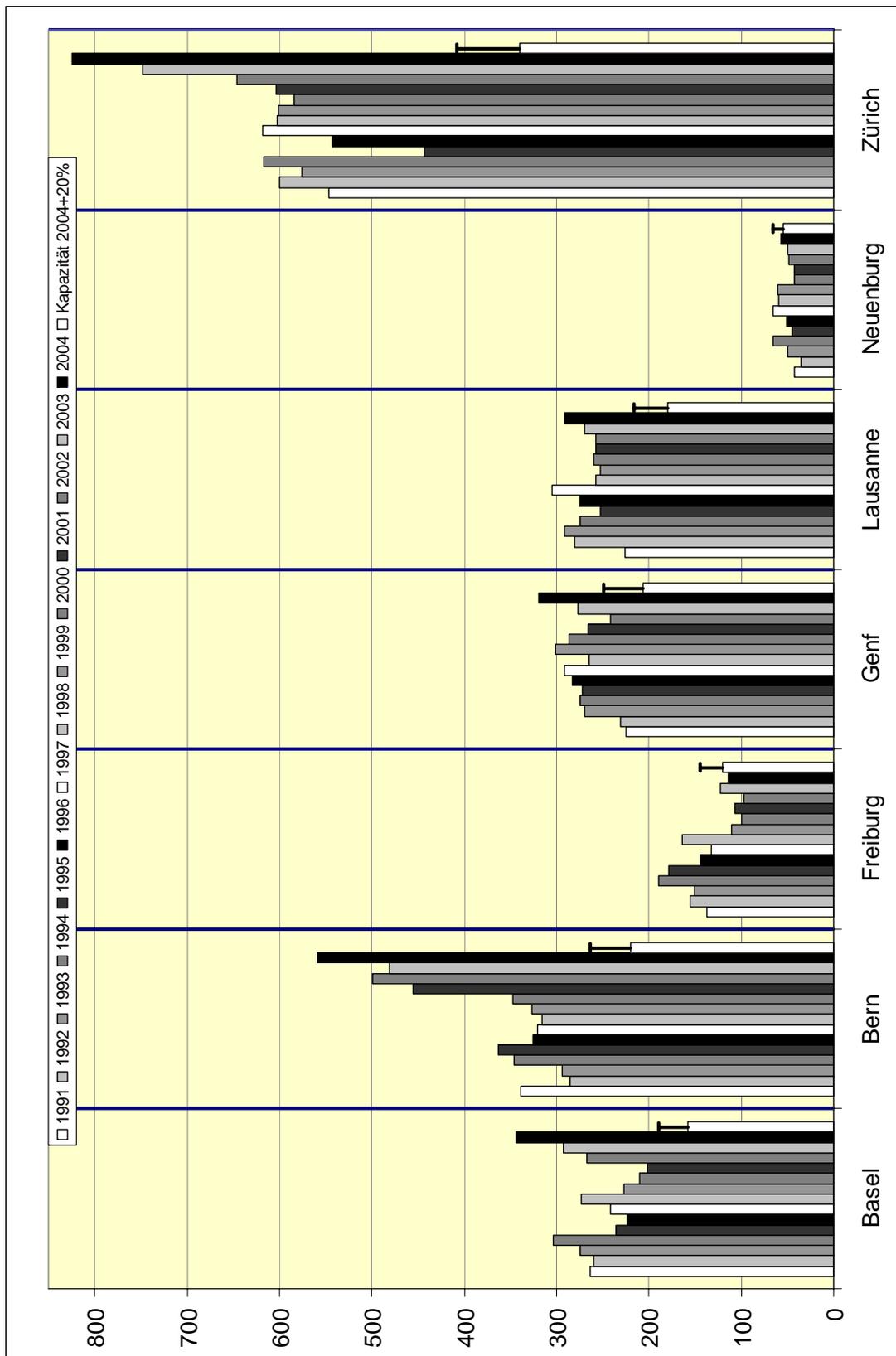


Abbildung 1: Anmeldungen zum Medizinstudium 1991 bis 2004, (letzte Säule Kapazität 2004 – Markierung des 20%-Überschreitungskriteriums als Grenze für NC)

Die Abbildung 1 verdeutlicht die Anmeldeverläufe 1991 bis 1997 (ohne NC) sowie 1998 bis 2004 (mit NC). Die letzte Säule zeigt die Kapazitäten 2004 und das 20%-Überschreitungskriterium, ab welchem ein NC zur Anwendung kommt.

In Basel, Bern und Zürich steigen die Anmeldezahlen weiter deutlich an. Ein Zusammenhang mit den Jahrgängen der Doppelmaturität ist nicht mehr erkennbar, die Nachfrage steigt unabhängig davon weiter an.

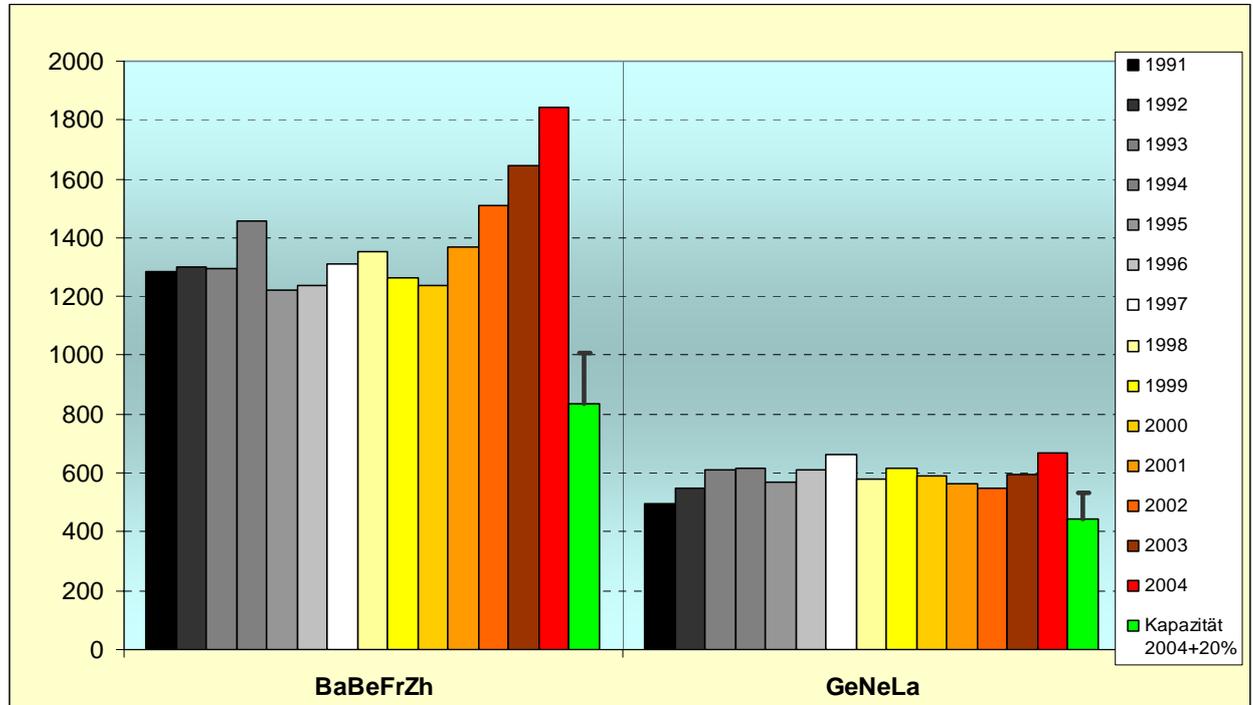


Abbildung 2: Anmeldungen zum Medizinstudium 1991 bis 2004 nach Gruppen (**Basel, Bern, Freiburg, Zürich** vs. **Genf, Neuenburg, Lausanne**)

Seit nunmehr vier Jahren steigt die Zahl der Anmeldungen zum Medizinstudium in den am NC beteiligten Universitäten an. Auch in der Westschweiz steigt die Zahl der Bewerbungen erneut. Es ist wie bereits beschrieben davon auszugehen, dass die Attraktivität des Medizinstudiums wieder zunimmt, was besonders in der Deutschschweiz zu vermehrten Studienbewerbungen beiträgt.

Die folgende Abbildung zeigt, dass mit einer Entspannung der Situation nicht zu rechnen ist. Das BFS veröffentlichte eine Prognose für die Studierendenzahlen der Schweiz. Eine wichtige Aussage ist, dass der nach den Jahrgängen mit Doppelmaturitäten vorhandene Nachfrageschub nicht mehr absinkt, das hohe Niveau bleibt erhalten.

Setzt man die Zahl der Anmeldungen bei der CRUS (bzw. früher der SHK) im Jahre 1998 zu 100% ergibt sich zunächst für 1999 und 2000 ein relativer Rückgang der Anmeldezahlen. Dass es sich hier um einen Dissuasionseffekt durch die Neueinführung des EMS handelt, wird durch das Gleichbleiben bzw. leichte Ansteigen der Studieneintritte 1999 und 2000 nahe gelegt.

Seit 2001 steigen die Anmeldezahlen jedes Jahr deutlich an – bis 2003 mit etwa gleichem Anstieg wie bei der Zahl aller Studieneintritte. Wenn es sich nur um eine einjährige Verzögerung bei den Anmeldungen zum Medizinstudium gegenüber den allgemeinen Studieneintritten handelt, würde für die Jahre 2005 und darauf folgende ein Gleichbleiben der Bewerberzahlen auf diesem hohen Niveau erwartet.

Möglich wäre es auch, dass das zunehmende Interesse für Medizinberufe ein weiteres Ansteigen bedeutet. Am wenigsten wahrscheinlich ist, dass sich die Situation von selbst entspannt.

Auch wenn die Zahl der Studienplätze erhöht würde, werden auf lange Sicht deutlich mehr Bewerbungen erfolgen, als Kapazitäten vorhanden sind. Es sollte damit gerechnet werden, dass ein Numerus Clausus und damit der EMS noch für längere Zeit notwendig bleiben.

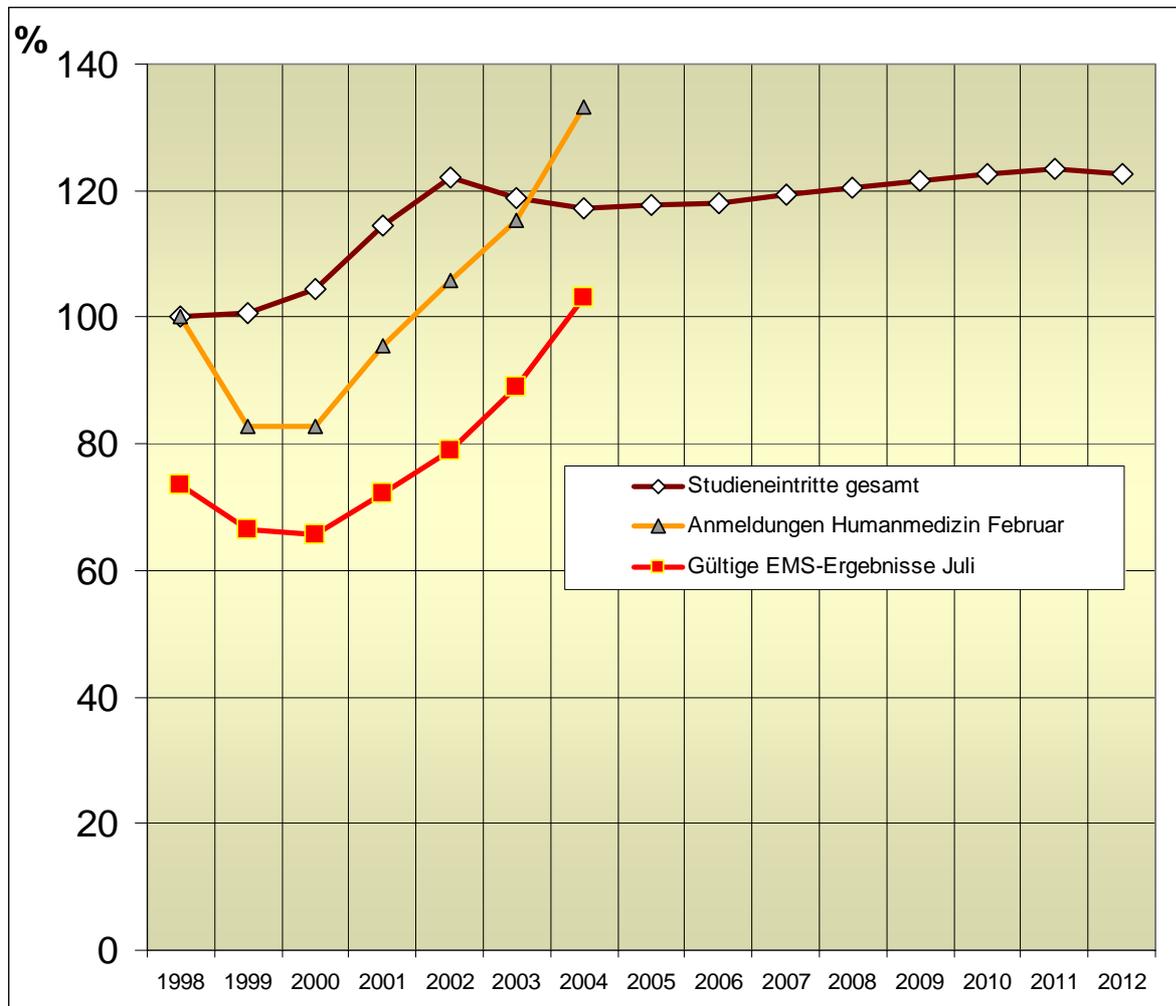


Abbildung 3: Studieneintritte Schweiz Universitäten und Hochschulen (indexiert, 1998 = 100%) im Vergleich zu Studienanmeldungen Februar und Bewerbungen mit gültigem EMS (indexiert, Anmeldungen 1998 = 100%) für Humanmedizin

5 Anmeldung zu Medizinstudium und Test 2004

5.1 Anmeldungen, Testantritte und Zulassungen

Im Jahr 2004 haben sich erstmals mehr als 200% Personen in Humanmedizin um einen Studienplatz auch nach dem Termin der Testanmeldung beworben, als Kapazitäten vorhanden sind. Die Rückzugsquote zwischen Anmeldung zum Studium und der Situation im Juli bleibt mit knapp 23% auf dem Niveau des Vorjahres.

	Humanmedizin						
	1998	1999	2000	2001	2002	2003	2004
Kapazitäten (am NC beteiligte Univ.)	620	546	546	583	648	598	546
Anmeldungen bei CRUS im Februar	1020	845	844	974	1080	1177	1360
... in % zu Kapazität	186	155	154	167	185	197	249
Anmeldungen EMS im Mai	797	717	678	752	827	917	1083
+ Übernahme Testergebnis aus Vorjahr	0	5	33	27	40	47	38
= Total Bewerbungen Mai	797	722	711	779	867	964	1121
... in % zu Kapazität	129	132	130	134	134	161	205
Erfolgreiche Absolvierung EMS	750	673	637	709	770	871	1026*
Gesamtbewerbungen Juli EMS / Ergebnis Vorjahr	750	678	670	736	805	907	1051*
Rückzugsquote zwischen Februar und Juli	26.5	19.8	20.6	24.4	25.5	22.9	22.7
Prozent Gesamtbewerbungen, die Studienplatz erhalten	82.7	80.5	81.4	79.2	80.5	65.9	51.9
NC-Index Bewerbungen/Studienplatz	1.21	1.24	1.23	1.25	1.24	1.52	1.92

Tabelle 1: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; 1998 bis 2004 für Humanmedizin. (* 16 EMS-Teilnehmer zählen nicht als Bewerbungen: 14 x von Zahn- in Humanmedizin, 2 x Studienbeginn in Neuenburg).

In Veterinärmedizin bleibt die Situation vergleichbar zu 2003, in Zahnmedizin war das Auslöskriterium für den NC erfüllt; die Situation hat sich durch Rückzüge deutlich entspannt.

	Veterinärmedizin						Zahn- medizin
	1999	2000	2001	2002	2003	2004	2004
Kapazitäten (am NC beteiligte Univ.)	150	150	165	185	170	150	142
Anmeldungen bei CRUS im Februar	231	213	217	254	271	273	208
... in % zu Kapazität	153	142	132	137	159	182	146
Anmeldungen EMS im Mai	194	176	179	206	225	207	163
+ Übernahme Testergebnis aus Vorjahr	0	5	5	3	11	13	1*
= Total Bewerbungen Mai	194	181	184	209	236	220	164
... in % zu Kapazität	129	121	112	113	139	147	115
Erfolgreiche Absolvierung EMS	182	164	171	203	217	196	151
Gesamtbewerbungen Juli EMS / Ergebnis Vorjahr	182	169	176	205	228	209	152
Rückzugsquote zwischen Februar und Juli	21.2	20.7	18.9	19.3	15.9	23.4	26.9
Prozent Gesamtbewerbungen, die Studienplatz erhalten	82.4	88.7	93.7	90.2	74.5	71.8	93.4
NC-Index Bewerbungen/Studienplatz	1.21	1.13	1.06	1.11	1.34	1.39	1.07

Tabelle 2: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; 1998 bis 2004 für Veterinär- und Zahnmedizin

*Bewerbung 2003 in Humanmedizin

Die Zahl der Anrechnungen des Testergebnisses aus dem Vorjahr bleibt stabil, dafür hat die Zahl der Wiederholungen der Testabnahme (siehe Seite 45) deutlich zugenommen. Da aufgrund der Anmeldezahlen mit einem verschärften NC zu rechnen war, wollten offenbar viele Personen die Chance auf eine Verbesserung wahrnehmen.

Nachfolgend sind die Anmeldungen und Rückzüge nach Disziplinen und Universitäten aufgeschlüsselt und grafisch im Verlauf dargestellt.

	Studienbeginn 2001			Studienbeginn 2002			Studienbeginn 2003			Studienbeginn 2004		
	Anmel- dungen am 15.2.	Pers. m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers. m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers. m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers.m. gültigem Testwert	Rück- züge in %
Basel HM	175	136 (7)	22.3	239	181 (6)	24.3	256	155 (13)	23.8	299	247 (7)	17.4
Bern HM	282	232 (5)	17.7	305	245 (16)	19.7	294	234 (13)	20.4	357	277 (5)	22.4
Freiburg HM	97	69 (1)	28.9	87	66 (5)	24.1	113	84 (4)	25.7	107	89 (8)	16.8
Zürich HM	420	299 (14)	28.8	449	313 (13)	30.3	518	394 (17)	23.9	597	450 (18)	24.6
Total	974	736	24.4	1080	805	25.5	1181	907	23.2	1360	1063	21.8
Bern VM	118	104 (4)	11.9	132	106 (2)	19.7	133	108 (4)	18.8	137	111 (6)	19.0
Zürich VM	99	72 (1)	27.3	122	99 (1)	18.9	141	120 (7)	14.9	136	98 (7)	27.9
Total	217	176	18.9	254	205	19.3	274	228	16.8	273	209	23.4
Basel ZM	kein NC bzw. kein EMS									45	30	33.3
Bern ZM										65	52 (1)	20.0
Freiburg ZM										7	6	14.3
Zürich ZM										91	63	30.8
Total										208	151	27.4

Tabelle 3: Anmeldeverlauf für Universitäten, Human- (HM), Veterinär- (VM) und Zahnmedizin (ZM). Personen mit gültigem Testwert; in Klammern: Übernahme aus dem Vorjahr

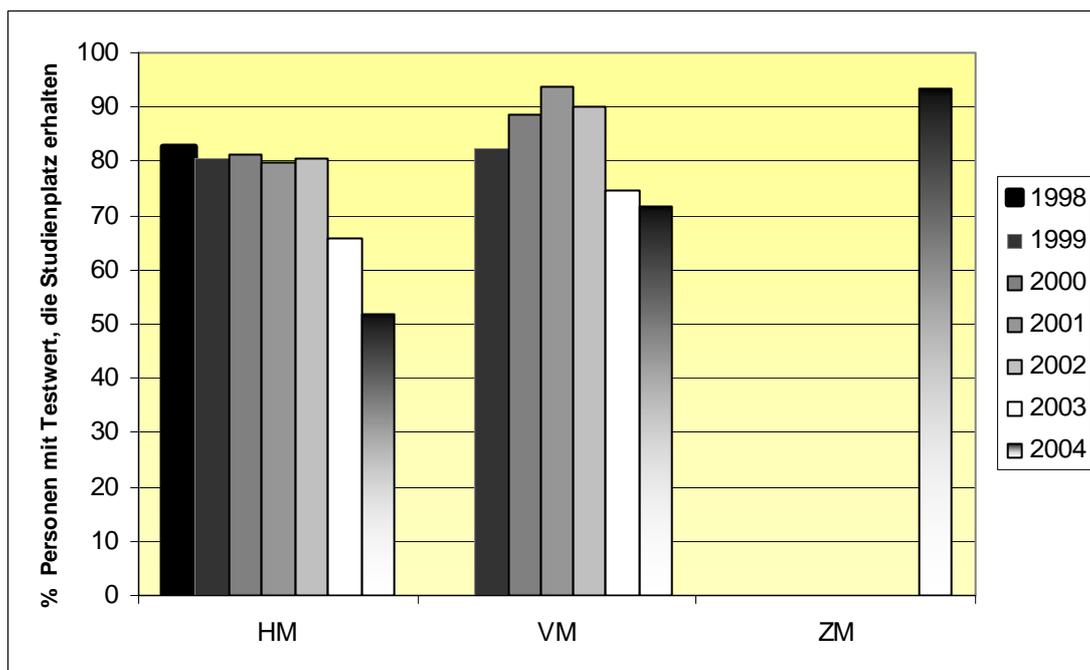


Abbildung 4: Prozent der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann – HM: Humanmedizin; VM: Veterinärmedizin; ZM: Zahnmedizin

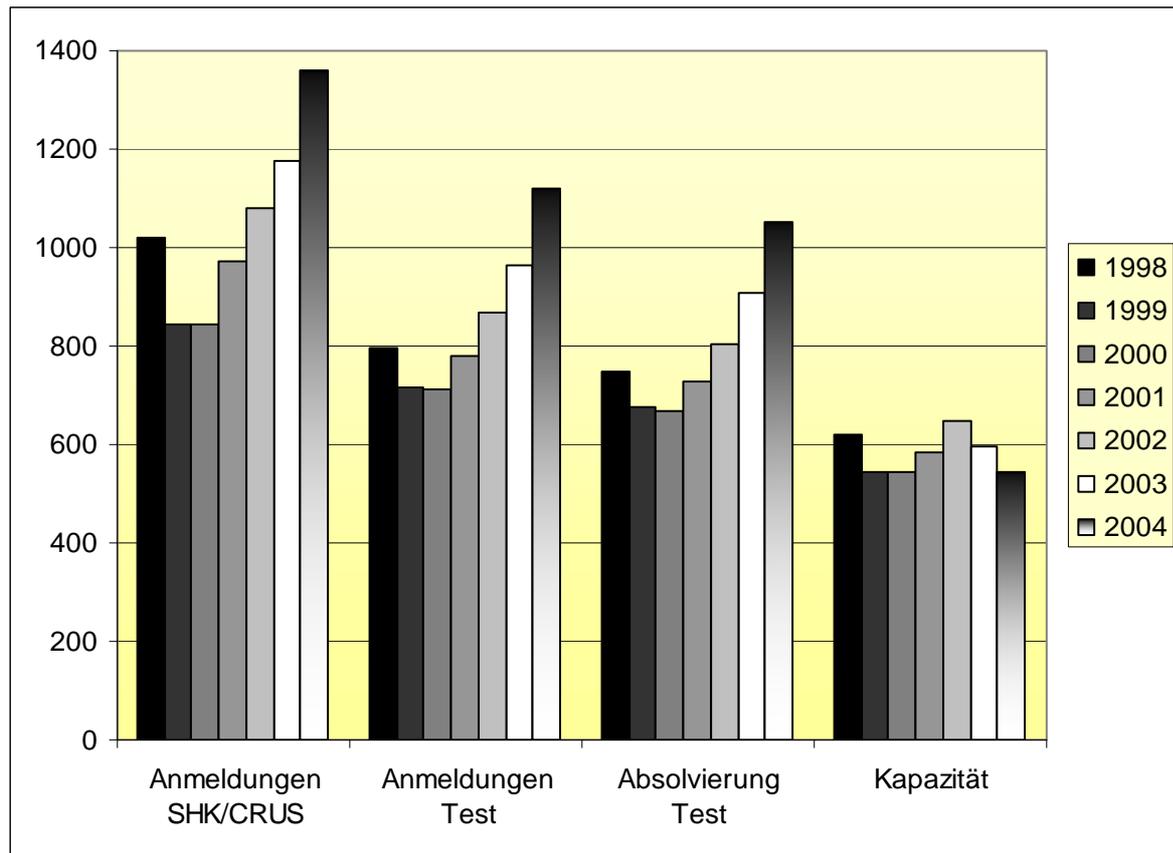


Abbildung 5: Anmeldeverlauf für Humanmedizin 1998 bis 2004 im Vergleich (mit NC)

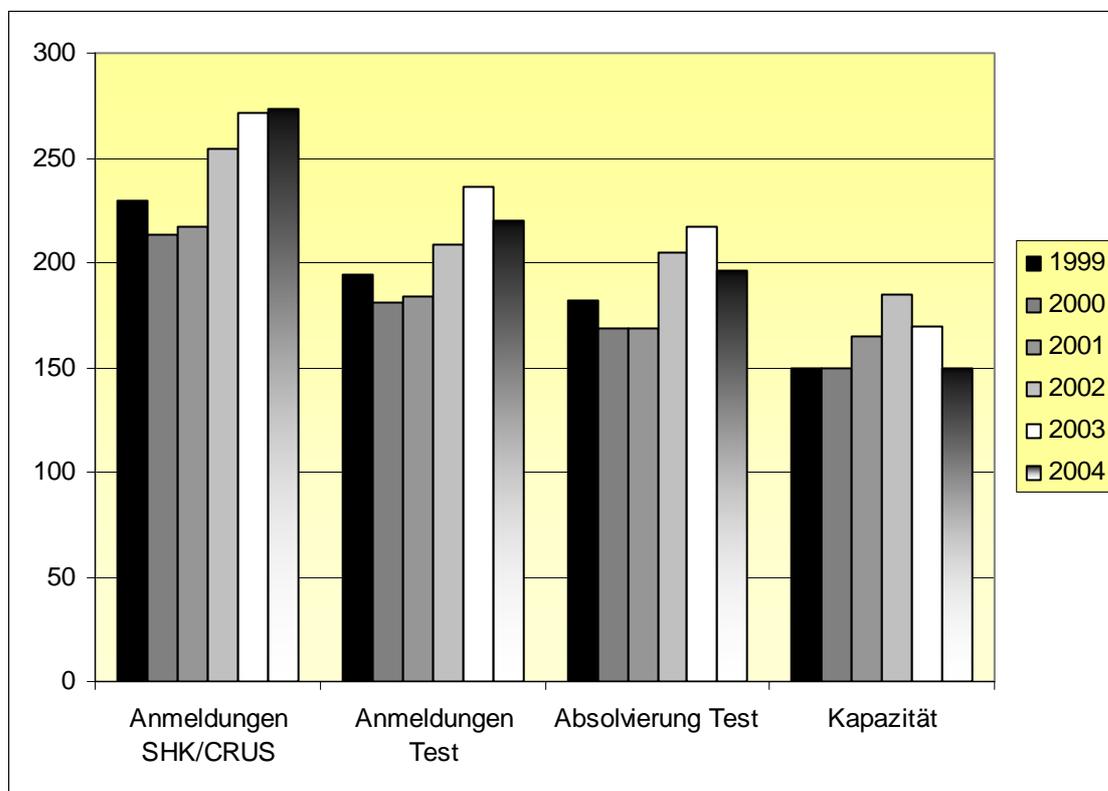


Abbildung 6: Anmeldeverlauf für Veterinärmedizin 1999 bis 2004 im Vergleich (mit NC)

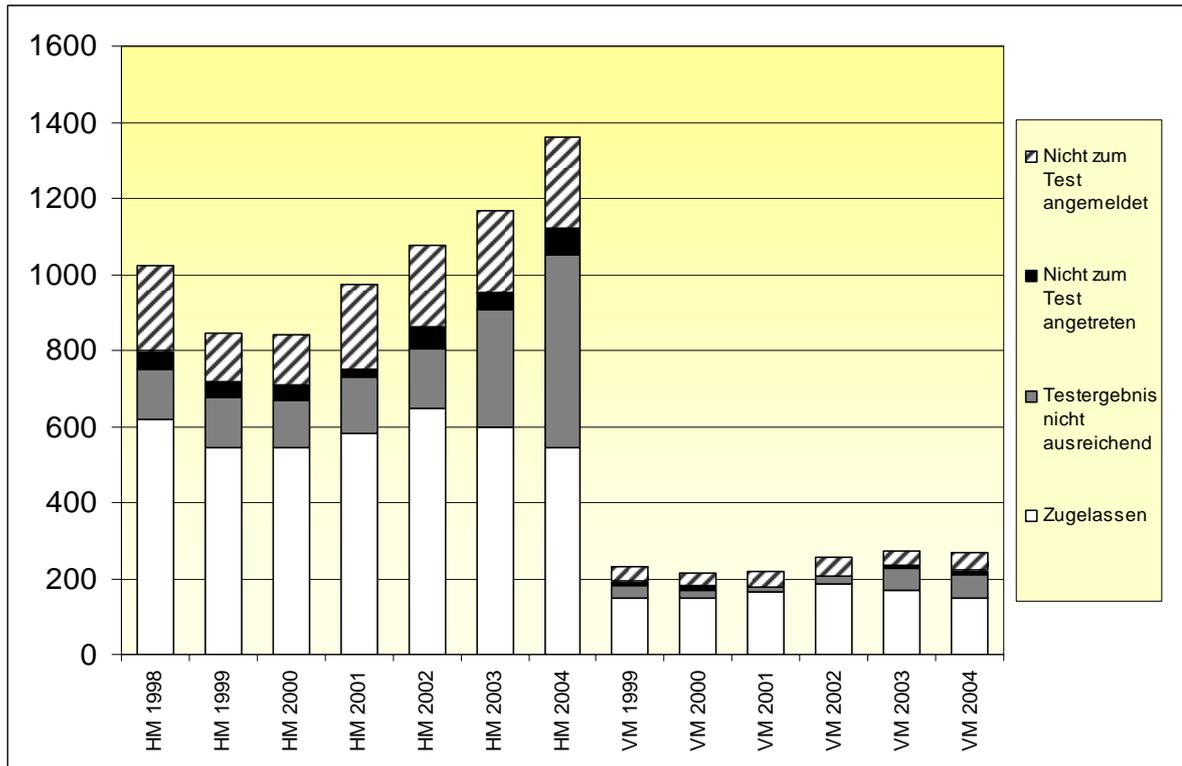


Abbildung 7: Kapazität, Anmeldungen und Testteilnahme für Human- (HM) und Veterinärmedizin (VM). Aufteilung der Voranmeldungen (Februar des jeweiligen Jahres)

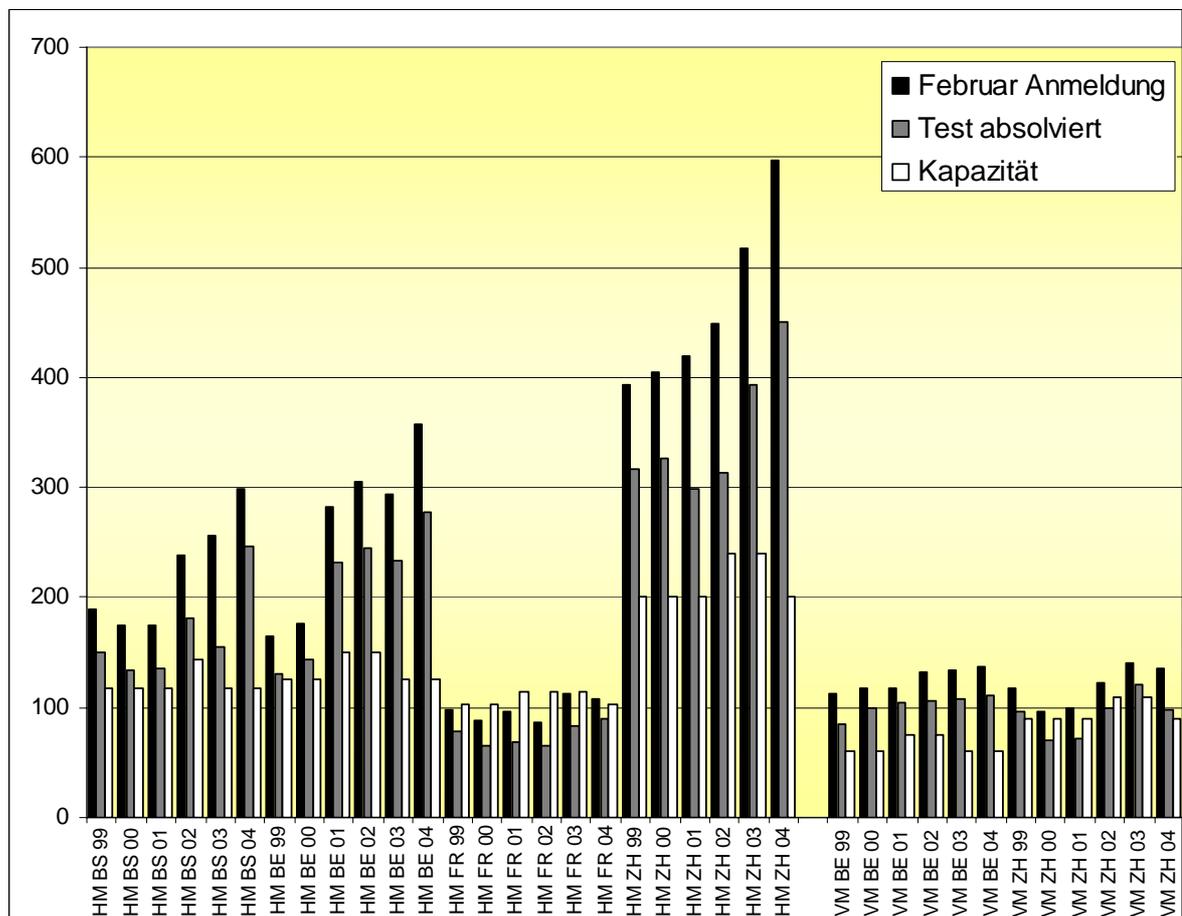


Abbildung 8: Kapazität, Anmeldungen und Testteilnahme für Fächer und Universitäten

5.2 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme in 15 Testlokalen in 7 Testorten. Die Grössen variierten im bisher üblichen Rahmen. Die Zahl der Betreuer war wie bisher proportional zur Teilnehmerzahl vorgegeben. Es gibt keine Hinweise seitens des Betreuungspersonals oder von Kandidaten, dass sich unterschiedliche Raumgrössen auf die Ergebnisse auswirken. In den grossen Räumen ist der zur Verfügung stehende Platz teilweise grosszügiger (wenn es sich um grössere Hörsäle handelt).

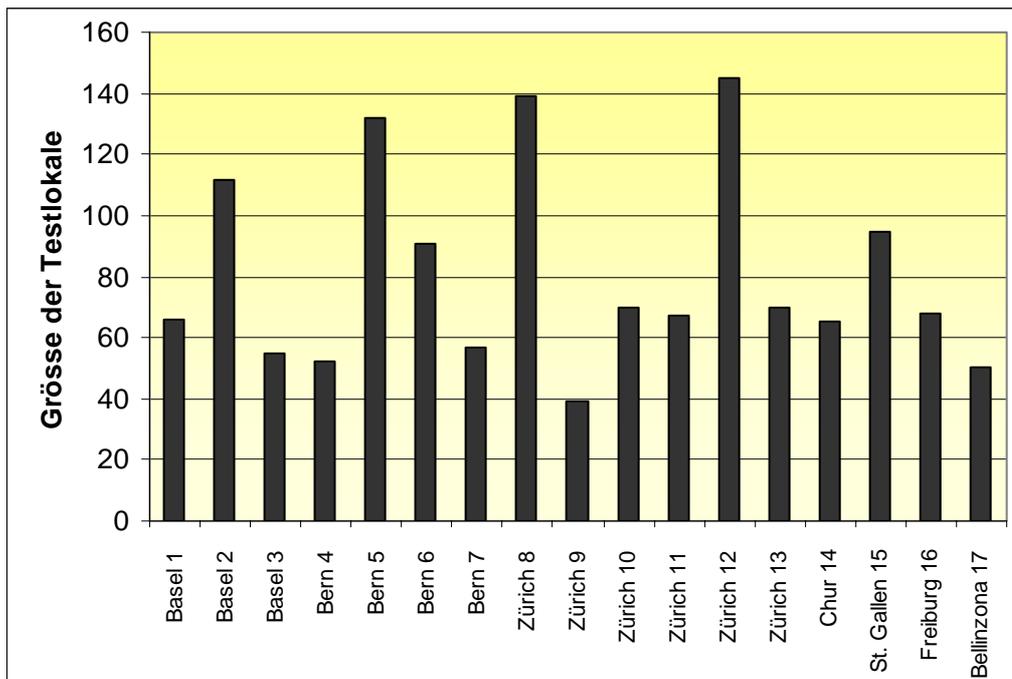


Abbildung 9: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2004

5.3 Testorte und Wunschuniversitäten

Die Wahl des Testortes hat keinen Einfluss auf den zukünftigen Studienort. Es wurde empfohlen, den Testort zu wählen, der dem Wohnort am nächsten liegt. Längere Anreisen könnten sich negativ auf die Leistungsfähigkeit auswirken.

Wie in den Vorjahren zeigen sich Übereinstimmungen zwischen Test- und Wunsch-Studienort, da auch die Wahl der Universität häufig nach der Wohnortnähe erfolgt. Bekanntlich kann auch eine Umleitung vermieden werden, wenn der Wohnkanton mit dem Kanton der gewünschten Universität identisch ist. Diesen Vorteil will man offenbar nutzen.

Testort	Wunschuniversität				Gesamt
	Bern	Basel	Freiburg	Zürich	
BASEL	22 5.1%	197 73.0%	4 4.6%	10 1.7%	233 17.0%
BELLINZONA	8 1.9%	5 1.9%	8 9.2%	29 4.9%	50 3.6%
BERN	299 69.5%	6 2.2%	22 25.3%	5 .9%	332 24.2%
CHUR	17 4.0%	9 3.3%	2 2.3%	37 6.3%	65 4.7%
FREIBURG	33 7.7%	1 .4%	31 35.6%	3 .5%	68 5.0%
ST. GALLEN	21 4.9%	13 4.8%	3 3.4%	58 9.9%	95 6.9%
ZÜRICH	30 7.0%	39 14.4%	17 19.5%	444 75.8%	530 38.6%
Gesamt	430	270	87	586	1373

Tabelle 4: Testorte und Wunschuniversität (erste Wahl) 2004

Abbildung 10 macht durch die „Streckung“ aller Universitäten auf 100% das Verhältnis der „lokalen“ Bewerber zu den übrigen vergleichbar.

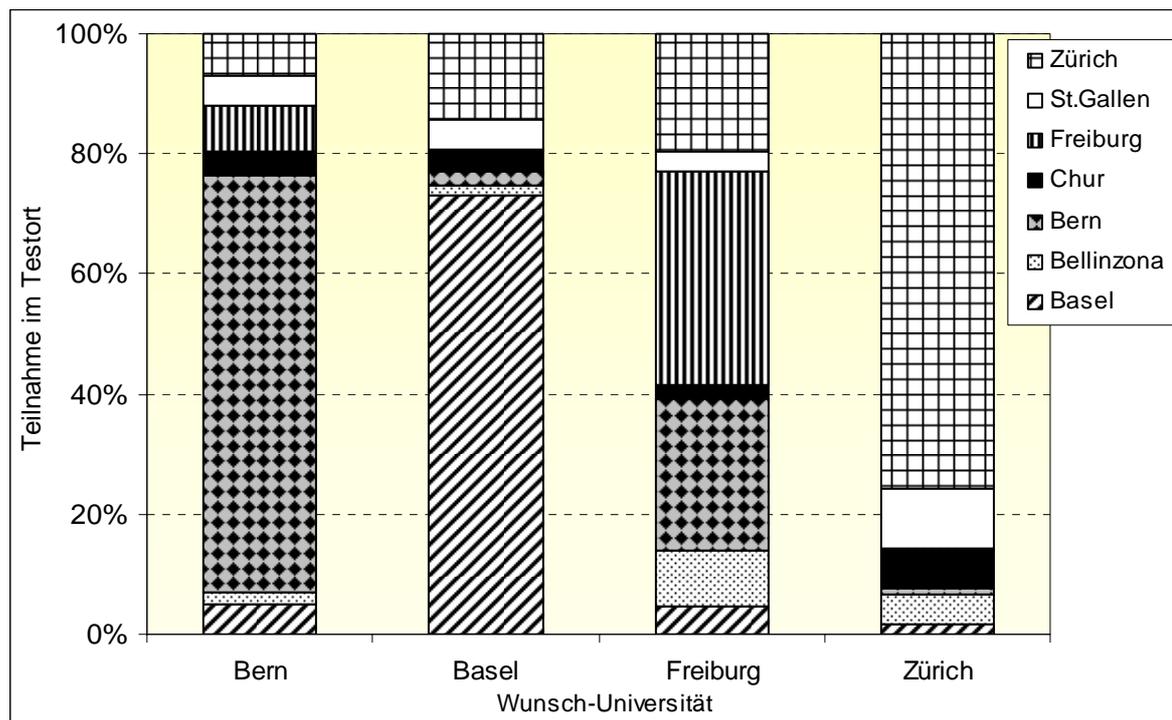


Abbildung 10: Aufteilung der Kandidaten nach Wunschuniversität auf die Testorte 2004

5.4 Wunschuniversitäten und Testorte nach Wohnkantonen

Nachfolgend sind die Testorte im Verhältnis zum Wohnkanton tabelliert. Angegeben sind auch die Vergleichswerte seit 1998.

Wohnkanton/-ort	Voranmeldung Humanmedizin 2004					Total bisheriger Jahre					
	Bern	Basel	Freiburg	Zürich	Total	2003	2002	2001	2000	1999	1998
AG	13	45	2	33	93	54	52	42	62	62	66
AI	2	1	0	1	4	0	2	1	3	0	2
AR		2	1	5	8	6	6	8	4	6	4
BE	167	7	4	1	179	135	139	148	88	87	101
BL	3	48	1	2	54	44	37	41	38	52	50
BS	1	46	0	1	48	41	35	31	26	37	40
FR	7	0	35	0	42	44	28	24	33	39	34
GE	0	0	0	0	0	1	1	1	0	0	1
GL	0	1	0	0	1	6	3	3	3	4	4
GR	11	10	1	19	41	46	18	30	16	22	19
JU	0	1	2	0	3	1	0	0	0	1	0
LU	19	12	6	16	53	52	62	42	29	34	43
NE	1	0	0	0	1	1	1	0	0	1	2
NW	0	3	0	2	5	9	4	3	2	3	6
OW	0	4	0	1	5	2	7	3	9	2	3
SG	9	10	3	33	55	48	50	41	51	51	64
SH	1	1	0	4	6	10	10	15	5	4	7
SO	17	20	3	2	42	38	35	21	13	22	26
SZ	2	1	0	15	18	11	11	15	16	16	21
TG	2	1	0	21	24	22	28	30	20	16	22
TI	6	4	7	17	34	30	18	31	21	23	30
UR	0	1	0	0	1	3	4	4	2	1	3
VD	0	0	1	0	1	2	1	2	2	0	3
VS	4	0	3	0	7	5	13	12	11	12	22
ZG	2	3	1	11	17	12	16	19	6	4	6
ZH	5	10	9	239	263	234	178	130	169	159	165
FL	0	1	0	4	5	3	3	5	4	2	3
Übriges Ausland	1	8	2	5	16	11	8	7	4	13	3
Total	273	240	81	432	1026	871	770	709	637	673	750

Tabelle 5: Wohnkanton und Voranmeldung Humanmedizin, Vergleich zu 1998 bis 2004. Das Fürstentum Liechtenstein (FL) wird als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wohnkanton/ Wohnort	Vor Anmeldung VM 2004			Total bisheriger Jahre				
	Bern	Zürich	Total	2003	2002	2001	2000	1999
AG	2	9	11	13	8	7	9	22
AI	1	0	1	0	3	0	1	0
AR	0	2	2	2	0	2	0	0
BE	29	0	29	32	30	39	26	26
BL	6	6	12	8	7	4	3	7
BS	2	2	4	2	4	7	2	2
FR	6	0	6	8	9	3	11	9
GE	2	0	2	4	6	9	11	13
GL	0	0	0	1	1	0	2	0
GR	1	9	10	13	4	6	3	7
JU	3	1	4	4	2	1	1	1
LU	10	1	11	7	13	5	12	10
NE	4	0	4	7	6	3	6	4
NW	0	1	1	2	2	0	0	0
SG	3	11	14	13	9	7	4	10
SH	0	1	1	2	0	3	3	0
SO	7	0	7	7	8	4	7	6
SZ	2	2	4	2	4	0	2	2
TG	1	4	5	11	6	8	3	4
TI	1	5	6	9	9	7	6	6
UR	0	2	2	1	1	1	0	0
VD	14	1	15	15	13	21	15	11
VS	7	0	7	4	7	7	7	3
ZG	0	4	4	6	3	4	3	6
ZH	0	29	29	42	46	19	25	32
FL	1	0	1	0	1	2	0	0
Übriges Ausland	3	1	4	2	1	2	2	1
Total	105	91	196	217	203	171	164	182

Tabelle 6: Gewünschte Studienorte Veterinärmedizin nach Wohnkantonen, Vergleich zu 1999 bis 2004. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wohnkanton / Wohnort	Vor Anmeldung Universität 2004 Zahnmedizin				
	Bern	Basel	Freiburg	Zürich	Total
AG	4	4	1	5	14
AR	0	0	0	4	4
BE	30	0	1	0	31
BL	0	10	0	0	10
BS	0	4	0	0	4
FR	0	0	1	0	1
GR	2	1	0	5	8
LU	1	1	0	0	2
NW	1	0	0	0	1
OW	1	1	0	0	2
SG	3	0	0	1	4
SH	1	1	0	1	3
SO	3	4	1	1	9
SZ	1	0	0	1	2
TG	1	0	0	5	6
TI	1	1	1	5	8
VS	3	1	0	0	4
ZG	0	0	0	1	1
ZH	0	1	1	31	33
FL	0	1	0	1	2
Übriges Ausland	0	0	0	2	2
Total	52	30	6	63	151

Tabelle 7: Gewünschte Studienorte Zahnmedizin nach Wohnkantonen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die nachfolgende Tabelle zeigt, welche Testorte gewählt werden. Im Jahr 2005 wird es, wenn ein NC notwendig ist, zusätzlich einen Testort im Aargau geben.

Wohnkanton/ -ort	Basel	Bellinzona	Bern	Chur	Freiburg	St.Gallen	Zürich	Total
AG	46	0	7	0	0	1	64	118
AI	0	0	0	0	0	5	0	5
AR	1	0	0	0	0	12	1	14
BE	6	0	228	0	3	0	2	239
BL	74	0	1	0	0	0	1	76
BS	54	0	1	1	0	0	0	56
FR	0	0	21	0	28	0	0	49
GE	0	0	0	0	2	0	0	2
GL	0	0	0	0	0	0	1	1
GR	0	1	1	52	0	1	4	59
JU	1	0	0	0	6	0	0	7
LU	6	0	22	0	0	0	38	66
NE	0	0	0	0	5	0	0	5
NW	2	0	2	0	0	0	3	7
OW	4	0	1	0	0	0	2	7
SG	1	0	0	6	0	59	7	73
SH	0	0	0	0	0	0	10	10
SO	21	0	33	0	0	0	4	58
SZ	0	0	2	0	0	1	21	24
TG	0	0	0	0	0	11	24	35
TI	0	47	0	0	0	0	1	48
UR	1	0	0	0	0	1	1	3
VD	0	0	0	0	16	0	0	16
VS	0	0	9	0	6	1	2	18
ZG	0	0	0	0	0	1	21	22
ZH	5	2	3	0	1	0	314	325
FL	0	0	0	6	0	1	1	8
Ausl.	11	0	1	0	1	1	8	22
Gesamt	233	50	332	65	68	95	530	1373

Tabelle 8: Testorte nach Wohnkantonen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

5.5 Testabsolvierung nach Alter und Geschlecht

Disziplin Humanmedizin	Geschlecht				Total
	männlich		weiblich		
	Anzahl	%	Anzahl	%	
Geburtsjahre 1951-1980	55	55.6%	44	44.4%	99
1981-1983	96	41.2%	137	58.8%	233
1984-1988	202	29.1%	492	70.9%	694
Total	353	34.4%	673	65.6%	1026
Disziplin Veterinärmedizin	Geschlecht				Total
	männlich		weiblich		
	Anzahl	%	Anzahl	%	
Geburtsjahre 1951-1980	4	14.8%	23	85.2%	27
1981-1983	10	22.2%	35	77.8%	45
1984-1988	10	8.1%	114	91.9%	124
Total	24	12.2%	172	87.8%	196
Disziplin Zahnmedizin	Geschlecht				Total
	männlich		weiblich		
	Anzahl	%	Anzahl	%	
Geburtsjahre 1951-1980	12	57.1%	9	42.9%	21
1981-1983	12	37.5%	20	62.5%	32
1984-1988	37	37.8%	61	62.2%	98
Total	61	40.4%	90	59.6%	151
Alle Bewerbungen	Geschlecht				Total
	männlich		weiblich		
	Anzahl	%	Anzahl	%	
Geburtsjahre 1951-1980	71	48.3%	76	51.7%	147
1981-1983	118	38.1%	192	61.9%	310
1984-1988	249	27.2%	667	72.8%	916
Total	438	31.9%	935	68.1%	1373

Tabelle 9: Testabsolventen nach Geburtsjahr und Geschlecht 2004 und im Vergleich zum ersten Jahr des NC pro Disziplin

Die Zahl der weiblichen Kandidatinnen im Fach Humanmedizin hat 2004 weiter zugenommen – die Zahl der männlichen Bewerber bleibt im Verhältnis dazu eher konstant. Vor allem die Zahl der Frauen mit einer Bewerbung unmittelbar nach der Mittelschule ist für den Anstieg verantwortlich (Abbildung 12). In Abbildung 13 wird deutlich, dass die grössten Zunahmen in den jüngeren Geburtsjahren liegen, tendenzielle Anstiege (d.h. die höchsten Werte) aber in allen Altersgruppen vorhanden sind. Die zunehmende Nachfrage ist letztendlich in allen Gruppen zu finden.

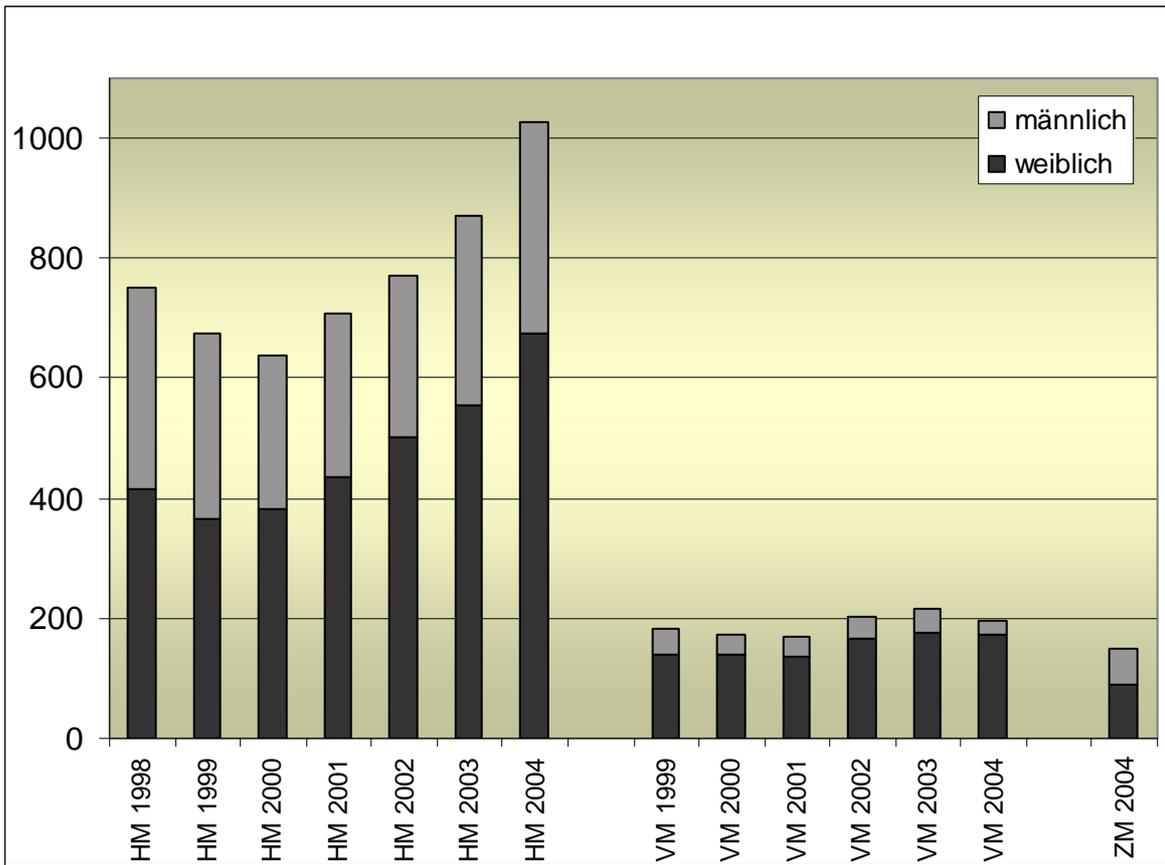


Abbildung 11: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM) 1998 bis 2004, Veterinärmedizin (VM) 1999 bis 2004 und Zahnmedizin (ZM) 2004

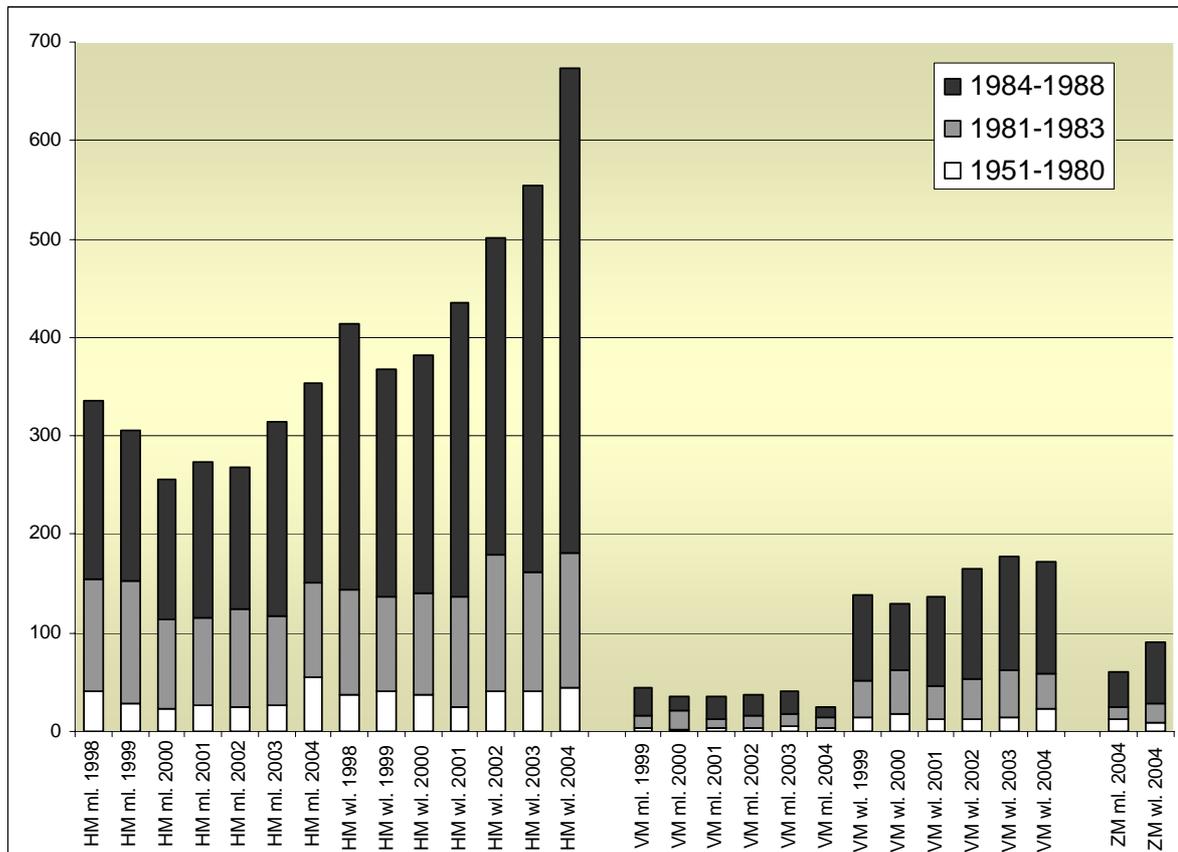


Abbildung 12: Zusammensetzung der Bewerbergruppen nach Geschlecht mi: männlich wi: weiblich und Alter (Geburtsjahre für Kohorte 2004 – 1998 bis 2003 jeweils entsprechend versetzt)

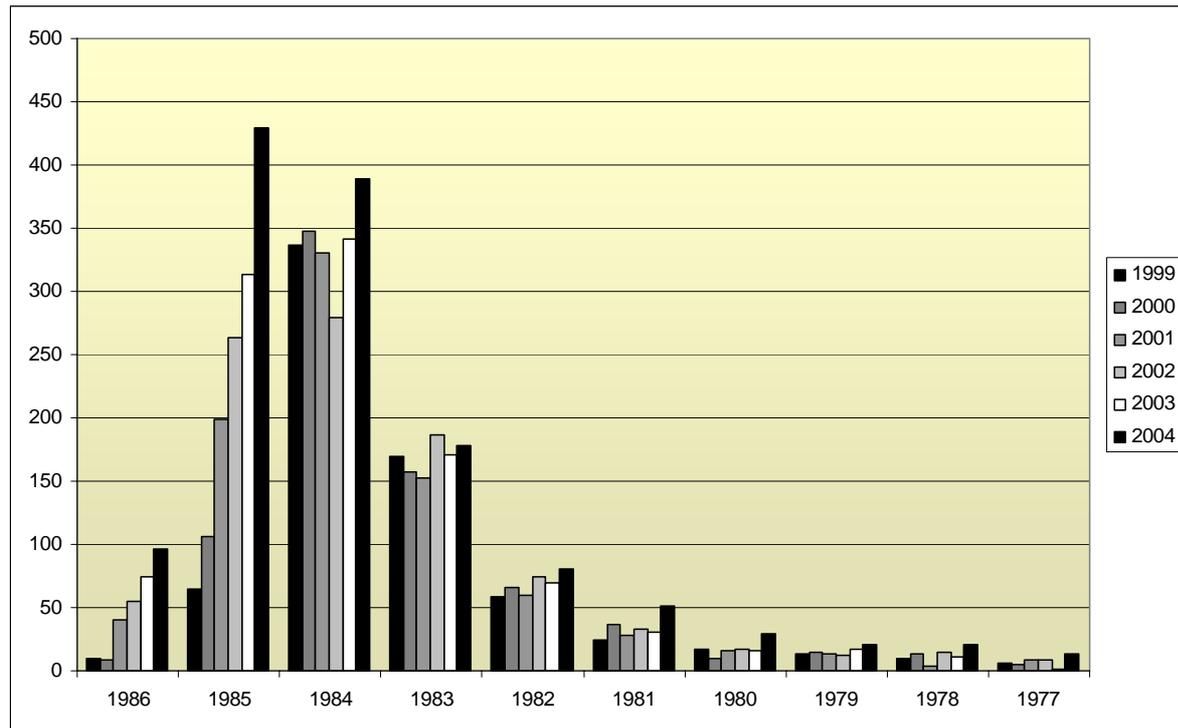


Abbildung 13: Bewerbungen nach Geburtsjahren für 2004 und Entsprechungen für die Jahre 1999 bis 2003 (jeweils um entsprechende Jahre versetzt)

5.6 Übernahme des Testergebnisses aus Vorjahren

Disziplin	Jahr	N	Minimum	Maximum	Mittelwert	Stand.- abw.	Zugelassen 1. Runde
Human- medizin	2000	33	85	119	102.85	7.25	33
	2001	27	94	119	102.96	7.28	27
	2002	40	83	118	104.03	9.16	40
	2003	47	91	121	103.70	8.27	44 (3 nicht)
	2004	37	90	119	105.9	7.7	31 (6 nicht)
Vete- rinär- medizin	2000	5	97	107	102.00	3.81	5
	2001	5	89	112	102.80	9.52	5
	2002	3	94	109	101.33	7.51	3
	2003	11	86	118	102.9	9.0	10 (1 nicht)
	2004	13	80	106	99.1	7.1	12 (1 nicht)
Zahn- medizin	2004	1			93.0		1

Tabelle 10: Testwert für Personen mit Übernahme des Testergebnisses aus dem Vorjahr

Die Zahl der Übernahmen wächst nicht – vielleicht, weil die Grenzwerte für die Zulassung seit 3 Jahren steigen und man lieber noch einmal antritt. 7 Personen werden aufgrund des

Testwertes des Vorjahres nicht zugelassen, weil der Grenzwert im Jahr 2004 höher liegt als 2003.

5.7 Erneute Testteilnahme 2004

Erste Testteilnahme	Zweite Testteilnahme					
	1999	2000	2001	2002	2003	2004
1998	12	2	1	1	1	1
1999		19	3	0	0	2
2000			12	0	3	1
2001				11	3	1
2002					32	13
2003						78
Gesamt	12	21	16	12	39	96

Tabelle 11: Wiederholte Testteilnahme: Anzahl der Personen

Es ist möglich, erneut zum Test anzutreten und dann das neuere Ergebnis für die Zulassung zu verwenden. Personen, die sich später als ein Jahr nach der ersten Testteilnahme erneut bewerben, müssen in jedem Falle noch einmal am Test teilnehmen. Vom Testergebnis hängt auch ab, ob ein Studium an der Wunschuniversität aufgenommen werden kann oder ggf. Umleitungen erfolgen. Es treten auch Personen erneut an, um ihr Ergebnis zu verbessern.

100 Personen hatten bis 2003 zweimal am Test teilgenommen. Im Jahr 2004 kamen 96 Personen, also noch einmal fast genauso viele, dazu. Mit 78 Testwiederholungen im Folgejahr liegt die Zahl deutlich höher als 2003 – diese Zahl ist offenbar auch zugunsten der Übernahme des gleichen Testergebnisses im Folgejahr gewachsen.

	N	Minimum	Maximum	Mittelwert	Stand- abweichung
Testwert 1. Teilnahme	196	72	112	87.9	7.5
Testwert 2. Teilnahme	196	72	127	96.8	9.6
Veränderungen	196	-16	29	8.8	7.4

Tabelle 12: Testwertvergleich zwischen erster und zweiter Testteilnahme

Die Korrelation zwischen den ersten und zweiten Testwerten beträgt 0.64. Die Wahrscheinlichkeit, sich bei der zweiten Testteilnahme zu verbessern beträgt 88.8%. Nur 9.7% verschlechtern sich, 1.5% erreichen genau das gleiche Ergebnis.

Die Tatsache, dass bei wiederholter Testung bessere Ergebnisse auftreten, spricht für die Empfehlung im Rahmen der Vorbereitung, einmal einen Probelauf unter „ernstfallnahen“ Bedingungen durchzuführen. Ein solcher wird in einigen Kantonen unter geeigneten Bedingungen angeboten. Es wäre denkbar, dass diejenigen Personen bevorzugt ein zweites Mal zum Test antreten, die sich auf die erste Teilnahme nicht ausreichend vorbereitet haben. Einige wenige Verschlechterungen sind in diesem Jahr allerdings auch feststellbar – im Vorjahr waren es lediglich 4 Personen, jetzt sind es 15 Personen.

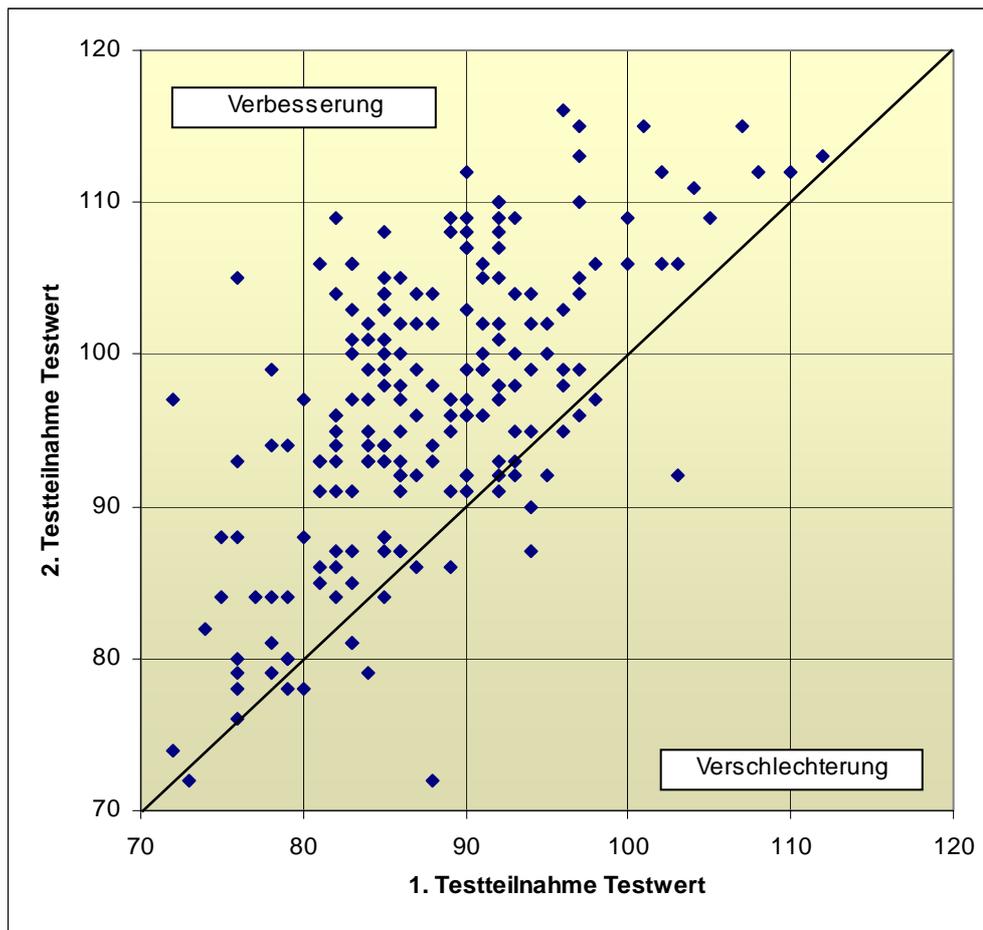


Abbildung 14: Testwerte für Personen, die zweimal am Test teilnahmen (1998 bis 2004)

5.8 Sprachgruppen

Nach wie vor absolviert die Mehrzahl der Personen den Test in deutscher Sprache. Der Prozentanteil von Personen mit einer Testabsolvierung in französischer Sprache nimmt eher etwas ab, während der Anteil italienischsprachiger Personen konstant bleibt.

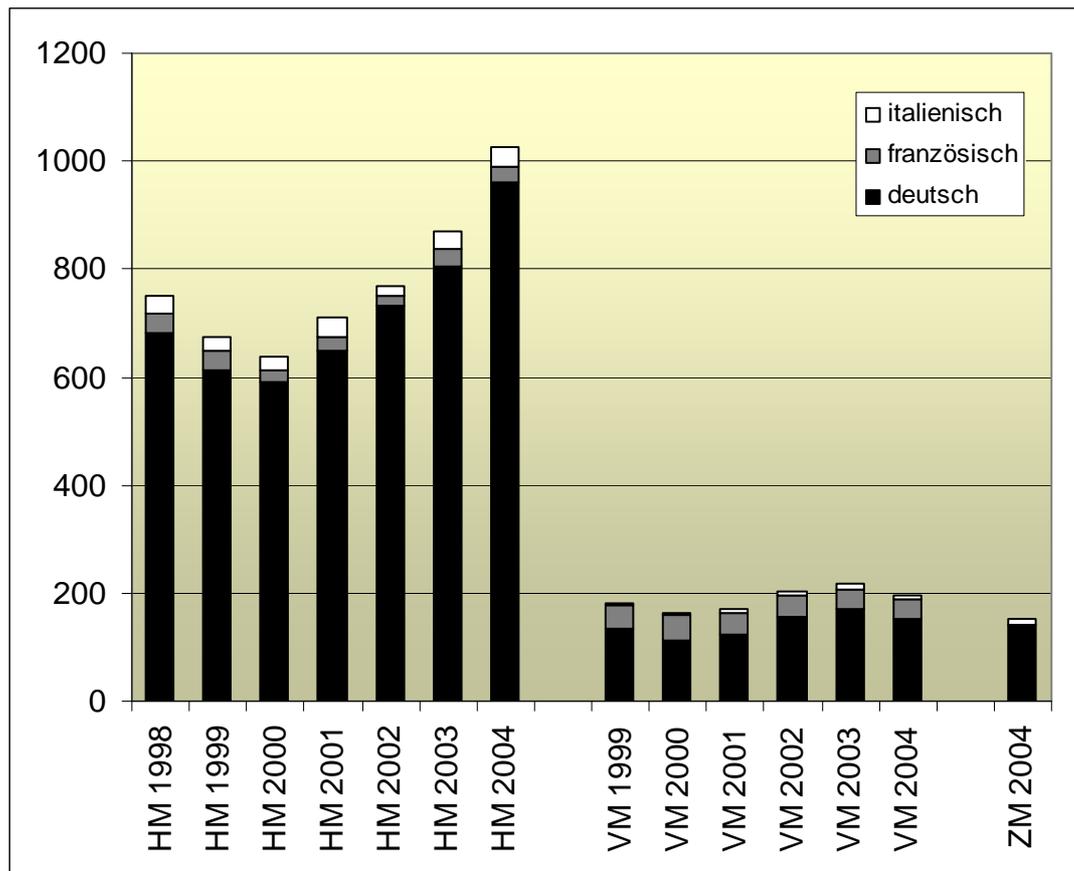


Abbildung 14: Vergleich der Sprachgruppen 1998 bis 2004, Absolutzahlen

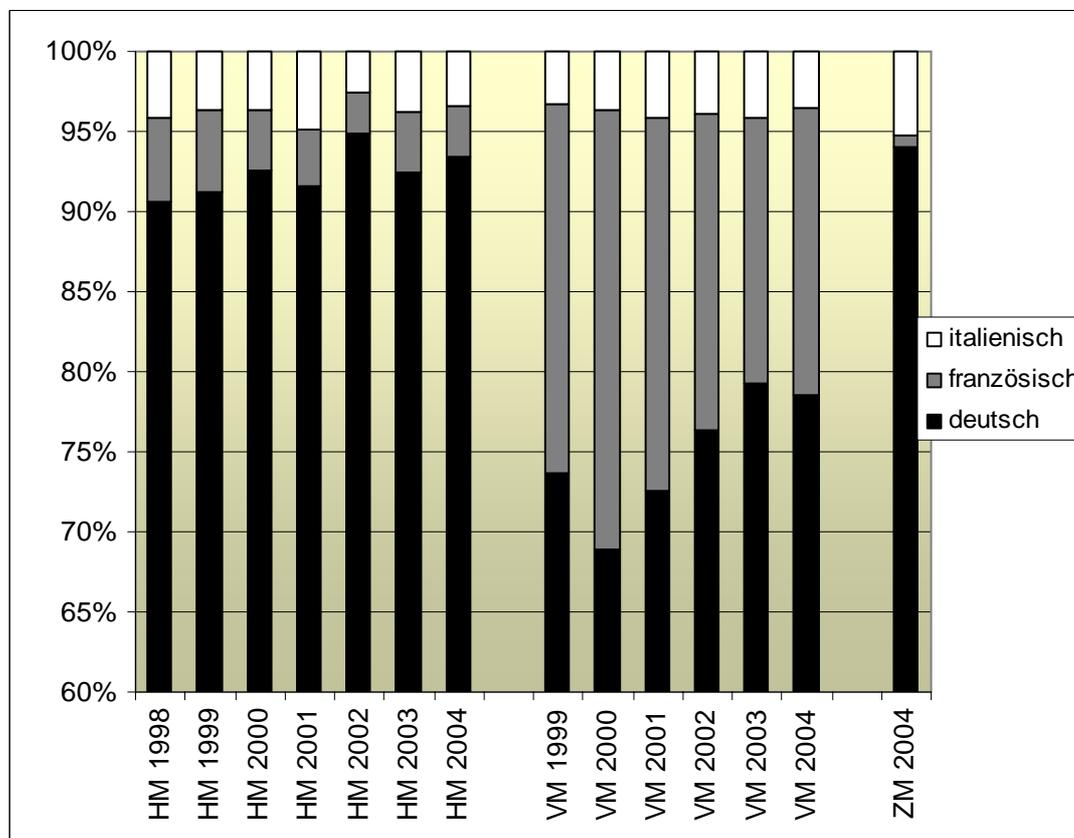


Abbildung 15: Vergleich der Sprachgruppen 1998 bis 2004, Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%)

			Geburtsjahre			Total
			1951-1980	1981-1983	1984-1988	
Humanmedizin	deutsch	männlich	51	94	192	337
		weiblich	41	132	449	622
		Total	92	226	641	959
	französisch	männlich		1	6	7
		weiblich		3	22	25
		Total		4	28	32
	italienisch	männlich	4	1	4	9
		weiblich	3	2	21	26
		Total	7	3	25	35
Veterinärmedizin	deutsch	männlich	3	7	7	17
		weiblich	18	29	90	137
		Total	21	36	97	154
	französisch	männlich	1	3	3	7
		weiblich	3	5	20	28
		Total	4	8	23	35
	italienisch	weiblich	2	1	4	7
		Total	2	1	4	7
	Zahnmedizin	deutsch	männlich	11	9	35
weiblich			8	20	59	87
Total			19	29	94	142
französisch		männlich			1	1
		Total			1	1
italienisch		männlich	1	3	1	5
		weiblich	1		2	3
		Total	2	3	3	8

Tabelle 13: Sprachgruppen 2004 nach Alter und Geschlecht aufgeschlüsselt

5.9 Alter und Maturitätsjahr

Die Mehrzahl der Personen hat sich gleich nach dem Mittelschulabschluss bzw. im darauf folgenden Jahr zum Medizinstudium beworben.

Die Personen werden später analog der Vorjahre zu Vergleichszwecken in drei Altersgruppen aufgeteilt. Die älteste Gruppe wird wiederum danach unterschieden, ob die Maturitätsprüfung unmittelbar nach der Mittelschule absolviert wurde oder noch eine Wartezeit dazwischen bestand. Bei drei Personen fehlt eine Angabe zum Maturitätsjahr, diese Personen werden hier nicht berücksichtigt.

G J	Maturitätsjahr																						T			
	74	76	77	81	82	83	84	85	86	87	91	92	93	95	96	97	98	99	00	01	02	03		04		
51	1	.	1		
55	1	1		
57	.	2	1	3		
58	.	.	1	1		
59	1	1	2		
60	1	.	1		
61	.	.	.	1	1		
63	1	1		
64	1	1	2		
65	1	1	2		
66	1	1		
67	1	2	1	1	1	6	
68	1	1	.	.	2	
69	1	1	
70	1	1	.	2	
71	1	.	.	1	1	3	
72	1	1	1	3	
73	1	2	3	
74	2	1	.	.	1	2	6		
75	3	2	1	.	1	1	.	.	.	5	13		
76	2	.	.	1	1	.	3	7	
77	1	1	1	4	6	13	
78	1	4	5	1	1	.	.	2	1	6	21	
79	5	1	1	.	.	4	10	21	
80	3	5	3	.	.	3	3	12	29
81	2	4	13	10	6	16	50	
82	2	16	34	15	13	81	
83	1	34	95	48	177	
84	13	181	4	389	
85	2	62	365	429	
86	4	92	96	
88	2	2	
T	1	2	1	1	2	1	1	1	2	5	2	2	0	8	4	6	7	14	14	34	101	380	781	1370		

Tabelle 14: Geburtsjahr (GJ) und Maturitätsjahr. Dicke Linien trennen die Altersgruppen (und bei der ältesten Gruppe zusätzlich eine Teilung nach dem Maturitätszeitpunkt), die in den nachfolgenden Statistiken unterschieden werden.

6 Beschreibung des verwendeten Eignungstests

6.1 Aufbau des Tests

Der EMS besteht aus neun verschiedenen Untertests, die zu einem Gesamtwert verrechnet werden. Die Tabelle 15 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Untertests. Aufgabenbeispiele finden Sie im Anhang ab Seite 116.

Bezeichnung der Untertests	Geprüfte Fähigkeiten	Zahl der Aufgaben	Bearbeitungszeit (in Minuten)
Muster zuordnen	Differenzierte visuelle Wahrnehmung	24 (20)*	22
Medizinisch-naturwissenschaftliches Grundverständnis	Verständnis für medizinisch-naturwissenschaftliche Problemstellungen	24 (20)*	60
Schlauchfiguren	Räumliches Vorstellungsvermögen	24 (20)*	15
Quantitative und formale Probleme	Quantitatives Problemlösen in medizinisch-naturwissenschaftlichen Kontexten	24 (20)*	60
Konzentriertes und Sorgfältiges Arbeiten	Konzentrationsfähigkeit, Aufmerksamkeit	1600 Zeichen 20 Punkte	8
MITTAGSPAUSE 60 Minuten			
Lernphase zu den Gedächtnistests: Figuren lernen Fakten lernen			4 6
Textverständnis	Verständnis und Interpretation medizinischer und naturwissensch. Texte	24 (18)*	60
Reproduktionsphase Gedächtnistests: Figuren lernen Fakten lernen	Behalten von figuralem Material Behalten von verbalem Material	20 20	5 7
Diagramme und Tabellen	Interpretation von Diagrammen und Tabellen	24 (20)*	60
Gesamttest		204 (178)*	5 Std. 7 Min.

Tabelle 15: Struktur und Ablauf des EMS *Aufgabenzahl und in Klammern die Zahl der gewerteten Aufgaben pro Untertest und für den gesamten Test, da zusätzliche (nicht gewertete) Einstreuaufgaben verwendet worden sind, Erläuterungen siehe Text

Beim EMS handelt es sich um eine Adaptation des deutschen Tests für Medizinische Studiengänge (TMS). Die Struktur des TMS hat sich in Deutschland im Rahmen von 12 Testeinsätzen bewährt (1986 zwei Testtermine und von 1987 bis 1996 je ein Termin pro Jahr). Beim EMS der Schweiz sind Struktur und Testlänge mit denen des deutschen TMS identisch, um dort langjährig erprobte Abläufe und damit auch die Güte-Kennwerte nicht zu verändern.

Die Vorteile des EMS lassen sich folgendermassen zusammenfassen:

- Auswahl der Testanforderungen aus einer grösseren Menge möglicher Studienanforderungen aufgrund einer Anforderungsanalyse;
- Wissenschaftliche Überprüfung der Prognosekraft dieser Anforderungen für den Studienerfolg, was die Zulassung nach der Eignung zum Studium erlaubt;
- Konstruktion der Aufgaben durch Experten UND anschliessende empirische Überprüfung, damit die Aufgaben alle Gütekriterien erfüllen, die gewünschte Fähigkeit tatsächlich messen und optimal „schwer“ sind;
- Beachtung, dass für die Beantwortung der Aufgaben kein spezielles fachliches Vorwissen notwendig ist, sondern tatsächlich die „Studierfähigkeit“ als aktuelle Fähigkeit zur Wissensaneignung und Problemlösung gemessen wird. Dadurch ist der Test auch wenig trainierbar, was sich positiv auf die Sozialverträglichkeit auswirkt (kein zusätzlicher Gewinn durch zusätzliche Trainingskurse nachgewiesen, wenn die empfohlene Vorbereitung erfolgt¹);
- Ökonomisches und genau kapazitätsentsprechendes Zulassungsverfahren möglich, die Rechtsfähigkeit des Verfahrens hat sich bei mehreren Überprüfungen (auch gerichtlich) bestätigt.

Wie wurden die Aufgaben entwickelt? Ausgangspunkt war eine differenzierte Anforderungsanalyse des Medizinstudiums, an der zahlreiche Lehrbeauftragte und Experten mitarbeiteten. Das Ziel war die Vorhersage des Studienerfolges, um ein faires und wissenschaftlich begründetes Zulassungsverfahren zu erhalten. Gemäss diesen Anforderungen wurden einzelne Aufgabengruppen (Untertests) konstruiert. Neun davon erfüllten alle notwendigen Anforderungen. Das Resultat ist also bereits eine Auswahl bewährter Aufgabentypen aus mehreren möglichen Alternativen. Jedes Jahr wurden neue Aufgaben für die Untertests entwickelt und in mehreren Schritten überarbeitet. An dieser Aufgabentwicklung nahmen zahlreiche Lehrbeauftragte und Experten teil. Die Aufgaben müssen sehr hohe Qualitätsstandards erfüllen, u.a. (1) müssen sie jedes Jahr die Studieneignung gleich zuverlässig messen, (2) muss das Schwierigkeitsspektrum aller Aufgaben annähernd vergleichbar sein, (3) darf kein spezielles Fachwissen vorausgesetzt werden, um die Trainierbarkeit des Tests gering zu halten, (4) muss eine eindeutige richtige Lösung existieren.

Die Erprobung neuer Aufgaben für sechs der neun Untertests (siehe Tabelle 15) erfolgte in Deutschland im Rahmen sogenannter „**Einstreuaufgaben**“. Nur bei ausreichender Bewährung wurden solche Aufgaben in nachfolgenden Testversionen für die Werteberechnung verwendet. Vier neue Aufgaben pro Untertest wurden in jeder Testform probeweise mitbearbeitet, ihr Ergebnis wurde jedoch nicht gezählt. Da in Deutschland acht verschiedene Testformen bei jedem Durchführungstermin zusammengestellt worden sind, konnten jeweils 32 neue Aufgaben pro Untertest an ausreichend grossen Stichproben erprobt werden. Maximal 20 davon wurden pro Jahr gebraucht. Dieser Überschuss war nach Aussagen der Entwickler auch notwendig, da nicht alle Aufgaben die Kriterien zufriedenstellend erfüllten. Es versteht sich, dass die Einstreuaufgaben nicht besonders gekennzeichnet worden sind – jede Aufgabe des Tests konnte eine solche sein.

¹ Zum Trainingsaspekt siehe Bericht 6 des ZTD, sowie die Argumentationen unter www.unifr.ch/ztd/ems

6.2 Berechnung der Werte

Alle Untertests, ausser dem „Konzentrierten und sorgfältigen Arbeiten“, liefern eine Summe („Punkte“) richtig gelöster Aufgaben zwischen 0 und 20 bzw. 18 bei „Textverständnis“. Summiert werden die gewerteten Aufgaben, nicht die Einstreuaufgaben.

Beim Test „Konzentriertes und Sorgfältiges Arbeiten“ müssen insgesamt 1600 Zeichen der Reihe nach bearbeitet werden – 400 davon sind anzustreichen. Es können in der zur Verfügung stehenden Zeit in der Regel nicht alle Zeichen bearbeitet werden. Die Position des letzten angestrichenen Zeichens bestimmt, wie viele Zeichen als bearbeitet gewertet werden. Alle übersehenen und fälschlich angestrichenen Zeichen **vor** diesem letzten bearbeiteten Zeichen zählen als Fehler und diese werden von der Menge der insgesamt angestrichenen Zeichen abgezogen. Die verbleibende Menge sind die „Richtigen“, die dann in eine Skala zwischen 0 und 20 transformiert werden, um mit den anderen Tests gleichgewichtig zum Punktwert addiert zu werden. Vierhundert Richtige wären das Maximum und entsprechen 20 Punkten.

Alle Punkte der Untertests werden zu einer Summe addiert (Punktwert, vgl. Abbildung 16). Dieser Wert hat den Nachteil, dass er nicht zwischen Tests verschiedener Geburtsjahre vergleichbar ist. Auch die Eichung verschiedener äquivalenter Sprachformen ist noch notwendig (s.u.). Deshalb findet eine Standardisierung auf den Mittelwert und die Standardabweichung der jeweiligen Testform statt. Dieser Testwert liegt zwischen 70 und 130 (der Mittelwert ist 100) und kann in einen Prozentrangwert umgerechnet werden. Prozentränge lassen sich am einfachsten veranschaulichen indem angegeben wird: x Prozent aller Teilnehmenden haben einen besseren oder schlechteren Testwert erreicht als die entsprechende Person.

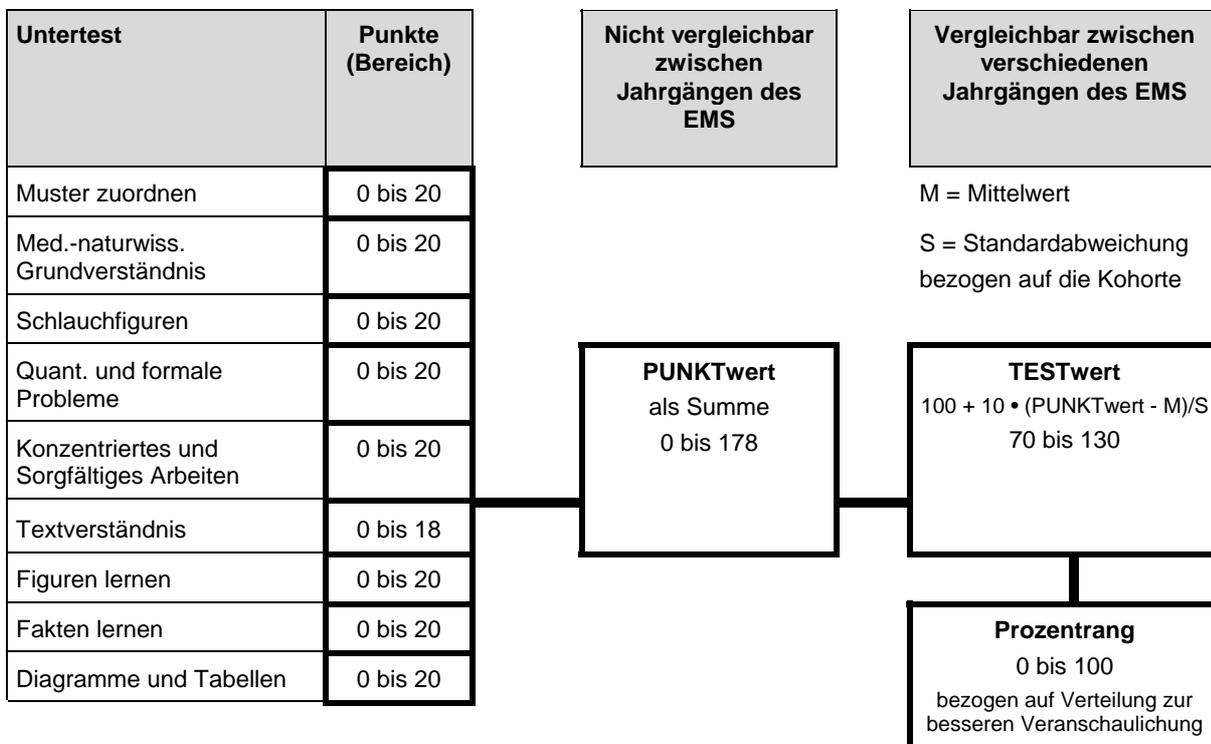


Abbildung 16: Punktwerte der einzelnen Untertests und ihre Zusammenführung über den Punktwert zum Testwert

6.3 Mittlerer Rangplatz der Untertests

Ziel des EMS ist es, eine kapazitätsentsprechende Vergabe der Studienplätze zu erreichen. Alle Plätze sollen vergeben werden. Es kann vorkommen, dass viele Kandidaten den niedrigsten für die Zulassung zu berücksichtigenden Testwert aufweisen – aber noch deutlich weniger Plätze zur Verfügung stehen. Da der Testwert normalverteilt ist, wird sich diese mangelnde Diskriminationsfähigkeit verstärken, je näher der kritische Testwert am Mittelwert 100 liegt, weil sich dort die maximalen Häufigkeiten pro Testwert befinden.

Vom Gesetz gefordert ist ein Parameter, welcher die Studieneignung vorhersagt. Dies erfüllt der bisher verwendete Testwert nachweislich. Beim Testwert werden die Punktwerte der einzelnen Untertests addiert und der Gesamtwert wird durch die Standardisierung auf $m = 100$ und $s = 10$ zwischen den Jahren vergleichbar. Das Gewicht der Untertests in diesem Gesamtwert wird letztendlich von der Varianz in diesen Untertests beeinflusst. Da die für die Studieneignung prognoserelevantesten Untertests zugleich die höchste Streuung aufweisen, erhöht dieser Effekt auch die Prognosekraft des Testwertes.

Wie die Tabelle A8 des Anhangs zeigt, haben bis zu 54 Personen im mittleren Testwertbereich den gleichen Testwert (Humanmedizin). Läge die Zahl der freien Plätze nun deutlich darunter, müsste eine Über- oder Unterschreitung der Kapazität in Kauf genommen werden. Beides hätte deutliche Nachteile.

Ein alternativer Parameter wäre der **mittlere Rangplatz** über alle Untertests. In jedem Untertest kann ein Rangplatz für jede Person berechnet werden. In der Praxis gibt es pro Untertest 21 verschiedene Rangplatz-Werte, also für jeden möglichen Punktwert einen. Der Rangplatz ist umso niedriger, je höher die Punktzahl ist und je weniger Personen insgesamt einen besseren Punktwert erreicht haben. Rangplätze gewichten besser, wie viele Personen besser oder schlechter als eine bestimmte Person sind.

Beispiel: Werden in 2 Untertests maximal 20 Punkte erzielt, gehen immer 40 Punkte in den Punktwert ein. Haben in einem Untertest 100 Personen diesen Wert erreicht, erhalten diese Personen den mittleren Rangplatz 50. Haben im anderen Untertest nur 10 Personen diesen Wert erreicht, ist der Rangplatz 5. Letzterer wird höher bewertet, weil die 20 Punkte im zweiten Untertest schwerer zu erzielen waren und deshalb wertvoller sind.

Die Rangplätze über alle neun Untertests können gemittelt werden und dieser „mittlere Rangplatz der Untertests“ kann als alternativer Zulassungsparameter verwendet werden.

Rangplätze aus dem Vorjahr können auch ins Folgejahr umgerechnet werden, indem sie auf die Zahl der Personen standardisiert werden. Folgende Ergebnisse gelten für 2004:

- Der Rangplatz bewegt sich zwischen 49 und 1236. Da er über 9 Untertests gemittelt ist, wird er nicht 1 oder 1373 (Zahl der Personen).
- Er korreliert mit dem Testwert mit -0.992 , was für praktische Äquivalenz spricht.
- Der Unterschied ist, dass er feiner differenziert als der Testwert. 720 unterscheidbare Rangplätze im Jahr 2004 stehen 57 unterscheidbaren Testwerten gegenüber.

Aus diesem Grunde hat der Beirat beschlossen, ab 2005 weiter den Testwert als Zulassungskriterium zu verwenden. Bei **gleichem Testwert** werden die Personen in der Reihenfolge des mittleren Rangplatzes aller Untertests berücksichtigt, wenn die Kapazität keine Zulassung aller Personen mit dem gleichen Testwert mehr zulässt. Dies wird für den jeweils niedrigsten Testwert gelten, der noch zur Zulassung ausreicht. Es werden allerdings nur noch so viele Personen zugelassen, wie freie Plätze tatsächlich vorhanden sind. Der mittlere Rangplatz wird auf dem Testbescheid mitgeteilt werden.

7 Testanwendung in der Schweiz 2004

7.1 Organisation der Testabnahme

Für die siebte Testanwendung in der Schweiz wurden 2 zusätzliche Testlokale notwendig. Trotzdem konnte in allen Testorten auf bewährte und erfahrene Personen und Strukturen zurückgegriffen werden. In jedem Testlokal waren entweder der Testleiter oder deren Stellvertreter mindestens einmal vorher an der Testabnahme beteiligt. Aus den Protokollen ergeben sich keine Hinweise auf irreguläre Vorkommnisse oder Bedingungen. Die Testabnahme kann daher in allen Testlokalen als uneingeschränkt gültig angesehen werden.

Vier Personen haben den Test während des Vormittagsteils abgebrochen: Alle haben im gleichen Testlokal den Test absolviert (eine „Ansteckung“ ist wahrscheinlich). Da nur zwei Verwarnungen wegen regelwidrigem Verhalten ausgesprochen wurden, weist dies auf eine sehr disziplinierte Testabnahme hin.

7.2 Verteilungsprüfung

Der Punktwert folgt wiederum einer Normalverteilung (Tabelle 16). Metrische Verfahren und die Standardisierung zum Testwert durch lineare z-Transformation sind daher zulässig. Die Punktwerte der Einzeltests sind auch dieses Jahr nicht normalverteilt (

Abbildung 17 und

Abbildung 18).

Von Bedeutung ist, dass „Konzentriertes und Sorgfältiges Arbeiten“ deutlicher einer Normalverteilung folgt. Die Häufung der maximalen Punktwerte ist nicht mehr vorhanden, was auf die Reduktion von Übungeffekten zurückzuführen ist.

	Punkt- wert	Muster zuord- nen	Med.- naturw. Grundv.	Schlauch- figuren	Quant.u. formale Probl.	Textver- ständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konz. u. sorgf. Arbeiten
Mittelwert	98.67	10.39	10.36	11.95	9.87	8.86	12.90	12.29	10.55	11.50
Stand.- abweich.	19.19	2.86	3.39	3.55	3.61	3.29	3.66	3.72	3.34	4.29
extremste Differenz	0.03	0.08	0.06	0.07	0.07	0.10	0.07	0.07	0.07	0.07
Kolmog.- Smirnov Z	0.97	3.09	2.36	2.52	2.54	3.71	2.52	2.73	2.46	2.59
Asymp. Sig.	0.30	.000	.000	.000	.000	.000	.000	.000	.000	.000

Tabelle 16: Verteilungsprüfung auf Normalverteilung (Kolmogorov-Smirnov Test)

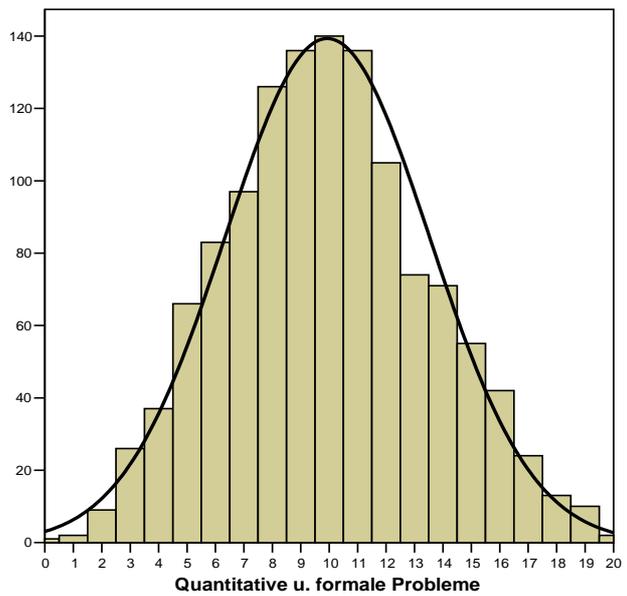
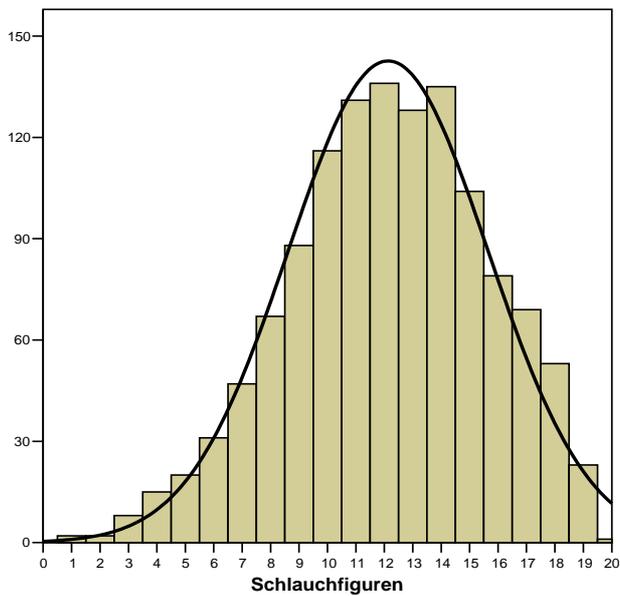
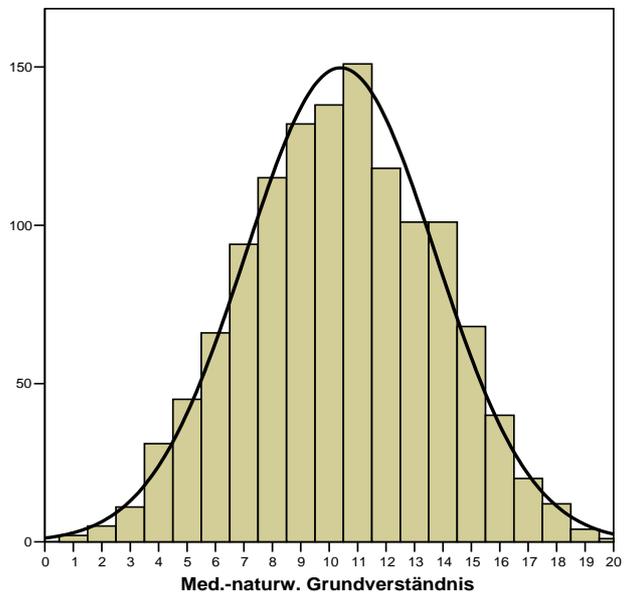
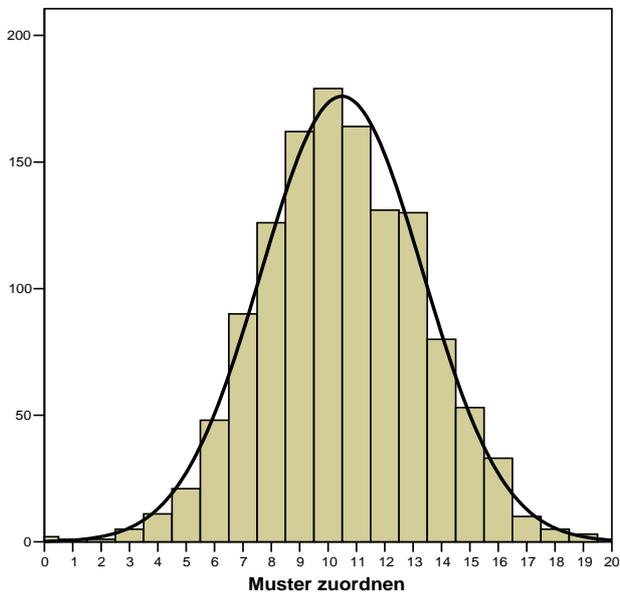
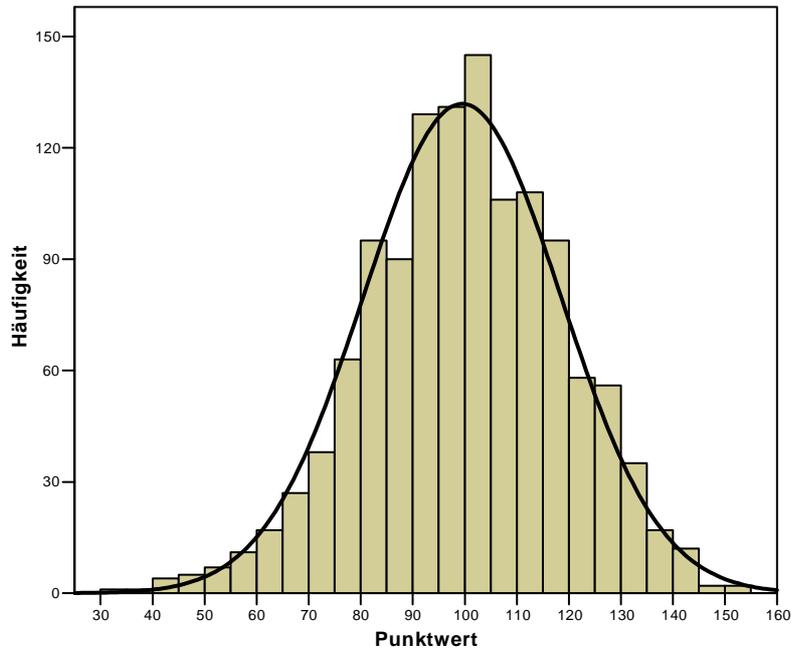


Abbildung 17: Häufigkeitsverteilungen für Punktwert und Punkte der Untertests 1-4

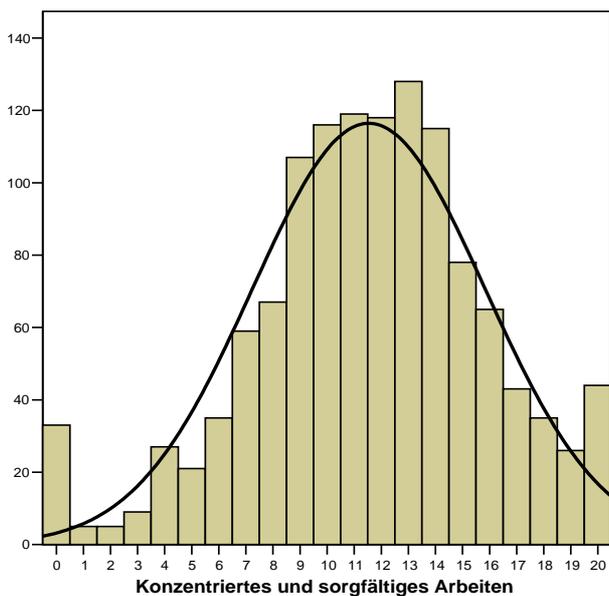
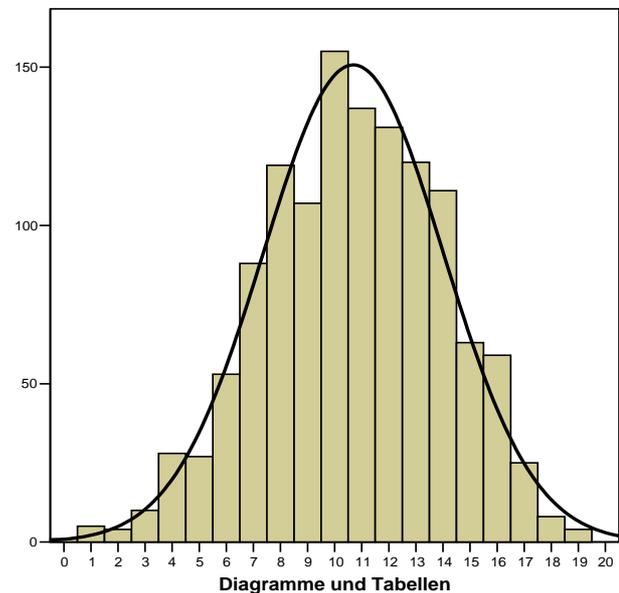
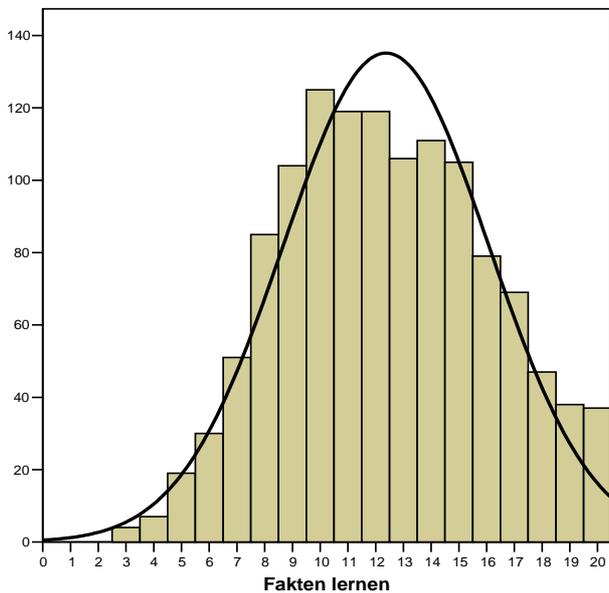
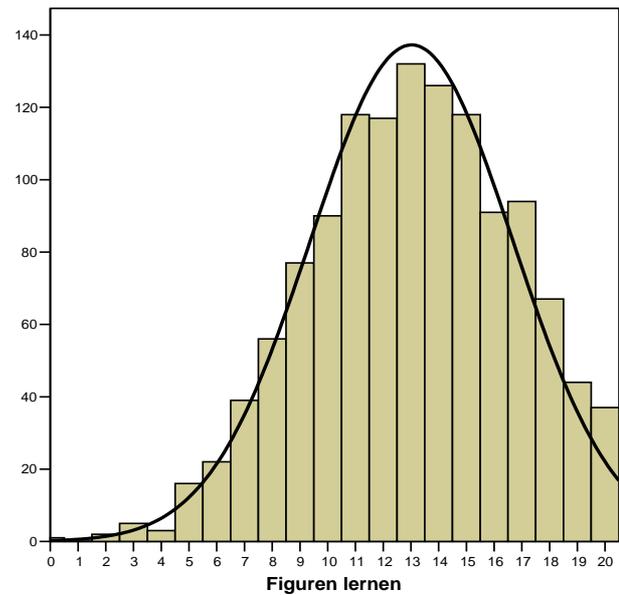
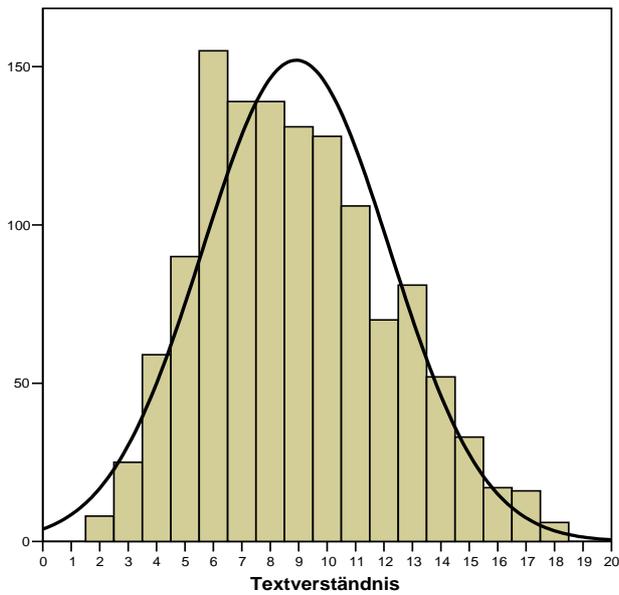


Abbildung 18: Häufigkeitsverteilungen für Punkte der Untertests 5-9

7.3 Vergleich der Testfassungen 1998 bis 2004

Durch die Transformation der Punktwerte in eine Skala, unter Berücksichtigung des jeweiligen Mittelwertes und der Standardabweichungen pro Jahr, ist gewährleistet, dass die **Testwerte** aus verschiedenen Jahren miteinander verglichen werden können. Dieses Vorgehen wird durch die jährlich nahezu identischen Gütekriterien zusätzlich legitimiert.

Die „absolute“ Schwierigkeit als **Punktwert** unterliegt über die Jahre Schwankungen. Tabelle 17 zeigt den Vergleich für alle bisherigen Testdurchführungen in der Schweiz mit den geschätzten Punktwerten aus Deutschland (Schätzung aufgrund der Item-Schwierigkeiten). Für die Berechnung der Schweizer Punktwerte wurden nur die Ergebnisse der deutschsprachigen Gruppe berücksichtigt.

	2004		2003		2002		2001		2000		1999		1998	
	D	CH	D	CH	D	CH	D	CH	D	CH	D	CH	D	CH
Muster zuordnen	10.3	10.5	10.2	10.8	10.9	11.3	9.7	10.1	9.3	10.3	11.0	11.6	12.2	13.2
Med.-naturwiss. Grundverst.	10.0	10.4	10.2	10.5	9.5	9.9	10.7	10.1	10.8	10.7	11.2	11.3	11.6	12.5
Schlauchfiguren	11.8	12.1	12.0	12.7	11.8	12.1	12.2	13.3	12.3	13.2	11.9	12.9	11.8	13.3
Quant. und form. Probleme	9.4	9.9	9.6	10.5	9.9	11.1	9.8	10.9	10.2	11.8	9.7	11.6	9.9	11.9
Textverständnis	9.0	8.9	8.9	10.2	8.6	9.2	8.6	8.5	9.3	9.0	8.9	8.9	10.2	10.3
Figuren lernen	11.5	13.0	10.7	13.6	11.8	15.7	9.2	10.6	9.3	10.2	11.8	12.7	11.6	12.6
Fakten lernen	11.3	12.4	10.3	11.4	9.7	10.7	9.3	10.8	8.9	10.1	11.0	11.8	11.2	11.6
Diagramme und Tabellen	10.1	10.7	10.7	11.7	10.5	10.8	10.0	10.3	9.7	10.4	10.3	10.5	11.0	11.4
Total (ohne Konzent. u. Sorgf. Arbeiten)	83.4	87.9	82.6	91.4	82.7	90.9	79.5	84.6	79.8	85.7	85.8	91.3	89.5	96.8

Tabelle 17: Mittelwerte (Punkte) der Untertests 1998 bis 2004 deutsche Sprachgruppe Schweiz und Schätzung aufgrund der Item-Schwierigkeitswerte aus Deutschland (HM: Humanmedizin)

Auch 2004 liegt das Gesamtergebnis aus der Schweiz (deutsche Sprachgruppe) über den Werten aus Deutschland. Diese Differenz zwischen den Gesamtergebnissen ist im Vergleich mit anderen Jahren allerdings geringer ausgefallen. Mit 88 Punkten liegt das Durchschnittsergebnis im Bereich früherer Durchführungen.

Im Untertest „Textverständnis“ erzielten die deutschen Teilnehmer einen etwas höheren Mittelwert, was den bisherigen Schwankungen in diesem Untertest entspricht.

Die folgenden Abbildungen veranschaulichen den Vergleich zwischen deutschen und Deutschschweizer Ergebnissen für die Untertests über die einzelnen Testdurchführungen.

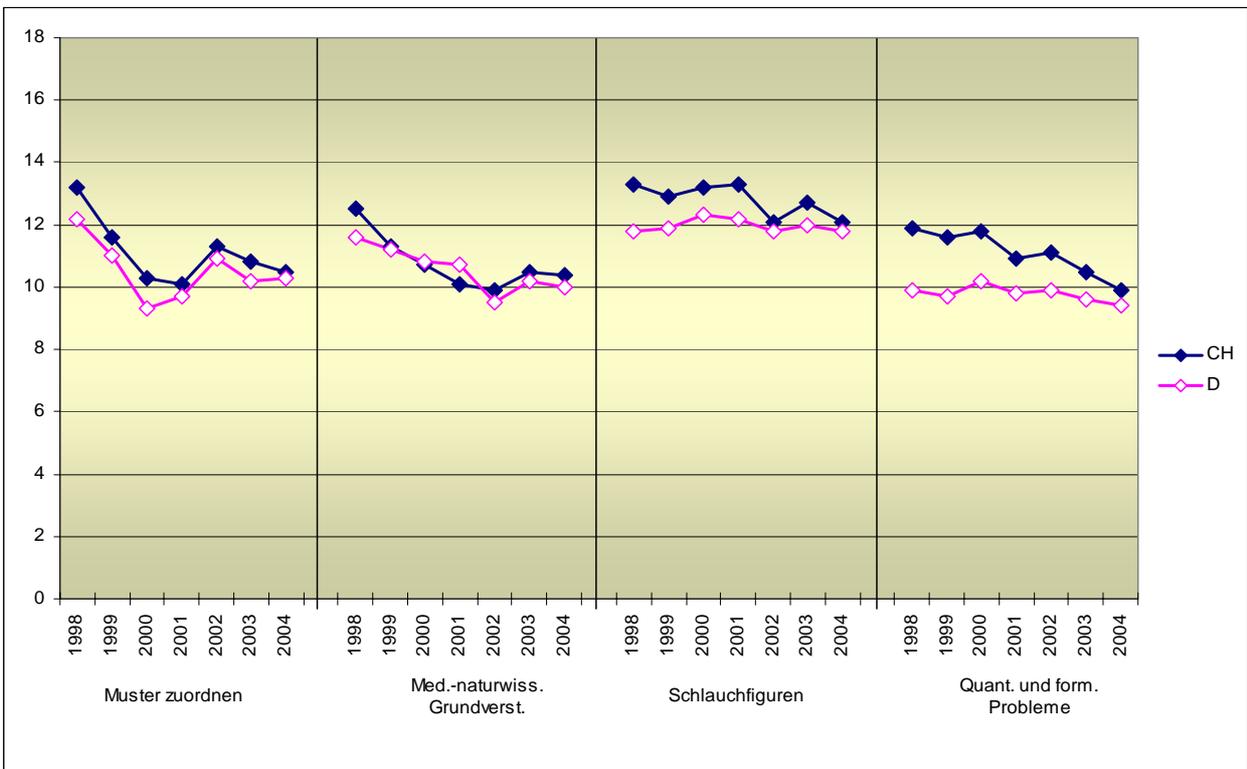


Abbildung 19: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2004 (Untertests 1-4)

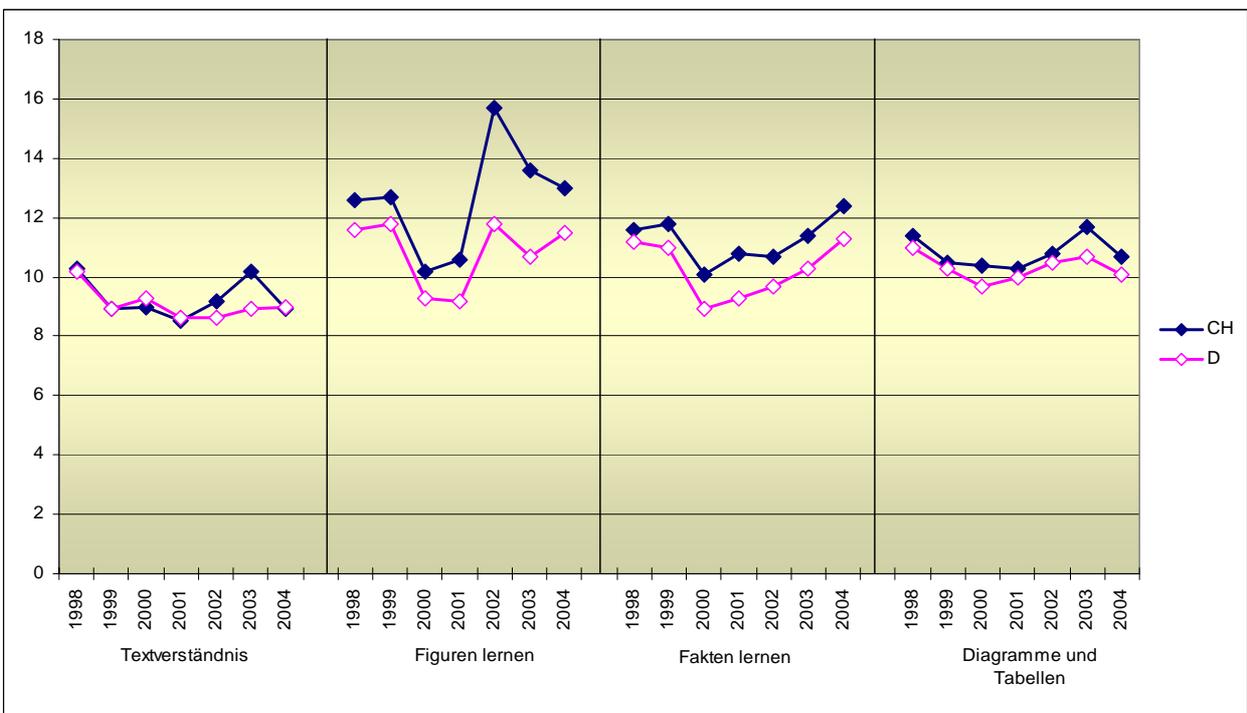


Abbildung 20: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2004 (Untertests 5-8)

7.4 Äquivalenz der Sprachversionen

Der Test wurde nach den gleichen Kriterien wie in den Vorjahren von der deutschen Vorlage adaptiert. Eine ausführliche Beschreibung des Verfahrens kann den Berichten 4 (1998) und 5 (1999) entnommen werden. Zielsetzung waren möglichst gut vergleichbare Aufgabenschwierigkeiten zwischen den Sprachversionen.

Chancengleichheit als Gleichbehandlung der Sprachgruppen muss gewährleistet sein, die Lösungswahrscheinlichkeit einer Aufgabe muss bei gleicher Fähigkeit tatsächlich auch vergleichbar sein.

Diese Gleichheit ist nicht allein durch Übersetzung zu beeinflussen, sondern hängt von mehreren Faktoren ab. Bei allen Adaptationen muss mit Unterschieden hinsichtlich der Aufgabenschwierigkeiten zwischen den Sprachversionen gerechnet werden. Die Abbildung 21 fasst zusammen, welche Ursachen für diese Differenzen verantwortlich sein können.

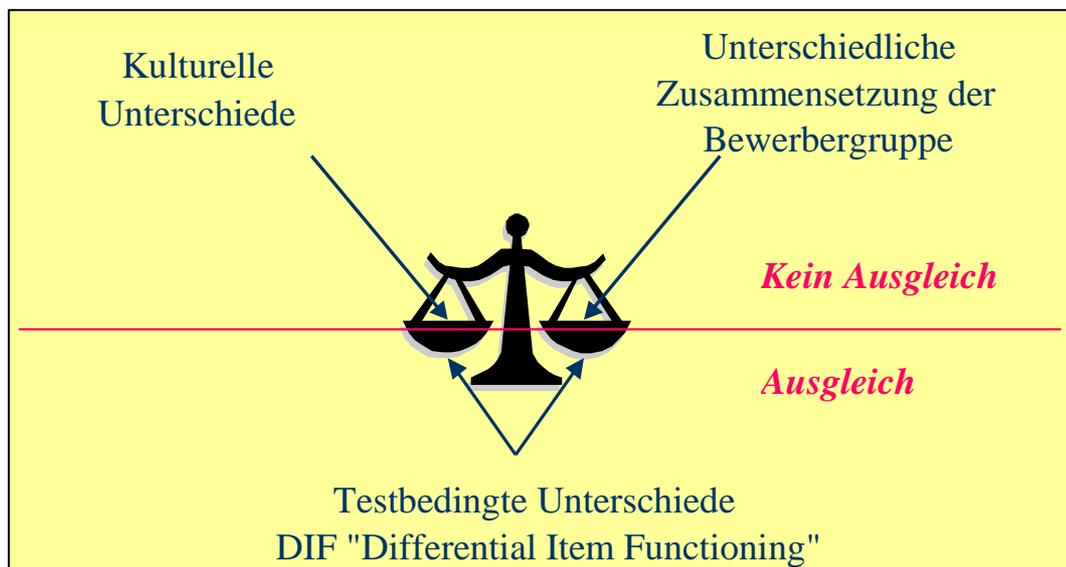


Abbildung 21: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen

Zwei Ursachen führen zu „echten“ Personen-Unterschieden, die nicht auf den Test oder seine Adaptation zurückgeführt werden können. Sie müssen durch die Ergebnisse abgebildet werden – hier erfolgt kein Ausgleich.

Testbedingte Unterschiede werden durch die Adaptation verursacht und sollen möglichst vollständig ausgeglichen werden. Beispiele wären eine Vereinfachung oder Erschwerung der Aufgabe bei der Übersetzung (Wortwahl, Satzgliederung o.ä.).

Wie bereits in den Vorjahren wird dazu das DIF-Verfahren¹ angewendet: Eine ausführliche Beschreibung kann beispielsweise in Bericht 5 (1999) nachgelesen werden.

¹ DIF als Abkürzung für Differential Item Functioning

Kulturelle Unterschiede zwischen den Sprachgruppen

- Mögliche generelle Fähigkeitsunterschiede oder unterschiedliche Fähigkeitsprofile (anderes Verhältnis der einzelnen Fähigkeiten zueinander) im Vergleich der Kulturen können vorhanden sein, die z.B. möglicherweise auf Unterschiede im Schulsystem zurückzuführen sind. Auch die Strategien, wie entsprechende Aufgaben gelöst werden, können sich unterscheiden und sich auf die Ergebnisse auswirken.
- Unterschiede in Leistungs-Voraussetzungen (z.B. Motivation, Belastbarkeit, Ausdauer) können zu unterschiedlichen Resultaten beitragen.

Unterschiedliche Repräsentativität der Stichproben für die Sprachgruppen

- Die jeweils untersuchten Stichproben können verschiedene Ausschnitte aus der jeweiligen Sprachgruppe sein. Die geringe Zahl der französisch- und italienischsprachigen Teilnehmer macht dies wahrscheinlich: Die Auswahl kann die „Spitze“ oder das „Ende“ der Leistungsrangreihe aller Maturanden der jeweiligen Sprachgruppe überrepräsentieren, etwa durch unterschiedliches „Wahlverhalten“ aufgrund der vorhandenen Alternativen für Studienorte.

Testbedingte Unterschiede

- Die Testaufgaben können nach der Übertragung eine unterschiedliche Aufgabenschwierigkeit aufweisen, indem durch Satzstellung, Wortwahl, Kompliziertheit des Satzes etc. ein Unterschied auftritt. Die Lösungsschwierigkeit eines Items wird unter anderem auch von der Formulierung und dem Satzaufbau einer Fragestellung beeinflusst. Bereits geringe Änderungen innerhalb einer Sprache können zu unterschiedlichen Schwierigkeiten führen. Allerdings sind diese Differenzen nicht vorherzusehen, sondern können erst empirisch nachgewiesen werden.

Der Beschluss der Schweizerischen Hochschulkonferenz aus dem Jahre 1999 ist weiterhin bindend. Er beinhaltet:

1. einen Ausgleich nur bei sprachabhängigen Tests vorzunehmen. Dies sind fünf von neun Untertests. Der Sprachausgleich beschränkt sich demnach auf die eher sprachabhängigen Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ sowie „Diagramme und Tabellen“;
2. den Ausgleich nur vorzunehmen, wenn Mittelwertunterschiede im entsprechenden Untertest vorhanden sind. Geprüft werden dazu die Abweichungen der jeweiligen Zielsprache von der deutschen Sprachgruppe.

Adaptationsbedingte Effekte sind praktisch nur in den fünf sprachabhängigen Untertests („Textverständnis“, „Diagramme und Tabellen“, „Fakten lernen“, „Quantitative und formale Probleme“, „Medizinisch-naturwissenschaftliches Grundverständnis“) zu erwarten, während die vier weitgehend sprachunabhängigen Untertests („Muster zuordnen“, „Figuren lernen“, „Schlauchfiguren“, „Konzentriertes und Sorgfältiges Arbeiten“) davon praktisch nicht betroffen sind.

Spricht man von DIF, so muss dieses klar von systematischen Mittelwertsdifferenzen, welche alle Aufgaben eines Untertests gleichermaßen betreffen, abgegrenzt werden. Die Ergebnisse der letzten Jahre haben gezeigt, dass sich die Sprachgruppen in manchen Untertestscores signifikant unterscheiden. Differenzen kommen nicht nur durch einzelne Items zustande, sondern auch durch eine systematische Verschiebung der Lösungsschwierigkeit aller Items eines Untertests. Es ist unwahrscheinlich, dass diese

systematischen Unterschiede testbedingt sind. Testbedingte Unterschiede als sprachliche Besonderheiten zeigen sich in spezifischen Abweichungen einzelner Items. Systematische Unterschiede sind dagegen mit grösserer Wahrscheinlichkeit den Faktoren „Kulturunterschiede“ und „unterschiedliche Repräsentativität“ zuzurechnen.

Es ergibt sich für das Ausgleichsverfahren die nachfolgende Zielstellung:

Das Risiko, dass testbedingte Unterschiede die Sprachunterschiede systematisch beeinflussen, soll verringert werden. Andererseits soll kein Ausgleich von Unterschieden aufgrund unterschiedlicher Zusammensetzung der Stichproben und der kulturellen Unterschiede erfolgen.

7.4.1 Sprachvergleich für die Untertests

Wie in den Vorjahren treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Untertests gleichermassen auf. Auch dieses Jahr liegen die Punktzahlen der deutschsprachigen Teilnehmer durchwegs über jenen der französisch- und italienischsprachigen Kandidaten. Ein signifikanter Unterschied zwischen französisch- und italienischsprachigen Teilnehmern ist im Untertest „Figuren lernen“ zu beobachten (Abbildung 22).

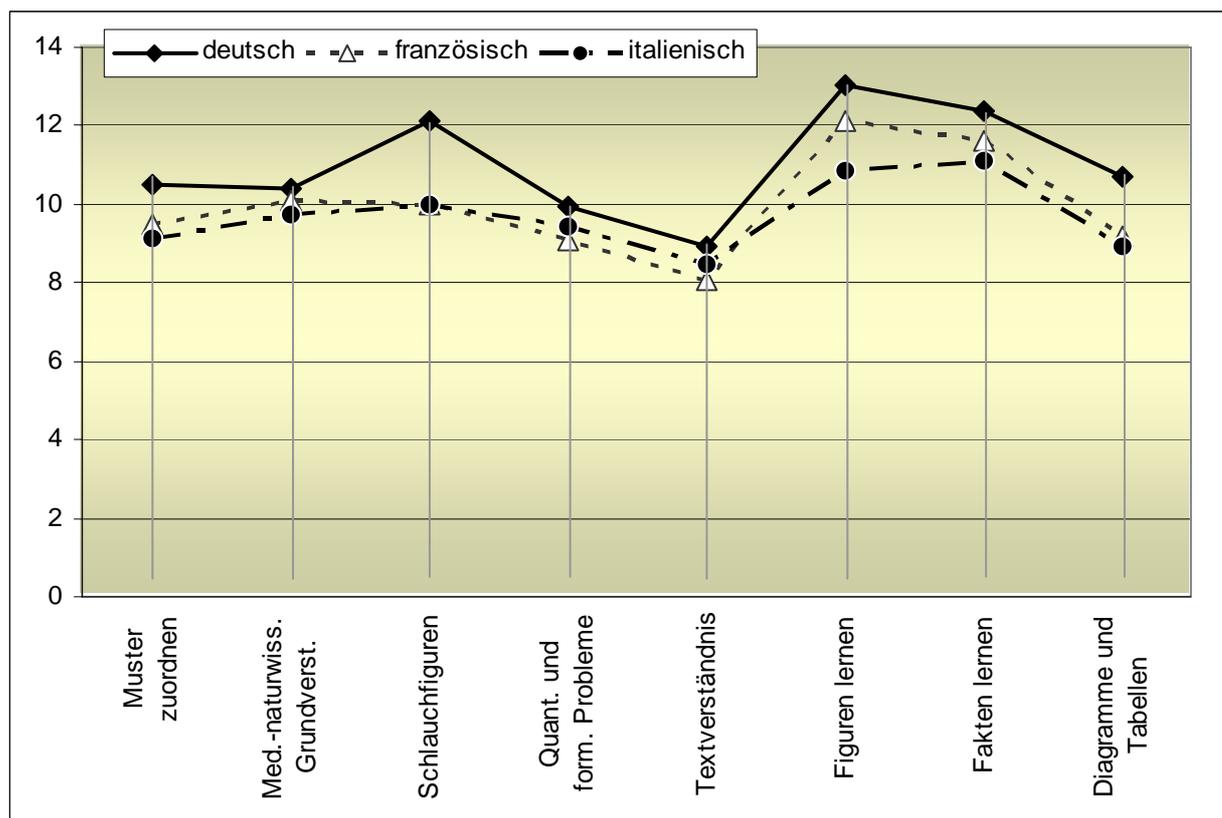


Abbildung 22: Mittelwerte der Untertests für die Sprachgruppen 2004 (unkorrigiert)

Die Abbildung 23 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern für alle Testanwendungen seit 1998. Die senkrechten Linien bezeichnen die Spannweite (Minimum bis Maximum) der Differenzen über sämtliche Jahrgänge, die Verbindungslinie steht für die Differenzen aus dem aktuellen Jahrgang. Die Unterschiede bewegen sich dieses Jahr im Vergleich zu früheren Jahrgängen im unteren Bereich.

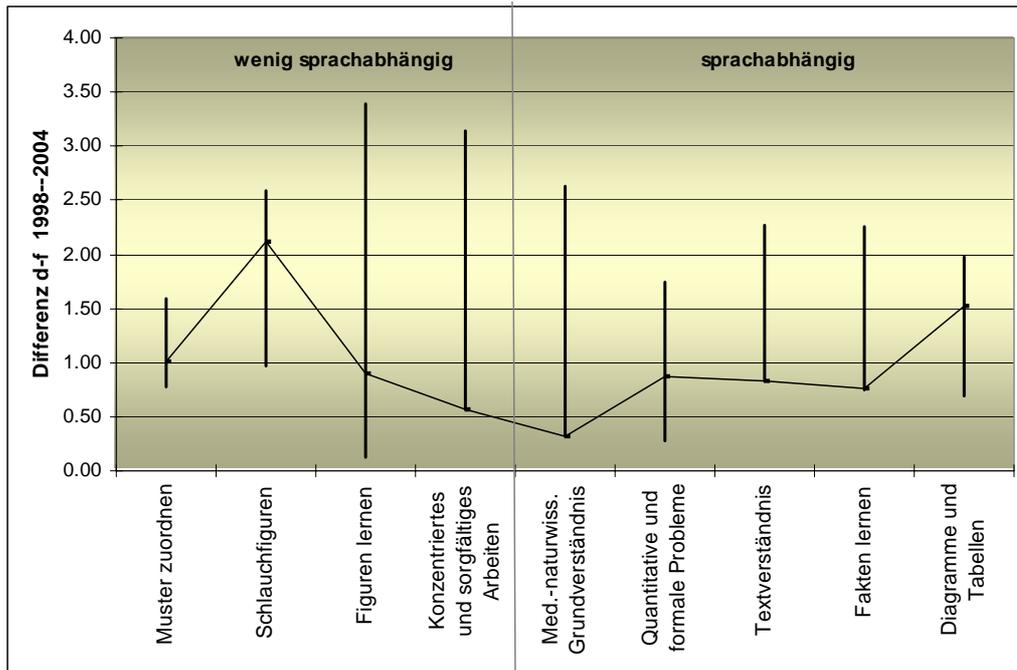


Abbildung 23: Differenz Testwert deutsch-französisch 2004 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2004

Im Vergleich zwischen deutsch- und italienischsprachigen Kandidaten wird im Untertest „Muster zuordnen“ ein neuer Höchstwert erreicht (Abbildung 24). Die Unterschiede in den sprachabhängigen Untertests fallen erneut tendenziell eher geringer aus. Im neu konzipierten Untertest „Konzentriertes und Sorgfältiges Arbeiten“ fallen die Differenzen kleiner aus als in den bisherigen Durchführungen mit dem ursprünglichen Konzept.

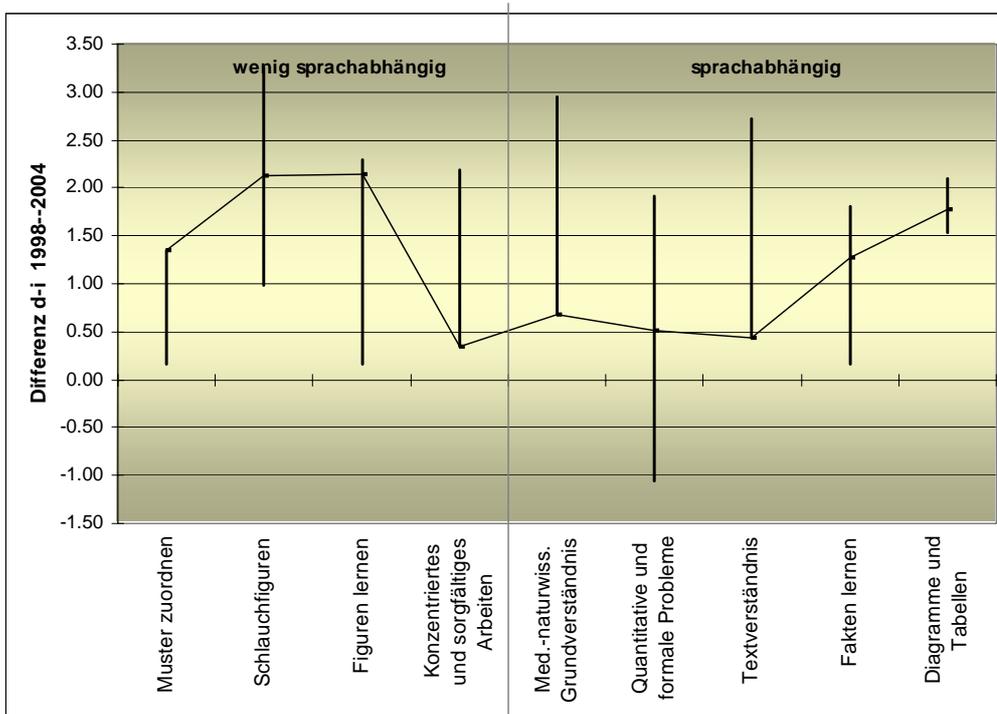


Abbildung 24: Horizontale Linie: Differenz Testwert deutsch-italienisch 2004 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2004 (vertikale Linien)

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Testwert ** $D > (F,I)$	deutsch	1255	100.44	9.90	70	129
	französisch	68	95.99	9.78	79	126
	italienisch	50	95.30	10.16	70	119
Muster zuordnen ** $D > (F,I)$	deutsch	1255	10.49	2.85	0	19
	französisch	68	9.49	3.05	3	16
	italienisch	50	9.14	2.56	3	15
Med.-naturwiss. Grundverständnis <i>homogen</i>	deutsch	1255	10.41	3.34	1	20
	französisch	68	10.09	3.47	3	20
	italienisch	50	9.72	3.56	3	18
Schlauchfiguren ** $D > (F,I)$	deutsch	1255	12.46	3.67	1	20
	französisch	68	10.31	3.50	2	19
	italienisch	50	10.12	3.37	4	17
Quantitative und formale Probleme <i>homogen</i>	deutsch	1255	9.94	3.58	1	20
	französisch	68	9.06	4.00	1	20
	italienisch	50	9.42	3.44	2	18
Textverständnis <i>homogen</i>	deutsch	1255	8.91	3.29	2	18
	französisch	68	8.09	3.36	1	17
	italienisch	50	8.48	3.12	3	16
Figuren lernen ** $(D,F) > I$	deutsch	1255	13.02	3.65	0	20
	französisch	68	12.12	3.11	5	19
	italienisch	50	10.88	3.91	2	19
Fakten lernen <i>homogen</i>	deutsch	1255	12.37	3.70	3	20
	französisch	68	11.62	3.46	4	17
	italienisch	50	11.43	4.09	4	20
Diagramme und Tabellen ** $D > (F,I)$	deutsch	1255	10.69	3.32	0	19
	französisch	68	9.48	2.96	3	18
	italienisch	50	9.39	3.28	2	16
Konzentriertes und sorgf. Arbeiten <i>homogen</i>	deutsch	1255	11.54	4.30	0	20
	französisch	68	10.97	4.33	0	20
	italienisch	50	11.20	4.06	1	20

Tabelle 18: Testwert und Punktwerte der Untertests für die Sprachgruppen

** : signifikanter Einfluss des Faktors „Sprache“ bei varianzanalytischer Prüfung (1%-Niveau); daneben sind signifikante Unterschiede bei multiplern Mittelwertvergleich angegeben. In französischer und italienischer Sprachgruppe korrigierte Punktwerte

Die Mittelwertsdifferenzen für jeden Untertest **nach** dem Sprachausgleich sind der Tabelle 18 zu entnehmen.

Die Ergebnisse multipler Mittelwertvergleiche können ebenfalls der ersten Spalte entnommen werden. Differenzen zwischen den Teilnehmern sind dort beispielsweise in der Form $D > (F,I)$ angegeben. Das Beispiel würde bedeuten, dass die deutschsprachigen Kandidaten signifikant höhere Werte erzielt haben als die beiden anderen Sprachgruppen.

7.4.2 Darstellung des Korrekturverfahrens

Die Identifikation und der Ausgleich testbedingter Unterschiede geschehen wie folgt:

- 1) Sprachausgleiche werden nur in den sprachabhängigen Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ sowie „Diagramme und Tabellen“ vorgenommen. Dass Unterschiede in den nichtsprachlichen Tests testbedingt sind (und ausgeglichen werden müssen), ist unwahrscheinlich. Die Aufgaben selber erfordern praktisch keine Sprachkompetenzen und die sprachspezifischen Anforderungen der Instruktion scheinen so einfach, dass Verständnisunterschiede unwahrscheinlich sind.
- 2) Damit ein Sprachausgleich in einem bestimmten sprachabhängigen Untertest erfolgt, müssen sich die entsprechenden mittleren Punktzahlen signifikant unterscheiden (Vergleich deutsch - französisch, beziehungsweise deutsch – italienisch mittels t-Test). Wenn keine Mittelwert- und Varianzunterschiede vorhanden sind, trägt der Untertest auch nicht zu Mittelwert-Unterschieden beim Testwert bei. Eine Korrektur allfälliger DIF-Items (die vorhanden sein können) würde zusätzliche Unterschiede produzieren.
- 3) Es werden jene Items ausgeglichen, welche über eine systematische Differenz hinaus DIF aufweisen. Für diese Items ist der Verdacht am ehesten begründet, dass es sich um testbedingte Differenzen handelt. Der Ausgleich wird nicht symmetrisch vorgenommen, da es darum geht, eine Benachteiligung der Vergleichsgruppe gegenüber der Referenzgruppe zu vermeiden.
- 4) Als Sprachausgleich wird den französisch- und italienischsprachigen Probanden die Differenz zur entsprechenden Itemschwierigkeit in der deutschsprachigen Version gutgeschrieben, falls sie das Item nicht gelöst haben. Um jede Benachteiligung auszuschliessen, wird nicht nur die Differenz zur Regressionsgeraden ausgeglichen. Die Lage dieser Regressionsgeraden wird auch von allen DIF-Items mitbestimmt. Man kann das so interpretieren, dass für DIF-Items bei Nichtlösung zumindest der Schwierigkeitsunterschied zur Referenzgruppe ausgeglichen wird.

Eine generelle Gutschrift des Mittelwertunterschiedes pro Untertest für die französisch- und italienischsprachige Gruppe könnte für einzelne Personen zu mehr Punkten pro Untertest führen als es theoretisch gibt und sich auf die Gewichtung der Untertests negativ auswirken. Aus diesem Grund wird insgesamt nicht mehr als ein Punkt pro Aufgabe vergeben, und nur Personen, welche das betreffende Item nicht gelöst haben, erhalten den Bonus.

7.4.2.1 Welche Untertests ausgleichen?

Die folgenden Tabellen zeigen die Sprachvergleiche für die eher sprachabhängigen Untertests. Referenz ist die deutschsprachige Testversion. Die französischsprachigen Teilnehmer (Tabelle 19) weisen in den Untertests „Textverständnis“ und „Diagramme und Tabellen“ signifikant geringere Ergebnisse als die deutschsprachigen Teilnehmer auf.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	125 5	10.41	3.34	.447	Nein
	f	68	10.09	3.47		
Quantitative und formale Probleme	d	125 5	9.94	3.58	.051	Nein
	f	68	9.06	4.00		
Textverständnis	d	125 5	8.91	3.29	.045	Ja
	f	68	8.09	3.36		
Fakten lernen	d	125 5	12.37	3.70	.101	Nein
	f	68	11.62	3.46		
Diagramme und Tabellen	d	125 5	10.69	3.32	.000	Ja
	f	68	9.16	3.03		

Tabelle 19: Signifikanzprüfung deutsche vs. französische Sprachform

Die italienische Form (Tabelle 20) weist in zwei sprachabhängigen Untertests signifikante Unterschiede auf: „Fakten lernen“ sowie „Diagramme und Tabellen“. Für die übrigen Untertests erfolgt kein Sprachausgleich.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	125 5	10.41	3.34	.156	Nein
	i	50	9.72	3.56		
Quantitative und formale Probleme	d	125 5	9.94	3.58	.314	Nein
	i	50	9.42	3.44		
Textverständnis	d	125 5	8.91	3.29	.362	Nein
	i	50	8.48	3.12		
Fakten lernen	d	125 5	12.37	3.70	.018	Ja
	i	50	11.10	4.24		
Diagramme und Tabellen	d	125 5	10.69	3.32	.000	Ja
	i	50	8.92	3.42		

Tabelle 20: Signifikanzprüfung deutsche vs. italienische Sprachform

Mit der Bestimmung der auszugleichenden Untertests kann noch nichts darüber ausgesagt werden, ob es sich bei den Differenzen um „reale“ oder um adaptationsbedingte Unterschiede handelt. Zu diesem Zweck ist eine Analyse der Aufgaben der betreffenden Untertests notwendig.

7.4.2.2 Identifikation von DIF-Items und Bestimmung des Korrekturwertes

Eine Möglichkeit zur Identifikation von DIF, auch bei kleinen Stichprobengrößen, ist die Methode „Delta-Plot“. Bei diesem Verfahren werden die zu vergleichenden Item-Schwierigkeiten z-standardisiert und anschliessend in „ Δ -Werte“ transformiert. Die Transformation erfolgt über die Formel

$$\Delta = 13 - 4z$$

Dies bedeutet, dass Δ einen Mittelwert von 13 und eine Standardabweichung von 4 aufweist. Hohe Werte stehen für „schwierige“ (von weniger Probanden gelöste) Items.

Die aus den Delta-Werten abgeleitete Regressionsgerade $Y = AX + B$ beschreibt die Beziehung zwischen den interessierenden Sprachversionen. Eine graphische Darstellung der Delta-Werte (Delta-Plot) würde im Idealfall eine ellipsenförmige Anordnung von Punkten entlang der Diagonalen ergeben (Abbildung 25). Dies würde bedeuten, dass sowohl die Itemschwierigkeiten wie auch deren Reihenfolge in beiden Sprachversionen vergleichbar sind.

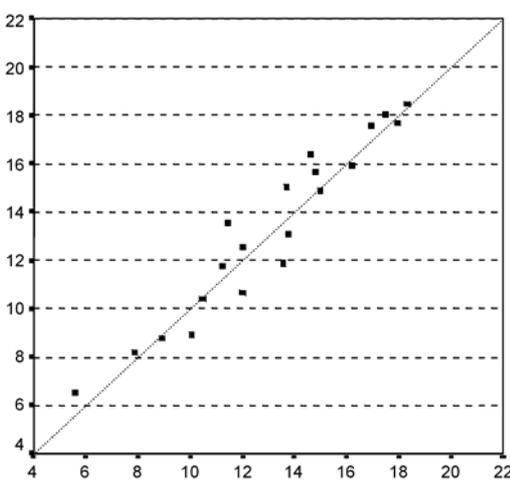


Abbildung 25: Beispiel für einen Delta-Plot

Die Diagonale wird in den folgenden Abbildungen durch eine durchgezogene Linie dargestellt. Ein Abweichen der Werte von dieser Diagonalen ergibt eine Regressionsgerade, welche sich sowohl in Steigung wie Achsenschnittpunkt von der Diagonalen unterscheiden kann. Solche Verschiebungen der Regressionsgeraden stehen für systematische Unterschiede in den Untertests, deren Ursachen eher nicht testbedingt (Übersetzung), sondern in realen Gegebenheiten zu vermuten sind. Um die Regressionsgerade gruppierte Items folgen also diesen Gegebenheiten und sind dementsprechend nicht auszugleichen.

Auffällige Abweichungen einzelner Punkte von dieser Regressionsgeraden hingegen weisen auf Items hin, welche zusätzlich zu einer eventuellen systematischen Verschiebung spezielle Eigenschaften aufweisen. Bei einer parallelen Verschiebung der Regressionsgeraden bleibt trotz einer Veränderung der absoluten Schwierigkeiten die „Schwierigkeits-Rangfolge“ der Items erhalten. Von der Regressionsgeraden abweichende Items stimmen aber in der untersuchten Sprachgruppe bezüglich der „Schwierigkeitshierarchie“ nicht mit der Referenzgruppe überein (sie sind im Vergleich zu den anderen Items zu schwer oder zu leicht ausgefallen). Die Ursache solcher Differenzen kann eher testbedingt, also beispielsweise in der Übersetzung vermutet werden. Betroffene Items müssen nicht in jedem Fall als „schlecht“ übersetzt betrachtet werden. Denkbar ist etwa auch, dass in der Fragestellung Konzepte enthalten sein könnten, welche in den Sprachgruppen unterschiedliche Bekanntheitsgrade aufweisen.

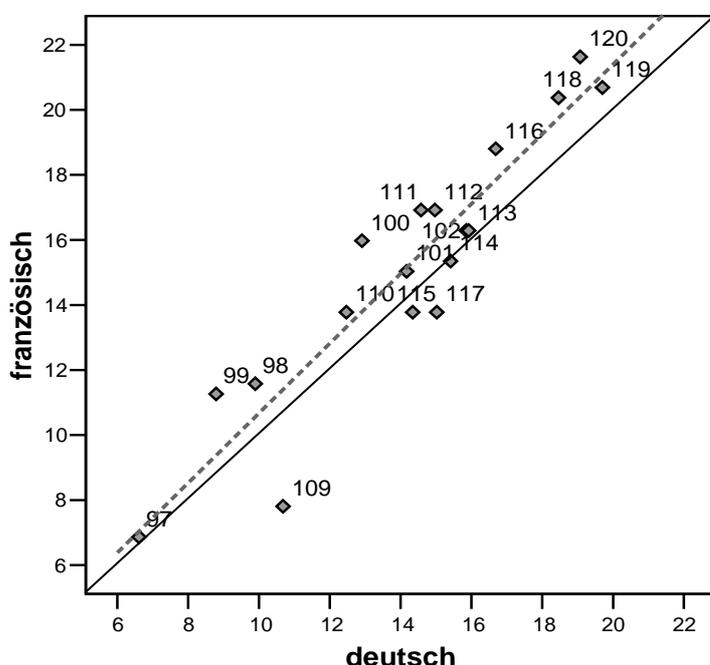
Regressionsgeraden werden in der Folge als durchgezogene Linien dargestellt. Massgeblich für die Entscheidung, ob bei einem Item DIF vorliegt, ist die Distanz des entsprechenden Punktes von dieser Geraden. Die Distanz D wird nach der Formel

$$D_i = \frac{AX_i - Y_i + B}{\sqrt{A^2 + 1}}$$

berechnet, wobei unter A die Steigung und B der Achsenschnittpunkt der Regressionsgeraden zu verstehen ist, X_i bezeichnet den Delta-Wert der Referenzgruppe, Y_i denjenigen der zu vergleichenden Gruppe.

In der Folge sollen innerhalb der auszugleichenden Untertests die kritischen Items identifiziert werden. Da, wie oben erwähnt, keine symmetrische DIF-Behandlung verwendet wird, ist die Richtung der Abweichung der Regressionsgeraden von der Diagonalen massgebend. Es werden also jene Items ausgeglichen, **welche mindestens 1.5 Punkte** (Longford, Holland & Thayer, 1993) von der Regressionsgeraden in entgegengesetzter Richtung zur Diagonalen abweichen. Als Sprachausgleich wird der benachteiligten Sprachgruppe die Differenz der betreffenden Itemschwierigkeit zur deutschen Itemschwierigkeit gutgeschrieben (also die Differenz der Lösungswahrscheinlichkeit). Dies betrifft nur jene Personen, welche das fragliche Item nicht korrekt beantwortet haben.

7.4.2.3 Textverständnis (TV)



Obwohl sich die Ergebnisse der beiden Sprachgruppen signifikant unterscheiden, ist diese Differenz nicht auf einzelne Items zurückzuführen, sondern kommt einzig aufgrund eines systematischen Unterschieds zustande. Am weitesten in „benachteiligender“ Richtung von der Regressionsgeraden entfernt ist die Aufgabe 100. Diese verfehlt jedoch knapp das Kriterium zum Ausgleich (-1.5). Der entsprechende Bonus hätte dort 0.15 Punkte betragen. Im Untertest „Textverständnis“ wurden trotz signifikantem Mittelwertsunterschied für die französischsprachigen Teilnehmer kein Bonus vergeben.

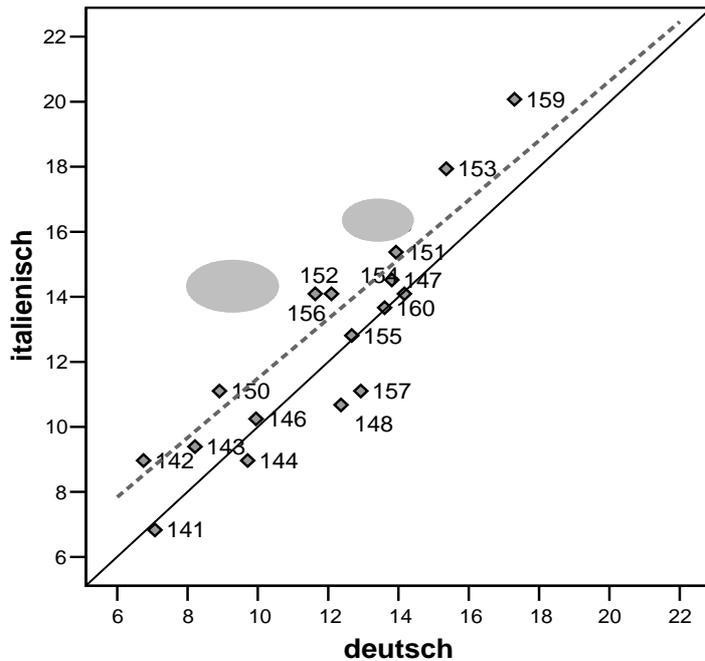
Abbildung 26: Delta-Plot für „TV“ (deutsch-französisch)

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
97	0.85	0.84	0.96	6.61	6.87	4.27	0.12			
98	0.70	0.62	0.60	9.90	11.58	11.96	-0.69			
99	0.75	0.63	0.64	8.79	11.27	11.10	-1.29			
100	0.56	0.41	0.52	12.91	15.98	13.67	-1.49			
101	0.50	0.46	0.42	14.17	15.03	15.80	0.08			
102	0.42	0.40	0.44	15.85	16.29	15.37	0.45			
109	0.66	0.79	0.68	10.68	7.81	10.25	2.45			
110	0.58	0.51	0.56	12.47	13.78	12.81	-0.31	Kein Ausgleich		Kein Ausgleich
111	0.48	0.37	0.52	14.58	16.92	13.67	-0.91			
112	0.46	0.37	0.46	14.97	16.92	14.95	-0.62			
113	0.41	0.40	0.42	15.92	16.29	15.80	0.50			
114	0.44	0.44	0.32	15.41	15.35	17.94	0.77			
115	0.49	0.51	0.46	14.34	13.78	14.95	1.06			
116	0.38	0.28	0.24	16.69	18.80	19.65	-0.65			

117	0.46	0.51	0.52	15.02	13.78	13.67	1.55			
118	0.30	0.21	0.24	18.46	20.37	19.65	-0.43			
119	0.24	0.19	0.14	19.70	20.69	21.78	0.27			
120	0.27	0.15	0.34	19.07	21.63	17.51	-0.84			

Tabelle 21: DIF-Analyse „Textverständnis“

7.4.2.4 Fakten lernen



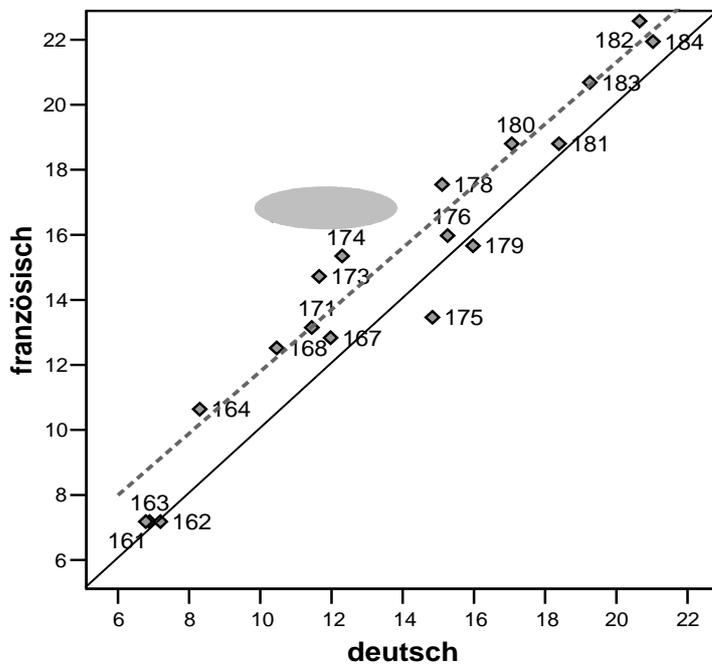
Im Vergleich zwischen deutscher und italienischer Fassung kann für die Fragen 145, 149 (identischer Wert) und 158 eine Distanz über dem kritischen Wert beobachtet werden. Für diese drei Fragen wird ein Bonus von maximal 0.63 Punkten vergeben.

Abbildung 27: Delta-Plot für „Fakten lernen“ (deutsch-italienisch)

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
141	0.83	0.84	0.84	7.07	6.87	6.83		1.47		
142	0.84	0.78	0.74	6.75	8.12	8.97		-0.33		
143	0.78	0.71	0.72	8.21	9.69	9.39		0.34		
144	0.71	0.71	0.74	9.71	9.69	8.97		1.67		
145	0.74	0.78	0.50	8.89	8.12	14.09		-2.67		0.24
146	0.69	0.76	0.68	9.95	8.44	10.25		0.88		
147	0.51	0.47	0.48	13.81	14.72	14.52		0.34		
148	0.58	0.66	0.66	12.36	10.64	10.67		2.20		
149	0.74	0.68	0.50	8.88	10.32	14.09		-2.68		0.24
150	0.74	0.66	0.64	8.91	10.64	11.10		-0.45		
151	0.51	0.57	0.44	13.93	12.52	15.37		-0.21		
152	0.62	0.38	0.50	11.63	16.61	14.09		-0.82		
153	0.44	0.32	0.32	15.36	17.86	17.94		-1.14		
154	0.50	0.63	0.50	14.17	11.27	14.09		0.89		
155	0.57	0.44	0.56	12.67	15.35	12.81		0.83		
156	0.59	0.57	0.50	12.09	12.52	14.09		-0.51		
157	0.55	0.51	0.64	12.93	13.78	11.10		2.26		
158	0.55	0.50	0.40	12.94	14.09	16.23		-1.51		0.15
159	0.35	0.21	0.22	17.30	20.37	20.07		-1.41		
160	0.52	0.43	0.52	13.61	15.66	13.67		0.83		

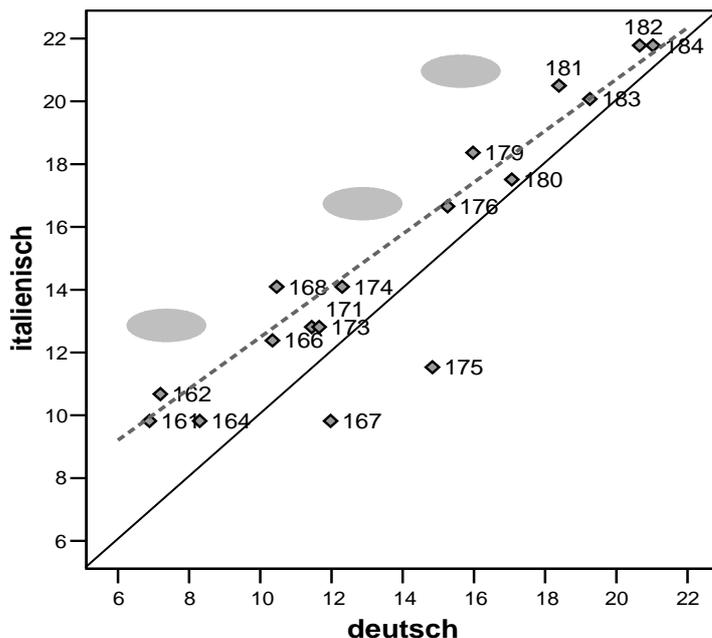
Tabelle 22: DIF-Analyse „Fakten lernen“

7.4.2.5 Diagramme und Tabellen



Im Vergleich zwischen deutscher und französischer Version des Untertests „Diagramme und Tabellen“ kann eine fast parallele Verschiebung der Regressionsgeraden beobachtet werden. Die beiden Aufgaben 166 und 170 weisen darüber hinaus starke Differenzen bezüglich ihrer Schwierigkeit auf und werden ausgeglichen.

Abbildung 28: Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch)



In der italienischen Fassung fällt vor allem eine „Erschwerung“ der leichteren Fragen auf. Über diese systematische Verschiebung hinaus weisen die Items 163, 170 und 178 höhere Schwierigkeiten auf. Für diese drei Aufgaben werden maximal 0.76 Punkte als Bonus gutgeschrieben.

„Diagramme und Tabellen“ (deutsch-italienisch)

Abbildung 29: Delta-Plot für

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
161	0.84	0.82	0.70	6.88	7.18	9.82	1.20	0.09		
162	0.82	0.82	0.66	7.19	7.18	10.67	1.41	-0.38		
163	0.84	0.82	0.56	6.78	7.18	12.81	1.13	-2.29		0.28
164	0.77	0.66	0.70	8.30	10.64	9.82	-0.33	0.98		
166	0.68	0.38	0.58	10.34	16.61	12.38	-3.25	0.30	0.29	
167	0.60	0.56	0.70	11.97	12.84	9.82	0.61	3.32		
168	0.67	0.57	0.50	10.46	12.52	14.09	-0.21	-0.95		
170	0.59	0.37	0.38	12.23	16.92	16.66	-2.18	-1.80	0.22	0.21
171	0.62	0.54	0.56	11.45	13.15	12.81	0.02	0.67		
173	0.61	0.47	0.56	11.65	14.72	12.81	-0.98	0.80		
174	0.58	0.44	0.50	12.30	15.35	14.09	-0.99	0.22		
175	0.47	0.53	0.62	14.83	13.46	11.53	2.12	3.81		
176	0.45	0.41	0.38	15.26	15.98	16.66	0.59	0.12		
178	0.45	0.34	0.18	15.11	17.55	20.93	-0.65	-3.28		0.27
179	0.41	0.43	0.30	15.97	15.66	18.36	1.31	-0.75		
180	0.36	0.28	0.34	17.06	18.80	17.51	-0.21	0.60		
181	0.30	0.28	0.20	18.39	18.80	20.50	0.70	-0.86		
182	0.19	0.10	0.14	20.65	22.57	21.78	-0.47	-0.42		
183	0.26	0.19	0.22	19.26	20.69	20.07	-0.06	0.02		
184	0.18	0.13	0.14	21.03	21.95	21.78	0.25	-0.18		

Tabelle 23: DIF-Analyse „Diagramme und Tabellen“

7.4.2.6 Effekte der Korrektur

Die durch den Ausgleich resultierenden Bonuspunkte können Tabelle 24 entnommen werden. Durchschnittlich werden den französischsprachigen Teilnehmern 0.32 Punkte, den italienischsprachigen Teilnehmern 0.80 Punkte gutgeschrieben. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der Bonus in beiden Gruppen beträgt also 0 bis 1 Punkte.

Franz.		n	Minimum	Maximum	m	s
		Diagramme und Tabellen	68	0.0	0.51	0.32
	GESAMT	68	0.0	0.51	0.32	0.18
Ital.	Fakten lernen	50	0.0	0.63	0.33	0.22
	Diagramme und Tabellen	50	0.0	0.76	0.47	0.24
	GESAMT	50	0.15	1.39	0.80	0.34

Tabelle 24: Mittelwerte und Standardabweichungen der Korrekturwerte für die Personen der französisch- und italienischsprachigen Gruppen

In Abbildung 30 und Abbildung 31 sind die Verteilungen der (ungerundeten) Bonuspunkte für die bisherigen Testdurchführungen dargestellt (1998 wurde kein Sprachausgleich vorgenommen). Die senkrechten Linien bezeichnen die Spannweite in jedem Jahr (Minimum bis Maximum), der jeweilige Mittelwert wird durch die waagerechte Markierung gekennzeichnet. Für die französischsprachige Gruppe wurde ein neuer Tiefstwert erreicht. Bei den italienischsprachigen Teilnehmern handelt es sich um den zweitniedrigsten Wert

nach 2003. Dies ist auch ein Zeichen dafür, dass die Übersetzungs- bzw. Adaptationsqualität immer besser wird.

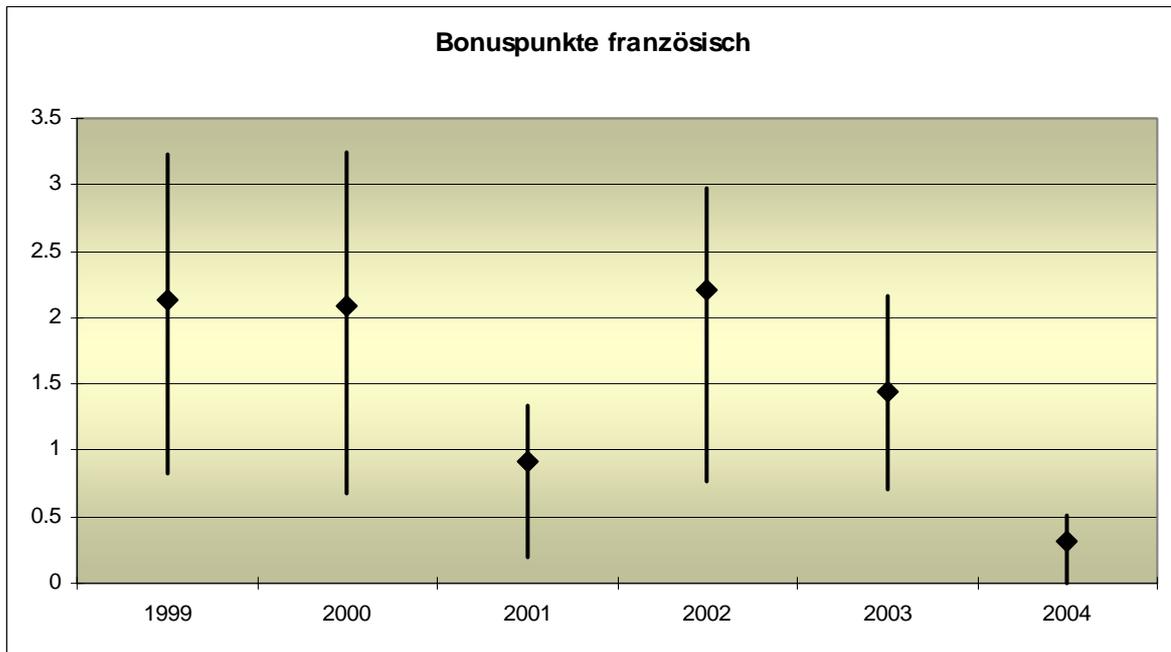


Abbildung 30: Auswirkungen der Korrektur über die Jahre für die französischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

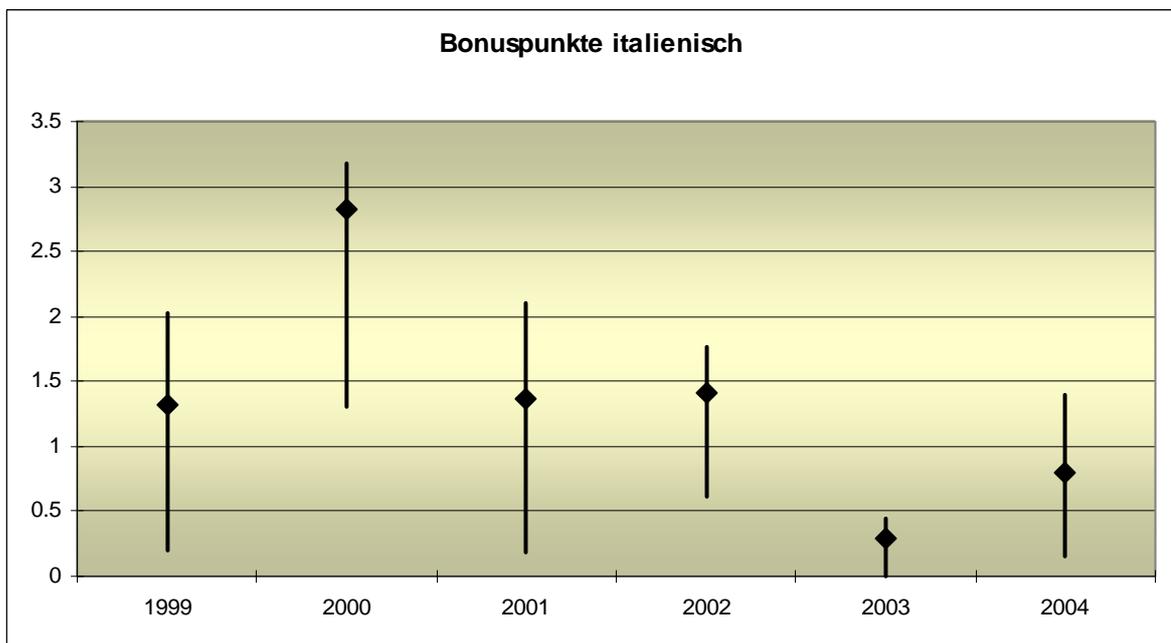


Abbildung 31: Auswirkungen der Korrektur über die Jahre für die italienischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

7.4.2.7 Evaluation: Gleichbehandlung der Sprachgruppen

Wirken sich die vorhandenen Unterschiede in den Punktwerten auf die Zulassung aus? In den Vorjahren zeigten sich bezüglich der Zulassungsquoten zur Humanmedizin signifikante Unterschiede zwischen den Sprachgruppen. Nachdem diese 2003 nicht mehr aufgetreten waren, sind sie diesmal in der Disziplin „Humanmedizin“ wieder zu beobachten (Tabelle 25 und Tabelle 26). Die Zulassungsquote ist 2004 für die deutsche Sprachgruppe höher als für die französisch- und italienischsprachigen Teilnehmer. Aufgrund des beschriebenen Vorgehens und der weiter unten aufgeführten Analysen kann jedoch ausgeschlossen werden, dass diese Unterschiede auf Eigenheiten des Tests zurückzuführen sind.

Disziplin	Sprache	Warteliste	zugelassen	Total
Humanmedizin	deutsch	383 39.9%	576 60.1%	959 100.0%
	französisch	20 62.5%	12 37.5%	32 100.0%
	italienisch	20 57.1%	15 42.9%	35 100.0%
	Total	423 41.2%	603 58.8%	1026 100.0%
Veterinärmedizin	deutsch	33 21.4%	121 78.6%	154 100.0%
	französisch	12 34.3%	23 65.7%	35 100.0%
	italienisch	3 42.9%	4 57.1%	7 100.0%
	Total	48 24.5%	148 75.5%	196 100.0%

Tabelle 25: Zulassung nach Sprache des Tests

Disziplin		Wert	df	Asymp. Sig. (2seitig)
Humanmedizin	Pearson Chiquadrat	10.294	2	.006
	Likelihood Ratio	10.103	2	.006
Veterinärmedizin	Pearson Chiquadrat	3.874	2	.144
	Likelihood Ratio	3.616	2	.164

Tabelle 26: Prüfung auf Verteilungsunterschiede für Sprachgruppen getrennt nach Disziplinen: kein signifikanter Unterschied für Humanmedizin und Veterinärmedizin

In den folgenden Abbildungen sind die Wartelisten-Quoten für Humanmedizin und Veterinärmedizin dargestellt. Hohe Werte bedeuten eine geringere Zulassungsquote. Die Rekordanmeldezahlen dieses Jahrganges wirken sich in massiv erhöhten Wartelistenquoten für alle Teilnehmer in Humanmedizin und Veterinärmedizin aus. Für die Zahnmedizin konnte jedem Teilnehmer mit gültigem Testergebnis ein Studienplatz vergeben werden.

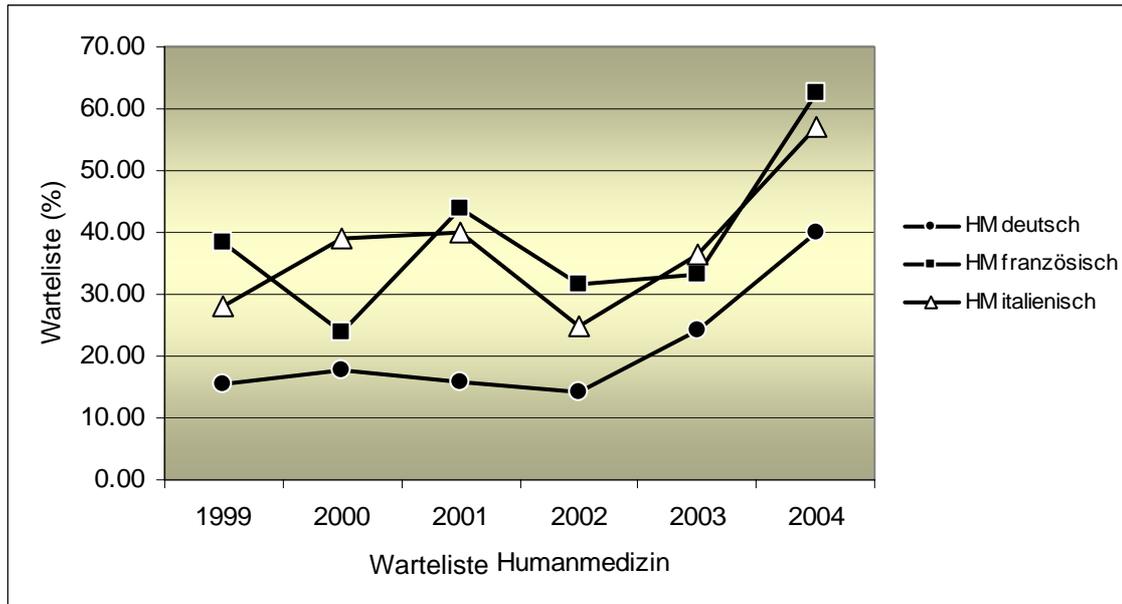


Abbildung 32: Wartelistenquoten Humanmedizin nach Sprache 1999 bis 2004

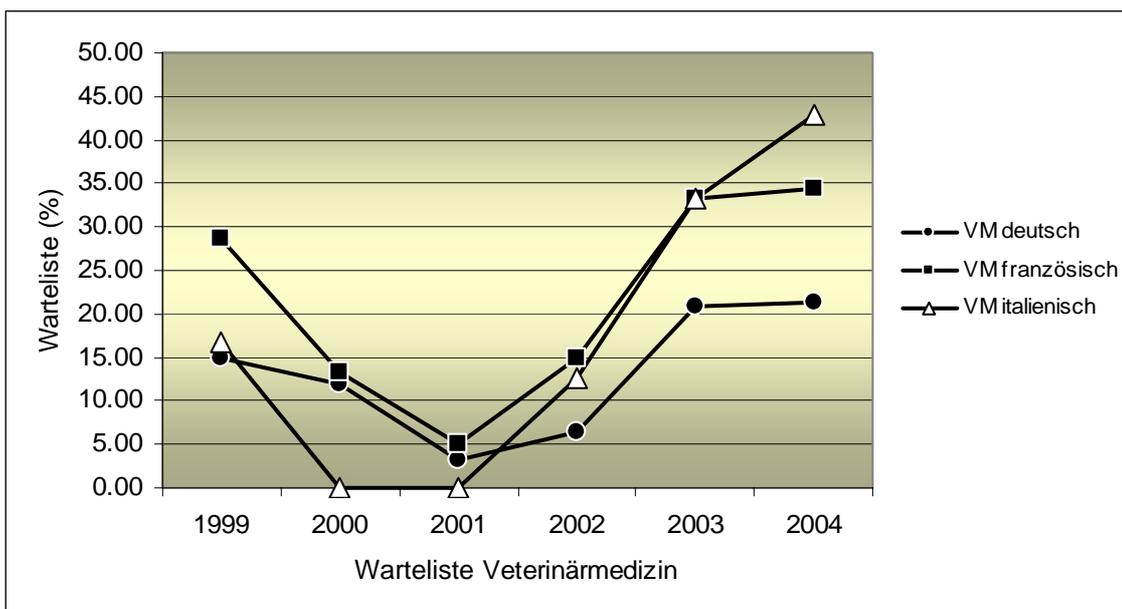


Abbildung 33: Wartelistenquoten Veterinärmedizin nach Sprache 1999 bis 2004

Es muss weiterhin sichergestellt sein, dass die unterschiedlichen Zulassungsquoten der Sprachgruppen nicht im Test selber begründet liegen. Dies kann durch eine Analyse des Verhältnisses zwischen sprachabhängigen und wenig sprachabhängigen Untertests recht gut belegt werden.

Die Differenzen der französischen und italienischen Teilnehmer von der deutschsprachigen Referenzgruppe werden nach folgender Formel berechnet:

$$\text{Delta } z = \frac{m_{\text{Sprachgruppe}} - m_{\text{deutsch}}}{S_{\text{deutsch}}}$$

m ist der Mittelwert und s die Standardabweichung der entsprechenden Sprachgruppe. Die resultierenden Werte „Delta z“ drücken die Abweichung der jeweiligen Sprachgruppe von der deutschen Sprachgruppe in Standardabweichungen aus. 0.5 hiesse, dass die betroffene Sprachgruppe um eine halbe Standardabweichung schlechter ist als die deutschsprachige Referenzgruppe.

In den wenig sprachabhängigen Untertests zeigen sich weiterhin für beide adaptierte Sprachversionen bereits vor der Korrektur geringere Differenzen zur deutschen Referenzversion als in den sprachabhängigen Untertests. Für die französischsprachige Version war einzig 1999 eine grössere Differenz in den sprachabhängigen Untertests zu beobachten. Im Jahr 2000 war das in der italienischen Version der Fall. In allen anderen Jahren fallen die Differenzen in den sprachabhängigen Untertests geringer aus. Weiterhin kann eher von einer Überkompensation als von einer Benachteiligung der französisch- und italienischsprachigen Kandidaten ausgegangen werden.

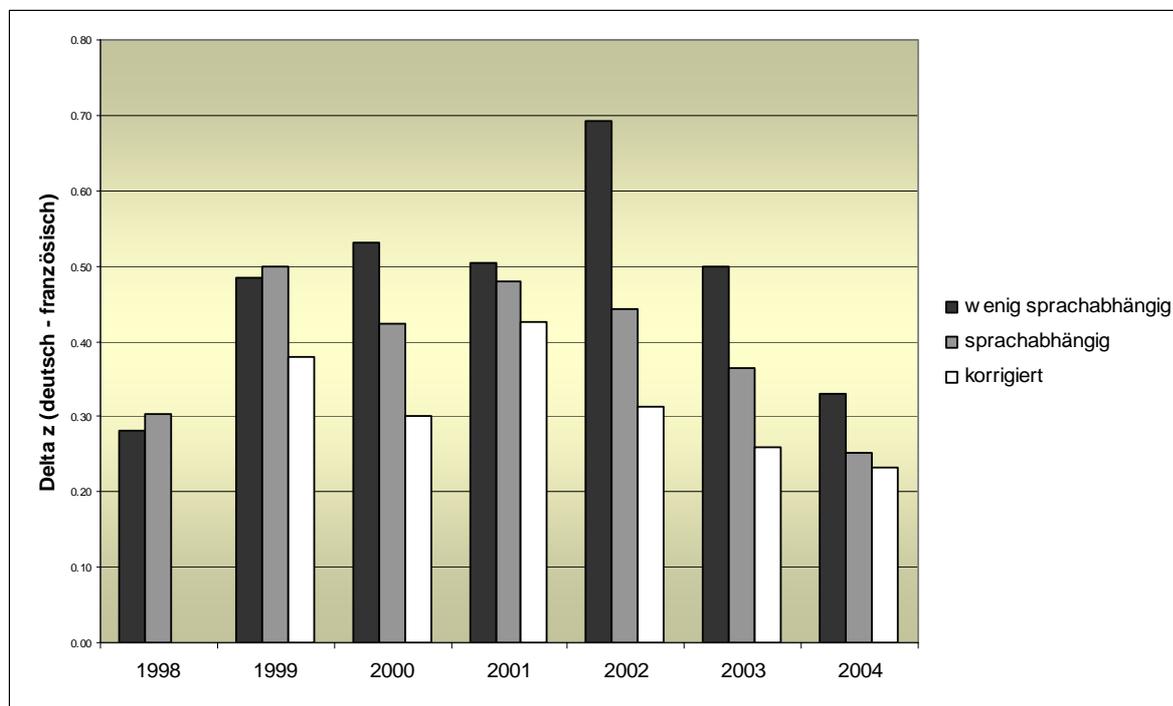


Abbildung 34: Abweichung (Delta z) für den Testwert zwischen deutscher und französischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

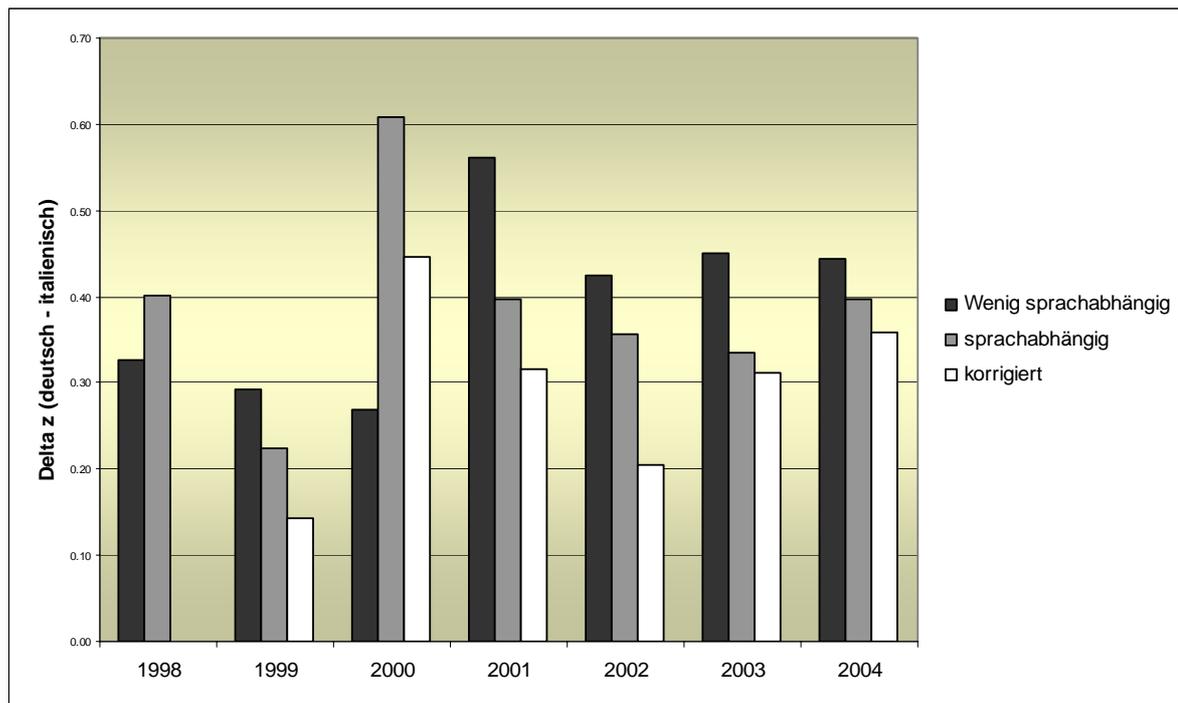


Abbildung 35: Abweichung (Delta z) für den Testwert zwischen deutscher und italienischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

7.4.3 Leistungen in identischen Testaufgaben 1998 – 2004

Zum Vergleich der Leistungen über die Jahre hinweg werden weiterhin „Einstreuaufgaben“ verwendet. Dabei handelt es sich um Items, welche sich im Vorjahr aufgrund ihrer Kennwerte bewährt haben. Die Werte in diesen Einstreuaufgaben lassen sich demnach zwischen zwei benachbarten Jahren vergleichen. Dabei ist zu beachten, dass ein anderer Aufgabenkontext und andere Positionen der Aufgaben im Test bereits für Unterschiede verantwortlich sein können. Da diese Positionen in den Sprachgruppen aber gleich sind, können die tatsächlichen Unterschiede der Sprachgruppen besser abgeschätzt werden.

	Jahr	n	Mittelwert	Standardabw.	Sig.
deutsch	2003	977	14.97	3.70	.000
	2004	1255	15.58	3.70	
französisch	2003	69	13.62	3.10	.258
	2004	68	14.28	3.72	
italienisch	2003	42	13.94	3.90	.697
	2004	50	14.24	3.90	

Tabelle 27: Verteilungskennwerte identischer Einstreuaufgaben 2003 und 2004

Abbildung 36 zeigt den Differenzverlauf für die Einstreuaufgaben. Die französischsprachige Gruppe zeigte 1999 eine Verschlechterung gegenüber 1998, die italienischsprachige Gruppe schnitt im Jahr 2000 schlechter ab als 1999. Ansonsten kann durchwegs eine Verbesserung der Leistungen beobachtet werden. Die Leistungen der deutschsprachigen Teilnehmer sind erneut statistisch signifikant besser ausgefallen als im Vorjahr. In der französischen Version

ist der Unterschied aufgrund der geringeren Fallzahl statistisch nicht signifikant, eine Verbesserung ist aber auch hier tendenziell zu beobachten. Die Leistungen der italienischsprachigen Gruppe unterscheiden sich leicht von denjenigen des Vorjahres. Die Tendenz zu besseren Leistungen bleibt insgesamt bestehen, nimmt allerdings in ihrem Ausmass ab.

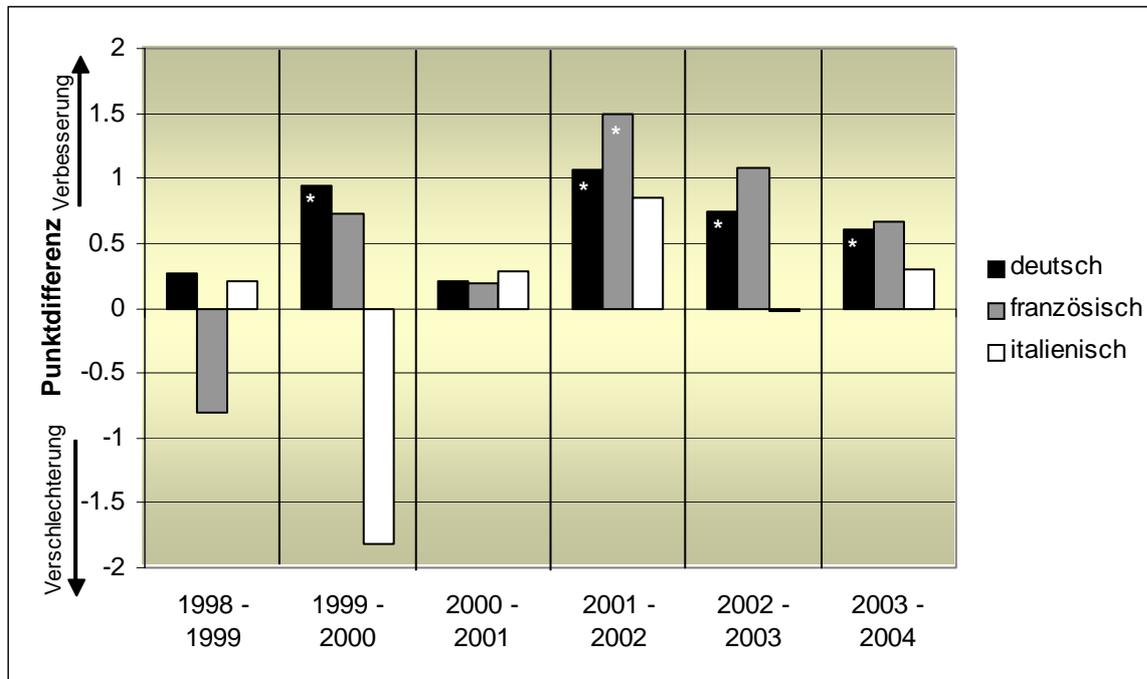


Abbildung 36: Verlauf der Leistungen in identischen Testaufgaben 1998 bis 2004 (Differenzverlauf); *= signifikante Differenz

7.5 Vergleichbarkeit der Testlokale

Die Durchführungsbedingungen in den einzelnen Testorten sind standardisiert. Dazu gehören beispielsweise die genaue Einhaltung von Zeiten, das wörtlich genaue Vorlesen von Instruktionen und einheitliches Reagieren auf mögliche Fragen und Probleme. Aus den Testprotokollen ergeben sich keine Hinweise, die als Einschränkungen dieser Chancengleichheit zu bewerten wären.

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testwert	Zwischen Gruppen	2305.595	14	164.685	1.715	.047
	Innerhalb Gruppen	119083.983	1240	96.035		
	Total	121389.578	1254			
Muster zuordnen	Zwischen Gruppen	226.210	14	16.158	2.018	.014
	Innerhalb Gruppen	9929.452	1240	8.008		
	Total	10155.662	1254			
Med.-naturwiss. Grundverständnis	Zwischen Gruppen	117.759	14	8.411	.751	.724
	Innerhalb Gruppen	13897.441	1240	11.208		
	Total	14015.200	1254			
Schlauchfiguren	Zwischen Gruppen	205.484	14	14.677	1.194	.273
	Innerhalb Gruppen	15244.027	1240	12.294		
	Total	15449.511	1254			
Quant. und formale Probleme	Zwischen Gruppen	460.890	14	32.921	2.595	.001
	Innerhalb Gruppen	15731.488	1240	12.687		
	Total	16192.378	1254			
Textverständnis	Zwischen Gruppen	317.606	14	22.686	2.118	.009
	Innerhalb Gruppen	13278.752	1240	10.709		
	Total	13596.359	1254			
Figuren lernen	Zwischen Gruppen	131.481	14	9.391	.704	.772
	Innerhalb Gruppen	16551.802	1240	13.348		
	Total	16683.283	1254			
Fakten lernen	Zwischen Gruppen	145.283	14	10.377	.754	.720
	Innerhalb Gruppen	17059.941	1240	13.758		
	Total	17205.224	1254			
Diagramme und Tabellen	Zwischen Gruppen	291.194	14	20.800	1.904	.022
	Innerhalb Gruppen	13545.232	1240	10.924		
	Total	13836.425	1254			
Konzentr. u. sorgf. Arbeiten	Zwischen Gruppen	386.145	14	27.582	1.500	.104
	Innerhalb Gruppen	22797.731	1240	18.385		
	Total	23183.876	1254			

Tabelle 28: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den Testlokalen

Die aufgetretenen Differenzen sind in erster Linie auf Unterschiede zwischen den Testorten zurückzuführen. Innerhalb der Orte zeigen sich keine Differenzen, welche auf systematische Benachteiligungen einzelner Lokale hinweisen würden.

7.6 Vergleich der Geschlechter

Die Forderung nach Gleichbehandlung der Geschlechter konnte in den bisherigen Testdurchführungen jeweils als erfüllt betrachtet werden. Chancengleichheit bedeutet, dass bei gleicher Eignung die gleichen Chancen auf eine Zulassung bestehen müssen. Beide Teilgruppen sind unterschiedlich repräsentativ in der Bewerberkohorte vertreten. Der Männeranteil in der Humanmedizin sank zwischen 1999 bis 2002 immer weiter ab (Abbildung 37). 2003 war erstmals wieder ein geringer Anstieg des Anteils an Kandidaten zu verzeichnen. 2004 wurde der Anteil erneut verringert und erreicht wieder den Wert von 2002.

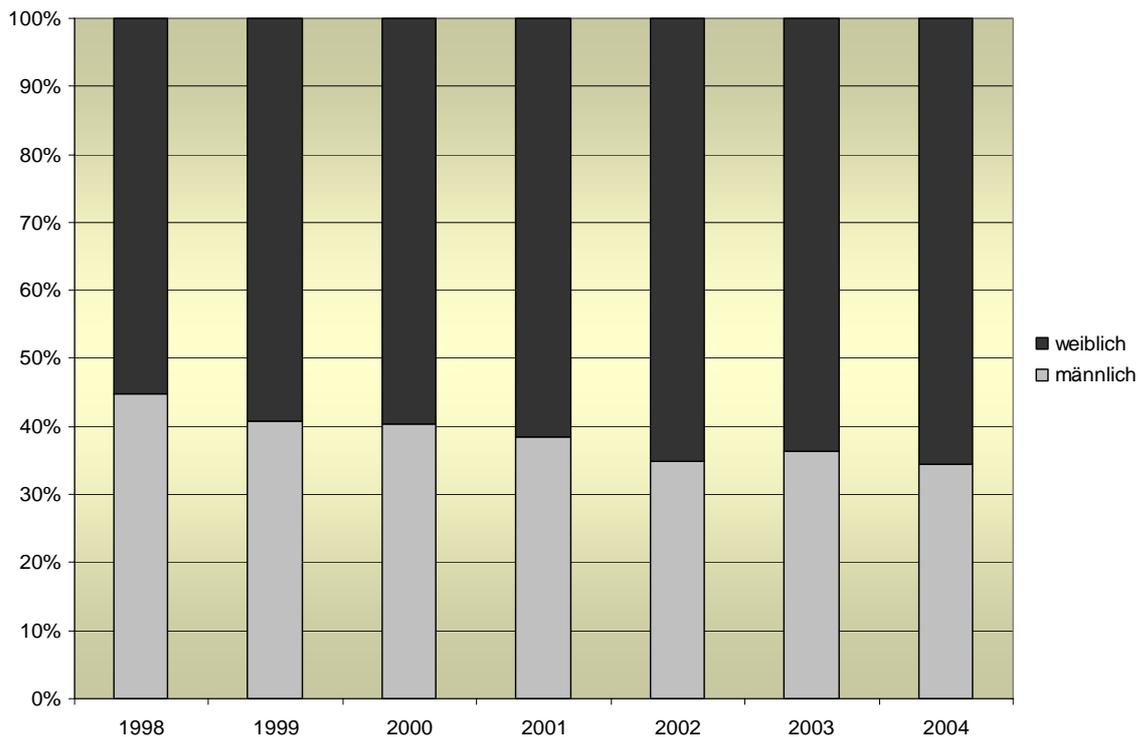


Abbildung 37: Bewerberkohorten (Humanmedizin) 1998 bis 2004 nach Geschlecht

Auch in diesem Jahr liegen die Testwerte der deutschsprachigen männlichen Teilnehmer in der Disziplin „Humanmedizin“ über denjenigen der Kandidatinnen (Tabelle 29). In den beiden anderen Disziplinen tritt der umgekehrte Fall ein.

Die Differenz bleibt wie bisher gering, auch wenn sie dieses Jahr signifikantes Niveau erreicht (Tabelle 30). Allein aufgrund der grösseren Personenzahlen sind kleinere Differenzen notwendig, um Signifikanzniveau zu erreichen. Unter Berücksichtigung der Erfolgswahrscheinlichkeiten für die Vorprüfungen kann eine Benachteiligung der Kandidatinnen durch den Eignungstest nach wie vor ausgeschlossen werden.

		Humanmedizin			Veterinärmedizin			Zahnmedizin		
		Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch
männlich	m	102.34	94.86	94.33	99.41	102.00		96.67	83.00	95.60
	s	9.692	11.246	9.618	6.384	9.434		11.598		10.383
	n	337	7	9	17	7		55	1	5
weiblich	m	100.21	97.48	96.77	100.09	93.89	94.29	97.85		87.33
	s	9.743	10.658	10.882	9.167	8.121	5.187	11.054		15.822
	n	622	25	26	137	28	7	87		3

Tabelle 29: Statistiken für den Testwert nach Geschlechtern und Disziplin

Sprache	Disziplin	Levene's Test Gleichheit der Varianzen			t-Test Gleichheit des Mittelwerts		
			F	Sig.	t	df	Sig.
deutsch	Humanmedizin	Varianzen homogen	.029	.864	3.237	957	.001
	Veterinärmedizin	Varianzen homogen	3.343	.069	-.298	152	.766
	Zahnmedizin	Varianzen homogen	1.068	.303	-.607	140	.545
französisch	Humanmedizin	Varianzen homogen	.013	.909	-.569	30	.574
	Veterinärmedizin	Varianzen homogen	.256	.616	2.291	33	.029
italienisch	Humanmedizin	Varianzen homogen	.182	.672	-.595	33	.556

Tabelle 30: Prüfung Varianzhomogenität und Mittelwert (t-Test) für Testwert zwischen Geschlechtern. Italienische Sprachgruppe Veterinärmedizin wg. Fallzahl nicht berechnet

Abbildung 38 zeigt den Vergleich zwischen männlichen Bewerbern und Kandidatinnen, aufgeteilt nach Disziplin, Abbildung 39 den Boxplot für die Testwerte nach Sprache und Geschlecht geschichtet.

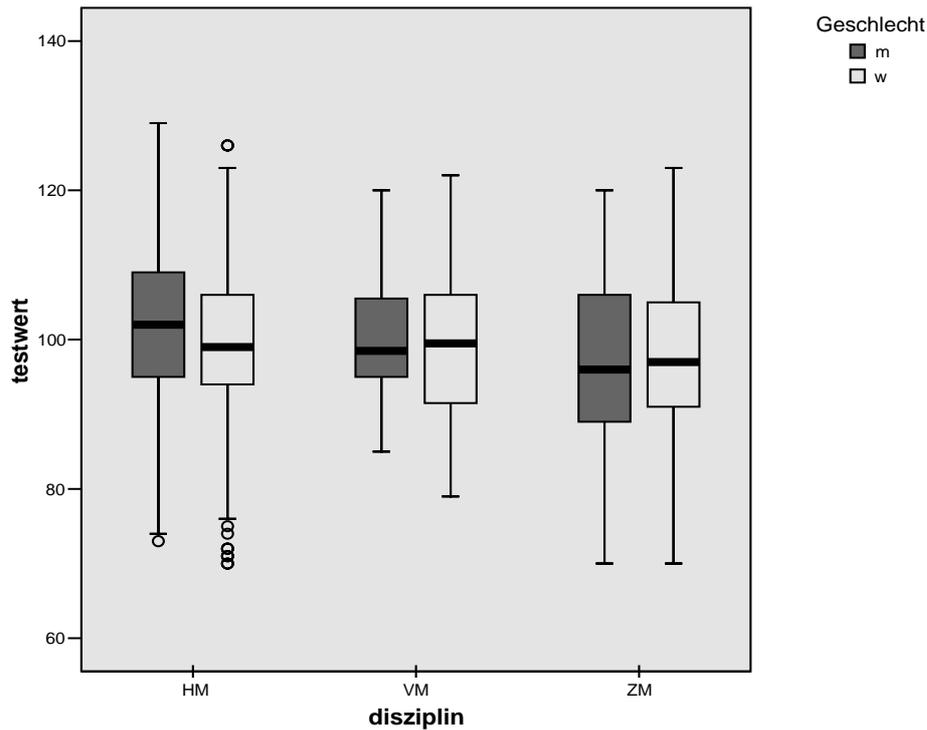


Abbildung 38: Boxplot für den Testwert nach Geschlecht und Disziplin (Humanmedizin und Veterinärmedizin)

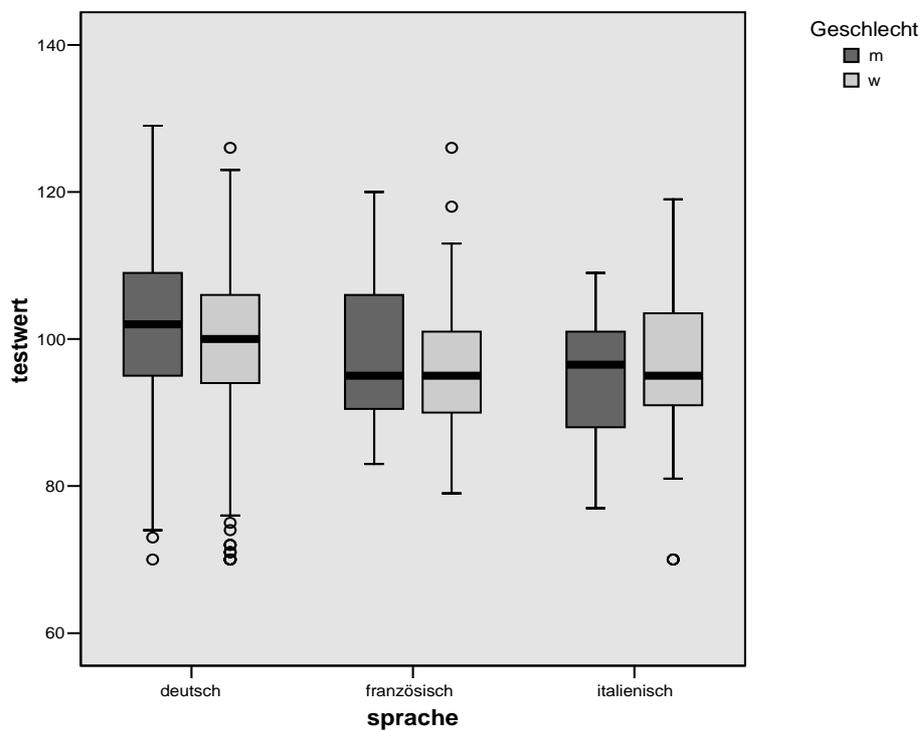


Abbildung 39: Boxplots für die Testwerte nach Geschlecht, getrennt für die drei Sprachgruppen

Untertest	Geschlecht	m	s	Levene's Test Varianz			EG	t-test Mittelwert		
					F	Sig.		t	df	Sig.
Muster zuordnen	männlich	10.41	2.92	homogen	0.00	0.974	0.01	.143	1371	.887
	weiblich	10.39	2.84							
Med.-naturwiss. Grundverst.	männlich	10.79	3.45	homogen	2.00	0.157	0.19	3.240	1371	.001
	weiblich	10.16	3.30							
Schlauchfiguren	männlich	13.17	3.62	homogen	0.16	0.685	0.36	6.204	1371	.000
	weiblich	11.85	3.67							
Quantitative und formale Probleme	männlich	11.24	3.63	homogen	3.62	0.057	0.57	9.945	1371	.000
	weiblich	9.24	3.41							
Textverständnis	männlich	9.05	3.39	homogen	1.54	0.215	0.09	1.480	1371	.139
	weiblich	8.77	3.24							
Figuren lernen	männlich	12.45	3.76	homogen	3.83	0.051	0.18	-3.107	1371	.002
	weiblich	13.11	3.59							
Fakten lernen	männlich	11.53	3.81	homogen	0.40	0.526	0.30	-5.204	1371	.000
	weiblich	12.64	3.63							
Diagramme und Tabellen	männlich	11.28	3.37	homogen	1.13	0.288	0.32	5.592	1371	.000
	weiblich	10.21	3.27							
Konzent. u. sorgfältiges Arbeiten	männlich	11.16	4.41	homogen	0.40	0.529	0.11	-1.988	1371	.047
	weiblich	11.66	4.23							

Tabelle 31: Mittelwerte (m) und Standardabweichungen (s) geschlechtsspezifisch und Ergebnisse der Prüfungen auf Varianzhomogenität und Mittelwertunterschiede mit Effektgröße (EG) (Näherung wenn keine Varianzhomogenität)

Die Geschlechtsunterschiede auf Untertestebene bleiben auch dieses Jahr konstant. Besonderes Augenmerk muss jeweils auf den neu konzipierten Untertest „Konzentriertes und Sorgfältiges Arbeiten“ gelegt werden. Auch hier zeigt sich eine zumindest mit den letzten drei Jahren vergleichbare Differenz.

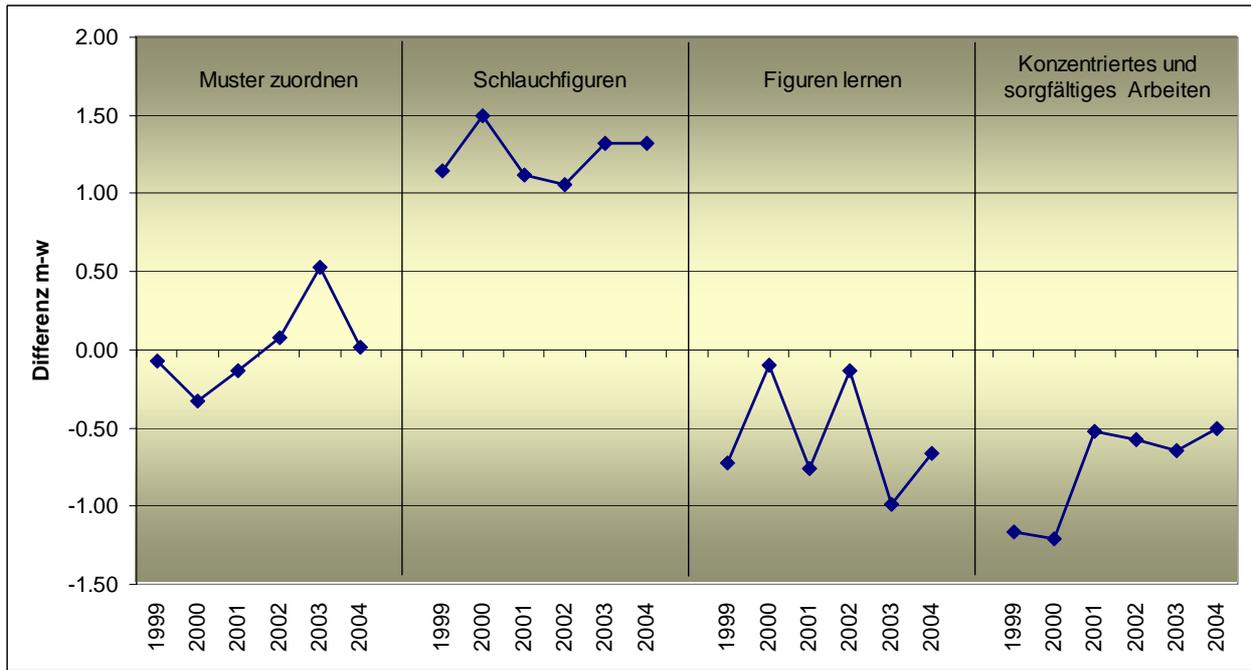


Abbildung 40: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (wenig sprachabhängige Untertests)

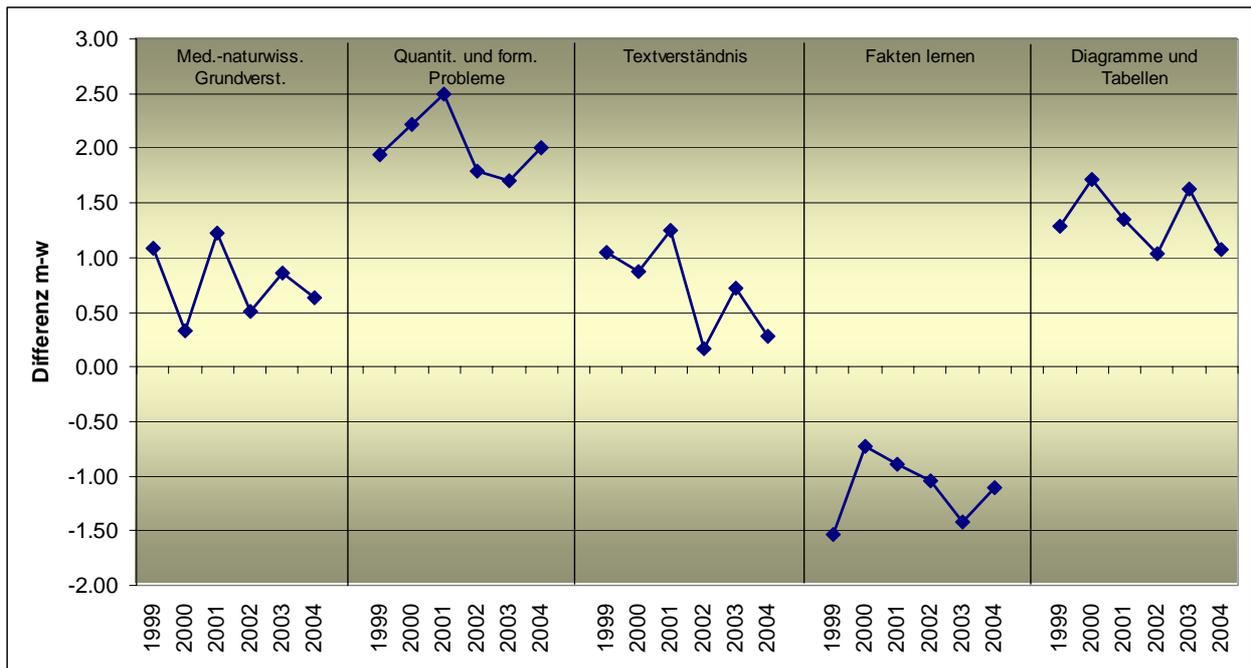


Abbildung 41: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (sprachabhängige Untertests)

Bezüglich der Zulassung sind in beiden Disziplinen signifikante Unterschiede zwischen den Geschlechtern festzustellen (Tabelle 33). Bei den Männern beträgt sie 65.4%, respektive 91.7%, bei den Frauen 55.3%, respektive 73.3% (Tabelle 32). In der Disziplin „Zahnmedizin“ beträgt die Zulassung 100%.

Disziplin	Geschlecht	Warteliste	zugelassen	Total
Humanmedizin	männlich	122 34.6%	231 65.4%	353 100.0%
	weiblich	301 44.7%	372 55.3%	673 100.0%
	Total	423 41.2%	603 58.8%	1026 100.0%
Veterinärmedizin	männlich	2 8.3%	22 91.7%	24 100.0%
	weiblich	46 26.7%	126 73.3%	172 100.0%
	Total	48 24.5%	148 75.5%	196 100.0%

Tabelle 32: Zulassungsquoten geschlechtsspezifisch

Disziplin		Wert	df	Asymp. Sig. (2-seitig)
Humanmedizin	Pearson Chiquadrat	9.872	1	.002
	Likelihood Ratio	9.975	1	.002
Veterinärmedizin	Pearson Chiquadrat	3.860	1	.049
	Likelihood Ratio	4.683	1	.030

Tabelle 33: Prüfung Unterschiede der Zulassung – geschlechtsspezifisch

7.7 Vergleich der Kantone

Weiterhin besteht ein grosses Interesse an Ergebnissen, die sich auf die Vergleiche zwischen den Kantonen beziehen. Wiederum ist vor unzulässigen Generalisierungen vorab zu warnen: Es handelt sich nicht um eine Repräsentativerhebung für „die Altersgruppe“ oder „die Maturanden“. Medizinbewerbungen werden nach verschiedenen Gesichtspunkten erfolgen und sich zwischen Kantonen auch möglicherweise unterscheiden.

Dennoch können die Ergebnisse Hinweise darauf liefern, ob sich bestimmte systematische Unterschiede zeigen. Die Interpretation muss mit der gebotenen Vorsicht für den jeweiligen Einzelfall vorgenommen werden.

7.7.1 Vergleich 2004 innerhalb der Deutschschweiz

hoch			mittel			niedrig		
Kanton	Maturitätsquote	Anzahl	Kanton	Maturitätsquote	Anzahl	Kanton	Maturitätsquote	Anzahl
BS	24.2%	56	ZG	16.6%	22	SG	13.9%	73
ZH	18.4%	325	AG	16.5%	118	NW	13.9%	7
BL	18.3%	76	SH	16.1%	10	SO	13.9%	58
			TG	15.3%	35	LU	13.4%	66
			GL	14.7%	1	OW	13.1%	7
			GR	14.5%	59	BE	13.0%	238
			AI	14.3%	5	UR	11.7%	3
			SZ	14.1%	24			
Summe		457	Summe		274	Summe		452

Tabelle 34: Gruppierung der Kantone nach der Maturitätsquote (Basis 2002)

Die Einteilung der Kantone nach „hohen“, „mittleren“ und „niedrigen“ Maturitätsquoten beruht auf den Quoten aus dem Jahr 2002 (aktuell verfügbare Zahlen). Gegenüber früheren Jahren ergaben sich in der Kantonszuordnung einige Unterschiede. Nach wie vor kann jedoch kein systematischer Unterschied bezüglich des Testwerts beobachtet werden, welcher in Zusammenhang mit der kantonalen Maturitätsquote stehen würde. Einzig im Untertest „Quantitative und formale Probleme“ zeigt sich ein signifikantes Ergebnis, wobei hier entgegen der Erwartung die Kantone mit niedriger Maturitätsquote schlechter abschneiden als die beiden anderen Gruppen (Tabelle 35).

				Zwischen Gruppen				
	Maturitäts- quote	Mittel- wert	Stand- abw.	Quadrat- summe	df	MQ	F	Sig.
Testwert	hoch	100.87	10.37	353.242	2	176.621	1.827	0.161
	mittel	101.20	9.80					
	niedrig	99.90	9.28					
Muster zuordnen	hoch	10.57	2.91	4.351	2	2.176	0.266	0.766
	mittel	10.41	2.95					
	niedrig	10.49	2.74					
Medizin.-natur- wiss. Grundver- ständnis	hoch	10.40	3.39	18.539	2	9.270	0.839	0.433
	mittel	10.65	3.27					
	niedrig	10.33	3.29					
Schlauchfigure n	hoch	12.78	3.73	79.352	2	39.676	2.998	0.050
	mittel	12.50	3.47					
	niedrig	12.19	3.64					
Quantitat. u. formale Probleme	hoch	10.09	3.71	130.274	2	65.137	5.174	0.006
	mittel	10.39	3.54					
	niedrig	9.56	3.38					
Textverständnis	hoch	8.94	3.34	2.416	2	1.208	0.112	0.894
	mittel	9.00	3.44					
	niedrig	8.88	3.12					
Figuren lernen	hoch	13.13	3.65	5.604	2	2.802	0.213	0.808
	mittel	13.19	3.63					
	niedrig	13.02	3.60					
Fakten lernen	hoch	12.59	3.87	19.500	2	9.750	0.709	0.493
	mittel	12.27	3.64					
	niedrig	12.38	3.58					
Diagramme und Tabellen	hoch	10.75	3.38	11.561	2	5.780	0.526	0.591
	mittel	10.81	3.22					
	niedrig	10.57	3.30					
Konzentr. u. sorgf. Arb.	hoch	11.46	4.66	103.278	2	51.639	2.809	0.061
	mittel	12.08	3.93					
	niedrig	11.33	4.10					

Tabelle 35: Mittelwerte und Standardabweichungen für drei Gruppen der Kantone nach der Maturitätsquote

Kanton		Testwert	Musterzuordnen	Medizin.-naturw. Grundv.	Schlauchfiguren	Quantit. u. form. Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arb.
AG	M	103.92	10.64	11.49	13.18	11.21	9.64	13.58	13.08	11.22	12.46
	S	9.22	2.61	2.96	3.30	3.75	3.54	3.52	3.69	3.23	3.80
BE	M	99.61	10.32	10.39	12.24	9.21	9.00	12.99	12.35	10.55	11.11
	S	9.31	2.68	3.37	3.62	3.32	3.15	3.53	3.64	3.36	4.32
BL	M	100.74	10.33	10.71	12.71	9.34	9.17	13.59	12.45	10.71	11.37
	S	10.52	2.98	3.33	3.53	3.44	3.22	3.35	4.08	3.51	4.47
BS	M	97.63	10.11	9.14	11.88	8.54	8.16	13.05	11.52	9.89	12.32
	S	9.92	2.95	3.29	3.22	3.52	3.01	3.50	3.98	3.18	4.73
LU	M	99.92	10.44	10.88	11.58	10.00	9.56	12.52	12.12	10.47	11.35
	S	9.23	2.79	3.35	3.31	3.31	3.31	4.07	3.16	3.35	3.70
SG	M	101.96	10.64	10.34	12.75	10.42	9.11	13.55	12.95	11.22	11.78
	S	9.09	2.70	3.21	4.16	3.28	2.72	3.48	3.34	3.33	3.66
ZH	M	101.46	10.70	10.54	12.96	10.53	9.02	13.04	12.80	10.90	11.33
	S	10.34	2.89	3.38	3.84	3.72	3.42	3.75	3.78	3.37	4.69

Tabelle 36: Vergleich der Kantone für Testwert und die einzelnen Untertests

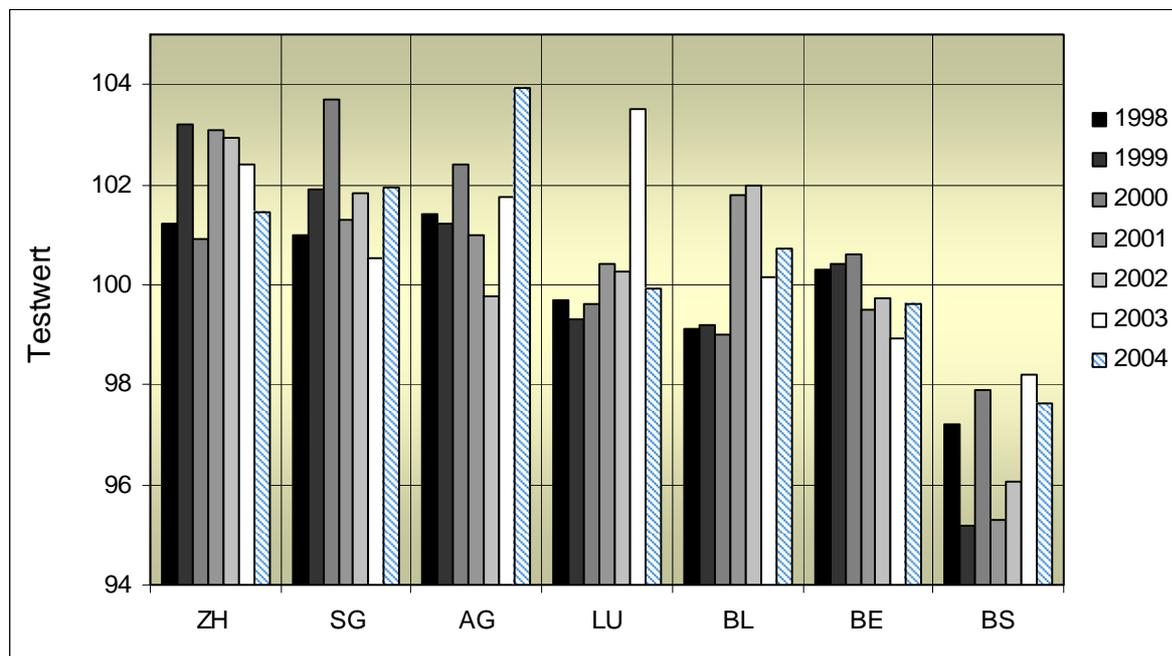


Abbildung 42: Testwert für Kantone im Vergleich

Tabelle 36 zeigt Testwerte und Untertestpunktwerte der sieben Kantone mit den grössten Teilnehmerzahlen. Abbildung 42 stellt die Testwerte grafisch dar. Der Kanton Aargau erreichte erstmals die höchsten mittleren Testwerte. Bei den Teilnehmern aus dem Kanton Zürich ist zum dritten Mal in Folge ein Rückgang der durchschnittlichen Testwerte zu beobachten. Luzern fiel nach einem Spitzenjahrgang auf das gewohnte Niveau zurück.

Besondere Beachtung findet jeweils das Abschneiden von Basel-Stadt. Eine nähere Betrachtung der Untertestergebnisse (Abbildung 43) macht deutlich, dass das tiefe Gesamtergebnis in erster Linie auf die sprachabhängigen Untertests zurückzuführen ist. In den weniger sprachabhängigen Untertests liegen die Ergebnisse im durchschnittlichen Bereich, im Untertest „Konzentriertes und Sorgfältiges Arbeiten“ wurde in Basel-Stadt sogar das zweitbeste Ergebnis erzielt.

Diese Abbildung zeigt auch, dass die Unterschiede zwischen den Kantonen insgesamt mehr auf Unterschiede in sprachabhängigen Tests zurückzuführen sind als auf jene in wenig sprachabhängigen Tests. Auch dafür sind verschiedene Ursachen denkbar. Unterschiede in der „Sprachkompetenz“ können sowohl auf Bildungsunterschiede, als auch auf die Vertrautheit mit der deutschen Sprache insgesamt zurückzuführen sein. Dieser Hinweis kann allerdings auch nur Anregung sein, über diese Unterschiede weiter nachzudenken, ohne die Ursachen pauschal in unterschiedlich effektiven Schulsystemen zu suchen.

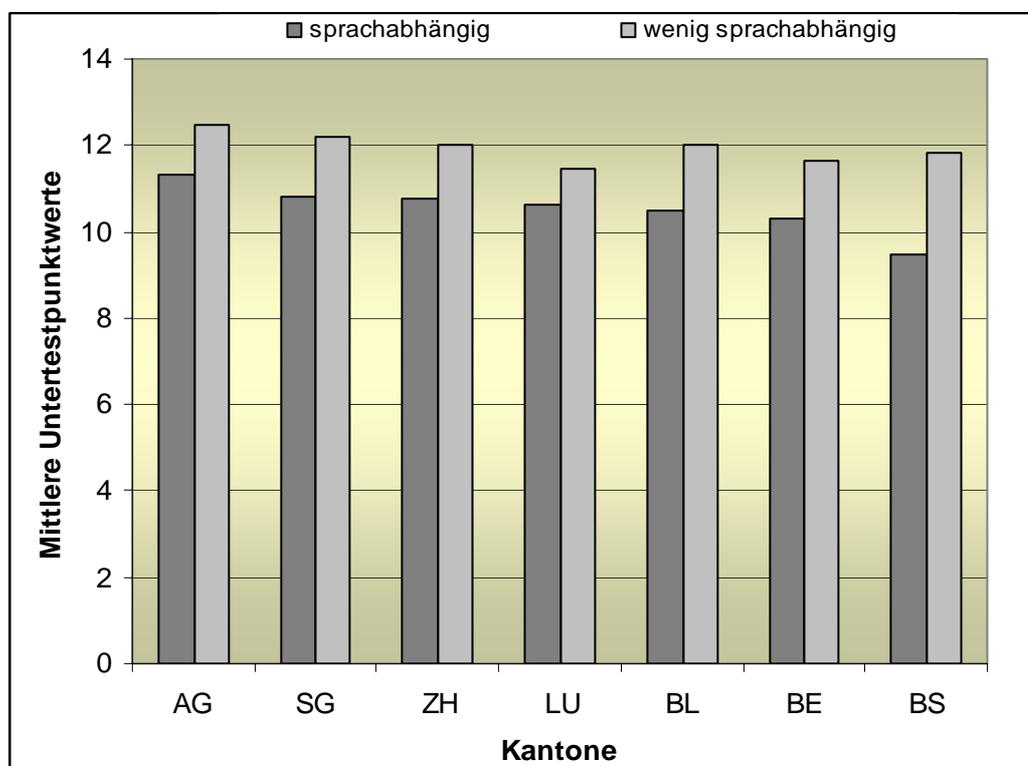


Abbildung 43: Mittelwerte sprachabhängiger und weniger sprachabhängiger Untertests 2004 für die Kantone im Vergleich

Abbildung 44 stellt den Verlauf über die Jahre anhand der aggregierten Mittelwerte dar. Deutlich sichtbar werden die drei Gruppen ZH-SG-AG / LU-BL-BE / BS, wobei es innerhalb der Gruppen zu Rangverschiebungen kommt, nicht aber dazwischen.

Mit dem diesjährigen Ergebnis konnte Aargau zu St. Gallen aufschliessen, welches 2000 die besten Mittelwerte erreichte. Luzern und Basel-Land hatten zwischenzeitlich die Plätze getauscht. Basel-Stadt liegt deutlich hinter den anderen Kantonen zurück, allerdings konnten die Leistungen gegenüber den schwächsten Jahrgängen 1999 und 2001 deutlich verbessert werden.

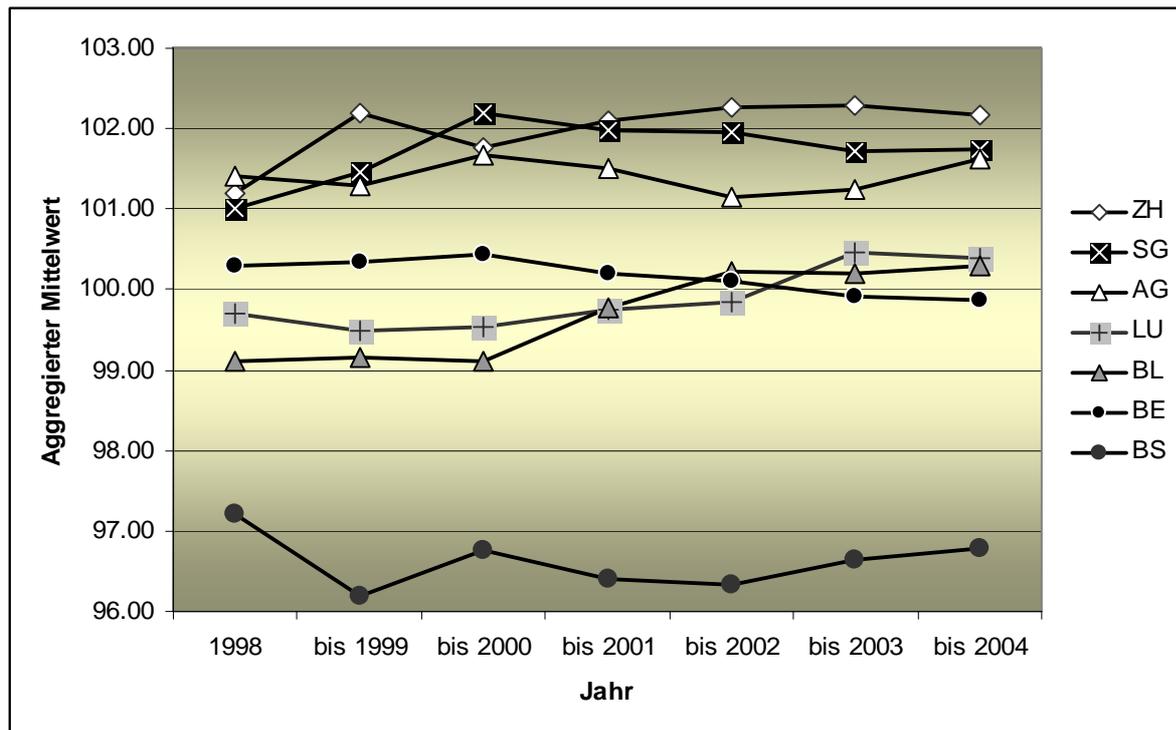


Abbildung 44: Aggregierte Testwerte für Kantone im Vergleich 1998 - 2004

7.8 Vergleiche für Altersgruppen

Für die Überprüfung der Alterseffekte im Jahr 2004 wurden die Gruppen auf der Basis der Tabelle 37 gebildet. Die älteste Gruppe wurde nahe dem Median des Maturitätsalters geteilt.

Maturitätsjahr	Geburtsjahr			Total
	1959-1979	1980-1982	1983-1987	
1974	1			1
1976	2			2
1977	1			1
1981	1			1
1982	2			2
1983	1			1
1984	1			1
1985	1			1
1986	2			2
1987	5			5
1991	2			2
1992	2			2
1993	0			0
1995	8			8
1996	4			4
1997	6			6
1998	7			7
1999	12	2		14
2000	8	6		14
2001	4	30		34
2002	8	78	15	101
2003	17	116	247	380
2004	51	77	653	781
	146	309	915	1370*

* 3 Personen mit fehlenden Angaben

Tabelle 37: Gruppenbildung für das Jahr der Maturitätsprüfung bezogen auf die Geburtsjahre

Erneut zeigen die älteren Jahrgänge mit später Maturität niedrigere Testwerte als alle anderen Gruppen (Tabelle 38 und Abbildung 45). Die älteren Jahrgänge mit früher Maturität unterscheiden sich hingegen statistisch nicht signifikant von den anderen Gruppen. Die besten Ergebnisse sind sowohl bezüglich des Testwerts wie auch der einzelnen Untertests bei den jüngsten Teilnehmern zu finden.

	Geburtsjahr und Maturität	Mittelwert	Standard-Abw.	Quadrat-summe	df	MQ	F	Sig.	Homogene Gruppen
Testwert	1951-1980, frühe Maturität	98.22	12.06	3531.61	3	1177.20	12.08	0.00	
	1951-1980, späte Maturität	95.05	9.84						
	1981-1983	98.84	10.05						
	1984-1988	101.02	9.64						
Muster zuordnen	1951-1980, frühe Maturität	9.40	2.97	137.32	3	45.77	5.63	0.00	
	1951-1980, späte Maturität	9.65	2.49						
	1981-1983	10.30	3.14						
	1984-1988	10.56	2.77						
Med.-naturwiss. Grundverständnis	1951-1980, frühe Maturität	9.69	3.41	381.15	3	127.05	11.54	0.00	
	1951-1980, späte Maturität	8.81	3.42						
	1981-1983	9.94	3.30						
	1984-1988	10.70	3.31						
Schlauchfiguren	1951-1980, frühe Maturität	11.75	3.82	73.01	3	24.34	1.77	0.15	
	1951-1980, späte Maturität	12.26	3.51						
	1981-1983	12.66	3.88						
	1984-1988	12.18	3.65						
Quantitative und formale Probleme	1951-1980, frühe Maturität	10.88	4.06	294.88	3	98.29	7.69	0.00	
	1951-1980, späte Maturität	8.73	3.48						
	1981-1983	9.37	3.53						
	1984-1988	10.09	3.56						
Textverständnis	1951-1980, frühe Maturität	9.15	3.30	272.33	3	90.78	8.50	0.00	
	1951-1980, späte Maturität	7.48	3.02						
	1981-1983	8.43	3.26						
	1984-1988	9.11	3.29						
Figuren lernen	1951-1980, frühe Maturität	12.16	4.89	185.77	3	61.92	4.67	0.00	
	1951-1980, späte Maturität	11.73	3.85						
	1981-1983	12.78	3.55						
	1984-1988	13.10	3.55						
Fakten lernen	1951-1980, frühe Maturität	11.49	4.26	156.10	3	52.03	3.78	0.01	
	1951-1980, späte Maturität	11.31	3.89						
	1981-1983	12.16	3.77						
	1984-1988	12.48	3.63						
Diagramme und Tabellen	1951-1980, frühe Maturität	10.70	3.76	282.23	3	94.08	8.57	0.00	
	1951-1980, späte Maturität	9.33	3.56						
	1981-1983	10.02	3.39						
	1984-1988	10.83	3.23						
Konzentriertes und Sorgfältiges Arbeiten	1951-1980, frühe Maturität	10.24	5.05	428.54	3	142.85	7.86	0.00	
	1951-1980, späte Maturität	10.08	4.93						
	1981-1983	11.10	4.51						
	1984-1988	11.85	4.05						

Tabelle 38: Testwert und Punktwerte für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung; Varianzanalytische Prüfung des Unterschiedes

In Tabelle 38 sind in den letzten Spalten ausserdem die bezüglich des Mittelwertes homogenen Gruppen dargestellt. Schattierungen in der gleichen **Spalte** bedeuten, dass sich die so markierten Gruppen NICHT voneinander unterscheiden. Bezüglich des Testwerts bedeutet dies beispielsweise, dass sich die ältesten Jahrgänge mit später Maturität signifikant von allen anderen Altersgruppen unterscheiden. Die Testwerte der Jahrgänge 1951-1980 mit früher Maturität unterscheiden sich nicht von jenen der Jahrgänge 1981-1983. Ebenso ist zwischen den Jahrgängen 1981-1983 und 1984-1988 kein Unterschied zu beobachten.

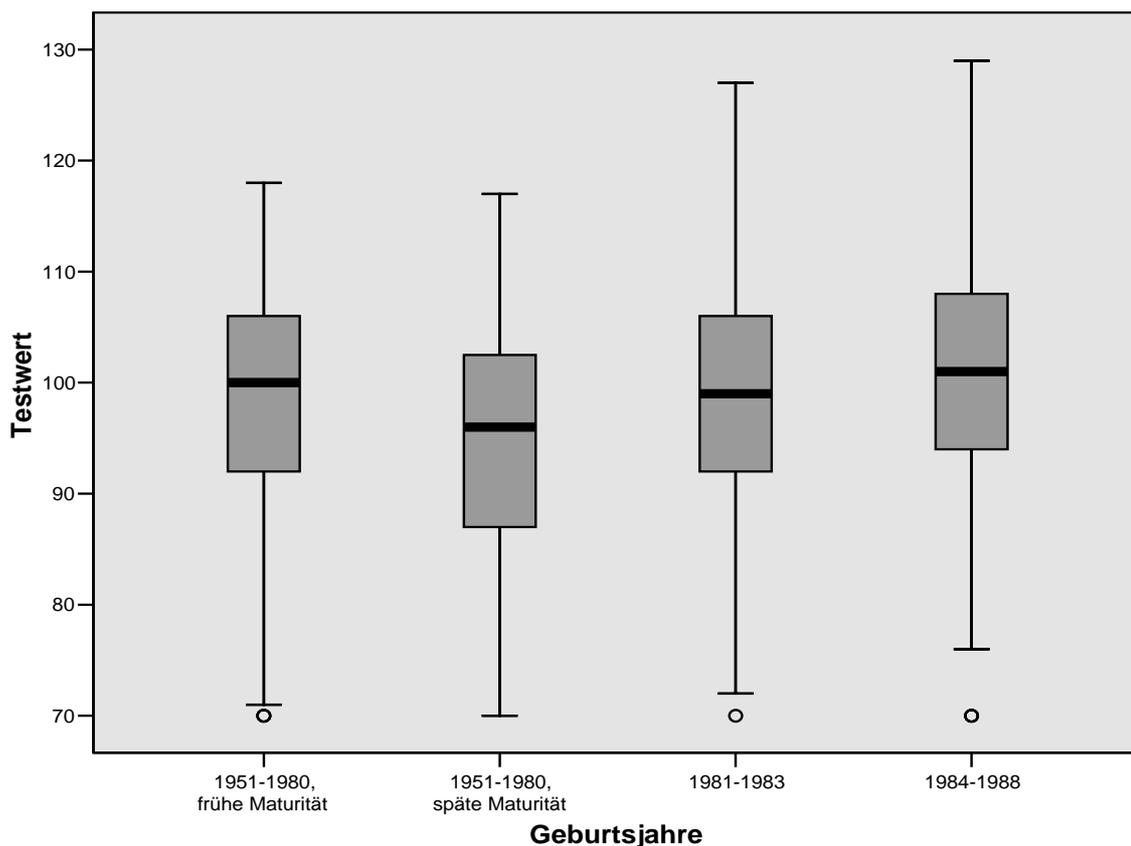


Abbildung 45: Boxplots für Testwert der Altersgruppen (Geburtsjahre) und Maturität

Geburtsjahr und Maturität	Warteliste	zugelassen	n
1951-1980, frühe Maturität	19 28.4%	48 71.6%	67 100.0%
1951-1980, späte Maturität	39 48.8%	41 51.3%	80 100.0%
1981-1983	112 36.2%	197 63.8%	309 100.0%
1984-1988	300 32.8%	615 67.2%	915 100.0%
Total	470 34.3%	901 65.7%	1371* 100.0%

*2 Personen mit fehlenden Angaben

Tabelle 39: Zulassungsquoten für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung

Hinsichtlich der Zulassungsquote (Tabelle 39) schneiden die ältesten Jahrgänge mit früher Maturität besser ab als alle anderen Gruppen. Unter diesen Teilnehmern sind keine sehr hohen Testwerte, aber auch nur wenige niedrige Testwerte zu finden.

7.9 Vergleiche nach Wunschuniversitäten

Die Zulassung zu den Universitäten erfolgt aufgrund des Testwerts. Je höher der Testwert, desto grösser ist die Wahrscheinlichkeit, zu der eigenen „Wunschuniversität“ zugelassen zu werden.

Die Zulassungsquoten zur Humanmedizin unterscheiden sich zum zweiten Mal in Folge signifikant, wenn auch in geringerem Ausmass als dies 2003 der Fall war (Effektgrösse = 0.97). Die höchste Zulassungsquote wurde wiederum unter jenen Personen erreicht, die Zürich als ihre Wunschuniversität angegeben hatten.

Disziplinen	Wunsch-Universität	Zulassung		Total
		Warteliste	zugelassen	
Humanmedizin	Bern	106	166	272
		39.0%	61.0%	100%
	Basel	108	132	240
		45.0%	55.0%	100%
	Freiburg	43	38	81
	53.1%	46.9%	100%	
Zürich	165	267	432	
	38.2%	61.8%	100%	
Total	423*	603	1026	
	41.2%	58.8%	100%	
Veterinärmedizin	Bern	30	75	105
		28.6%	71.4%	100%
	Zürich	18	73	91
		19.8%	80.2%	100%
	Total	48	148	196
24.5%		75.5%	100%	

* 1 Anmeldung für Neuenburg

Tabelle 40: Zulassungsquoten für Disziplin und Wunschuniversität (erste Wahl)

	Value	df	Asymp. Sig.
Pearson Chi-Square	9.75	4	.045
Likelihood Ratio	10.01	4	.040
Cramer's V	.097		.045

Tabelle 41: Signifikanz für Zulassungsquoten nach Wunschuniversität (HM)

Bezüglich der Testwerte erreichen nur die Unterschiede zwischen Freiburg und Bern/Zürich statistisch signifikantes Ausmass. Die Ergebnisse der Bewerber für Basel unterscheiden sich nicht von allen anderen Gruppen.

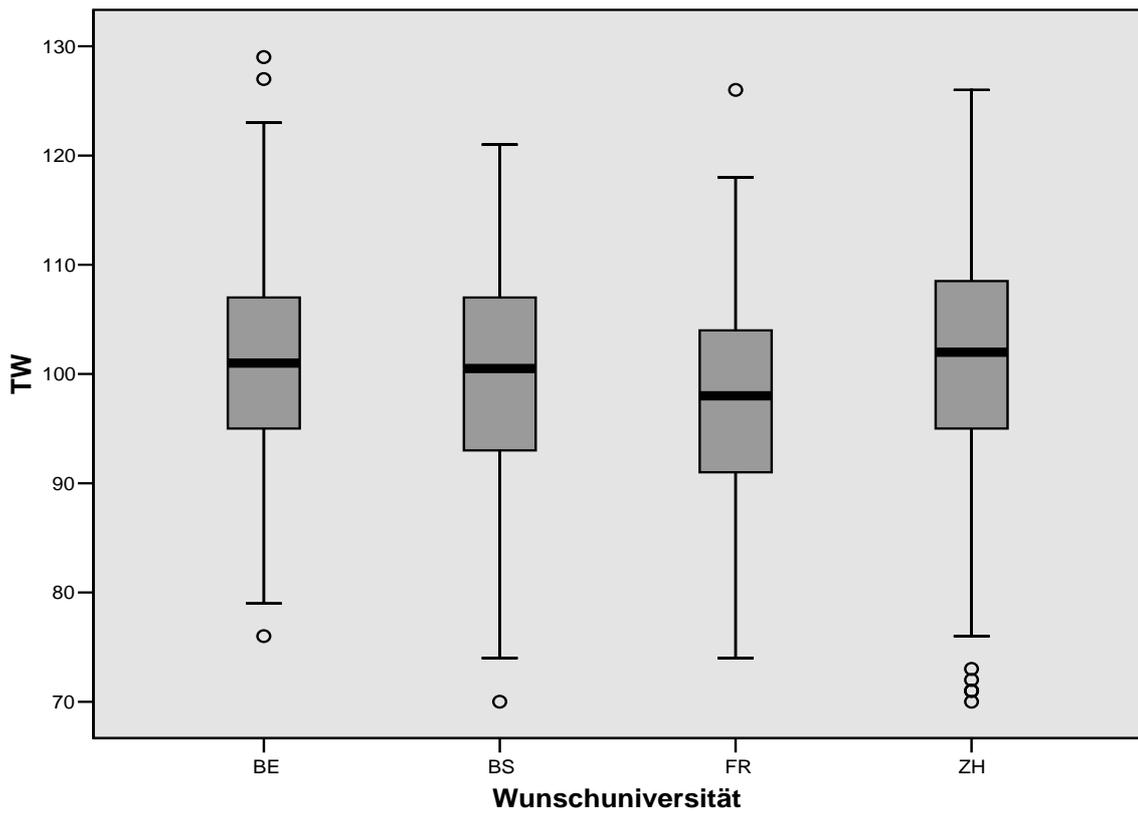


Abbildung 46: Boxplots für Testwerte aller Teilnehmer nach Wunschuniversitäten

Wunsch- universität	Teilnehmer gesamt			zugelassen		
	N	Mittel- wert	Standard- abw.	N	Mittel- wert	Standard- abw.
Bern	272	100.91	9.41	166	106.70	6.50
Basel	240	100.01	9.78	132	107.12	5.62
Freiburg	81	98.35	10.15	38	106.92	6.38
Zürich	432	101.38	10.03	267	107.57	6.25
Total	1025*	100.67	9.88	603	107.19	6.20

*1 Bewerber für Neuenburg

Tabelle 42: Testwerte nach Wunschuniversitäten, alle Teilnehmer und Zugelassenen (HM)

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	81.08	3	27.03	0.7	0.55
Innerhalb der Gruppen	23031.37	599	38.45		
Gesamt	23112.45	602			

Tabelle 43: Signifikanzprüfung nach Wunschuniversitäten der zugelassenen Personen (HM)

Betrachtet man nur die zum Studium zugelassenen Personen, so treten zwischen den Universitäten keine signifikanten Unterschiede mehr auf (Tabelle 42 und Tabelle 43).

8 Ergebnisse zur Testgüte

8.1 Zuverlässigkeit

Die Reliabilität kann anhand zweier Koeffizienten verglichen werden. Die internen Konsistenzen (Cronbachs Alpha) schätzen die Messgenauigkeit anhand der Korrelationen jeder Aufgabe mit allen anderen des entsprechenden Untertests. Wie sich zeigt, liegen weiterhin einige der in der Schweiz ermittelten Werte unterhalb der deutschen Vergleichsangaben, insbesondere der Untertest „Muster zuordnen“ erreicht seit dem Jahrgang 2000 nicht mehr die früheren Werte. Eine Erklärung hierzu konnte bisher nicht gefunden werden. Die Qualität des gesamten Tests ist jedoch durch dieses Ergebnis nicht in Frage gestellt, weil die Gesamtreliabilität nicht beeinträchtigt wird.

Untertest	Innere Konsistenz (Cronbachs Alpha)							
	Deutschland	Schweiz (d)						
		1998	1999	2000	2001	2002	2003	2004
Muster zuordnen	.64 - .69	.69	.60	.53	.56	.57	.53	.54
Med.-nat. Grundv.	.64 - .73	.70	.70	.61	.66	.57	.72	.66
Schlauchfiguren	.76 - .80	.71	.74	.71	.73	.68	.74	.72
Quant. u. form. Prob.	.69 - .71	.76	.75	.76	.74	.71	.67	.71
Textverständnis	.68 - .71	.73	.75	.74	.70	.68	.78	.67
Figuren lernen	.69 - .74	.72	.70	.62	.64	.82	.77	.71
Fakten lernen	.62 - .70	.68	.72	.64	.69	.70	.73	.72
Diagr. und Tab.	.67 - .71	.67	.75	.63	.61	.62	.70	.66

Tabelle 44: Zuverlässigkeit 1998 bis 2004; Vergleich der Skalenkonsistenzen für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

Untertest	Reliabilität nach Testhalbierungsmethode							
	Deutschland	Schweiz (d)						
		1998	1999	2000	2001	2002	2003	2004
Muster zuordnen	.67 - .73	.72	.61	.62	.59	.56	.55	.57
Med.-nat. Grundv.	.65 - .72	.72	.71	.60	.69	.57	.73	.67
Schlauchfiguren	.75 - .82	.75	.74	.74	.72	.67	.76	.74
Quant. u. form. Prob.	.70 - .75	.78	.75	.74	.74	.71	.68	.72
Textverständnis	.69 - .74	.75	.70	.75	.75	.71	.78	.70
Figuren lernen	.72 - .75	.72	.70	.62	.62	.81	.78	.73
Fakten lernen	.68 - .73	.70	.70	.61	.72	.69	.64	.72
Diagr. und Tab.	.68 - .72	.66	.62	.61	.63	.64	.62	.68
Gesamter Test	.91 - .93	.92	.91	.90	.91	.91	.92	.91

Tabelle 45: Zuverlässigkeit 1998 bis 2004; Vergleich der Reliabilitäten nach der Testhalbierungsmethode für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

Die Reliabilitäten nach der Testhalbierungsmethode (Teilung nach gerad- und ungeradzahligen Aufgaben) wurden ebenfalls verglichen (Tabelle 45). Wegen der Stichprobengrößen werden nur die Ergebnisse des deutschsprachigen Tests der Schweiz verwendet.

Die **Reliabilität des Punktwertes** insgesamt – geschätzt nach der Testhalbierungsmethode (gerade vs. ungerade Itemnummern) auf der Basis aller neun Untertests liegt mit 0.91 wie gewohnt im sehr hohen Bereich.

Ein Weglassen einzelner Untertests würde nicht zu einer Verbesserung der Gesamtkonsistenz des Testprofils führen. Diese ist mit 0.78 etwas niedriger als in früheren Jahren, aber immer noch ausreichend hoch.

Untertest	Kons. ohne den Untertest 1998	Kons. ohne den Untertest 1999	Kons. ohne den Untertest 2000	Kons. ohne den Untertest 2001	Kons. ohne den Untertest 2002	Kons. ohne den Untertest 2003	Kons. ohne den Untertest 2004
Muster zuordnen	.82	.80	.80	.78	.79	.80	.77
Med.-nat. Grundverständnis	.79	.78	.78	.77	.77	.78	.75
Schlauchfiguren	.80	.79	.79	.78	.78	.79	.75
Quant. u. form. Probleme	.80	.79	.78	.78	.77	.78	.75
Textverständnis	.80	.79	.78	.77	.78	.77	.75
Figuren lernen	.81	.80	.80	.78	.78	.79	.76
Fakten lernen	.82	.81	.80	.79	.79	.80	.77
Diagramme und Tabellen	.80	.78	.78	.77	.77	.78	.74
Konz. u. sorgf. Arbeiten	.82	.81	.80	.79	.79	.80	.78
TESTPROFIL	.83	.81	.81	.80	.80	.81	.78

Tabelle 46: Innere Konsistenzen, wenn der jeweilige Untertest weggelassen worden wäre und Konsistenz des Testprofils („standardized alpha“)

8.2 Binnenstruktur

Aufgrund der Korrelationen zwischen den Untertests kann mittels Strukturanalyse geprüft werden, ob sich die einzelnen Untertests bestimmten Dimensionen zuordnen lassen. Die so gewonnene Struktur bietet Vergleichsmöglichkeiten mit theoretischen Vorstellungen zum Fähigkeitsbereich. Die Gruppierung der Untertestleistungen kann zu den bekannten Faktormodellen der Intelligenz in Beziehung gesetzt werden.

Von besonderem Interesse dürfte dieses Jahr die Einordnung des neu konzipierten Untertests „Konzentriertes und Sorgfältiges Arbeiten“ sein.

Die Korrelationen der Untertests mit dem Punktwert sind hoch. Ausnahme ist weiterhin der Untertest „Muster zuordnen“, wo der Zusammenhang wie 2003 unter 0.50 fällt (Tabelle 47). Alle Korrelationen mit dem Punktwert sind auf dem 1%-Niveau signifikant.

	Muster zuordnen	Med.-naturwiss. Grundv.	Schlauchfiguren	Quantitat. und formale Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arbeiten	Punktwert
Muster zuordnen		0.14	0.38	0.18	0.11	0.31	0.23	0.20	0.22	0.49
Med.-naturw. Grundverst.	0.14		0.27	0.48	0.60	0.21	0.23	0.51	0.19	0.66
Schlauchfiguren	0.38	0.27		0.34	0.24	0.38	0.33	0.28	0.16	0.62
Quant. u. formale Probl.	0.18	0.48	0.34		0.43	0.19	0.18	0.58	0.24	0.66
Textverständnis	0.11	0.60	0.24	0.43		0.23	0.23	0.52	0.21	0.65
Figuren lernen	0.31	0.21	0.38	0.19	0.23		0.42	0.23	0.19	0.59
Fakten lernen	0.23	0.23	0.33	0.18	0.23	0.42		0.21	0.17	0.56
Diagramme und Tabellen	0.20	0.51	0.28	0.58	0.52	0.23	0.21		0.28	0.69
Konzent. u. sorgfält. Arbeiten	0.22	0.19	0.16	0.24	0.21	0.19	0.17	0.28		0.52

Tabelle 47: Korrelationen zwischen Punktwerten und Untertests

Abbildung 47 zeigt, dass die Korrelationen mit der Gesamtpunktzahl für die einzelnen Untertests über die Jahre mehr oder weniger stabil bleiben. Der Untertest „Muster zuordnen“ hat 2003 ein geringeres Gewicht erhalten, was sich in diesem Jahr bestätigt.

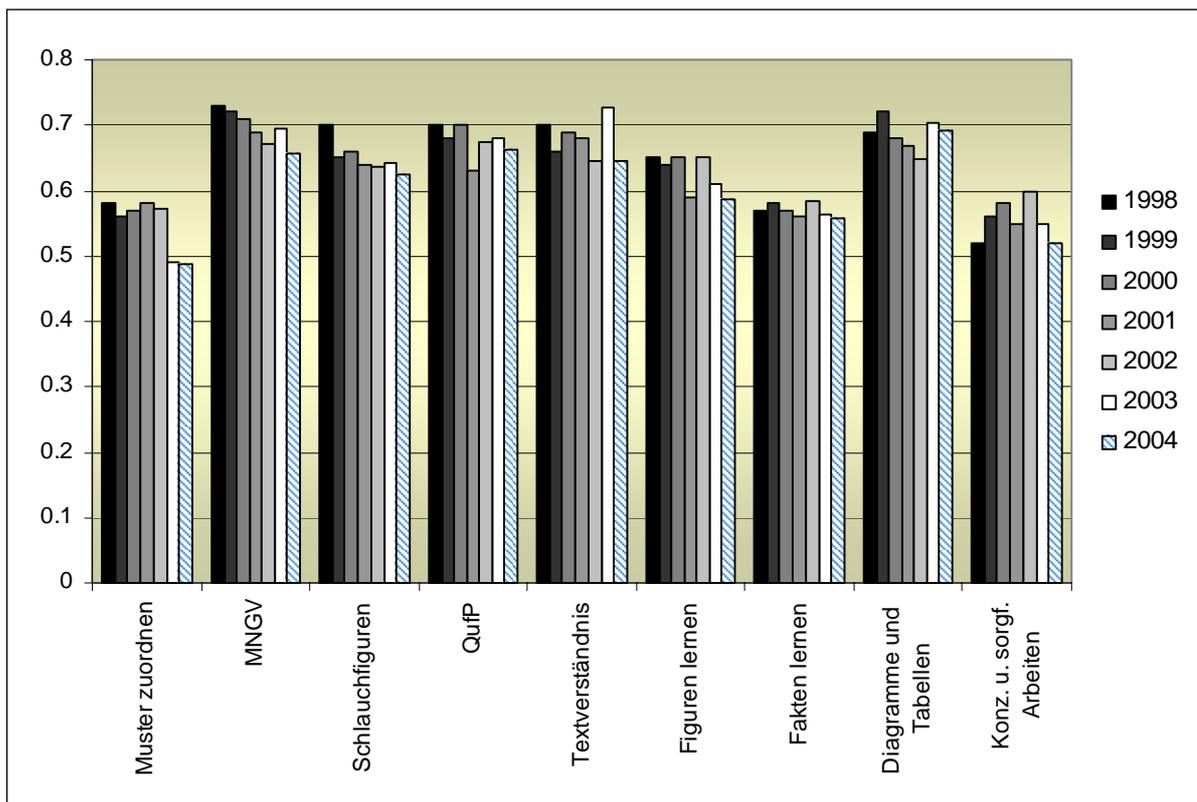


Abbildung 47: Korrelationen zwischen Punktwerten und Untertests nach Jahren

Die Drei-Faktoren-Struktur bleibt weiterhin stabil (Tabelle 48). Die Ladungsmuster stimmen über die Jahre gut überein, fielen in ihrem Ausmass dieses Jahr jedoch etwas tiefer aus.

Unrotierte Lösung															
Faktor	Eigenwert					% Varianz					Kumuliert %				
	2004	2003	2002	2001	2000	2004	2003	2002	2001	2000	2004	2003	2002	2001	2000
1	3.36	3.64	3.61	3.53	3.70	37.33	40.48	40.14	39.21	41.12	37.33	40.48	40.14	39.21	41.12
2	1.39	1.60	1.54	1.63	1.46	15.42	17.78	17.12	18.14	16.18	52.74	58.26	57.26	57.35	57.30
3	.90	0.92	0.77	.85	.84	10.05	10.28	8.53	9.39	9.31	62.79	68.54	65.79	66.74	66.61
4	.82	0.67	0.68	.62	.66	9.06	7.50	7.51	6.93	7.34	71.85	76.03	73.30	73.67	73.95
5	.62	0.54	0.58	.57	.58	6.90	5.95	6.41	6.35	6.40	78.75	81.98	79.71	80.03	80.35
6	.52	0.50	0.53	.55	.53	6.36	5.54	5.94	6.06	5.93	85.11	87.52	85.64	86.09	86.28
7	.54	0.48	0.49	.50	.47	6.03	5.38	5.40	5.53	5.23	91.14	92.90	91.05	91.61	91.51
8	.47	0.35	0.42	.42	.39	4.623	3.93	4.67	4.63	4.31	95.76	96.83	95.72	96.24	95.82
9	.38	0.29	0.39	.34	.38	4.24	3.17	4.28	3.76	4.18	100	100	100	100	100

Tabelle 48: Verlauf der unrotierten Lösung

Die rotierten Lösungen sind in Tabelle 49 dargestellt. Die Zweifaktorenlösung klärt 53% der Gesamtvarianz auf, die Dreifaktorenlösung 63%. Der Scree-Plot (Abbildung 48) stellt den Eigenwertverlauf der Faktoren grafisch dar. Die Ein-Faktorenlösung bleibt deutlich hervorgehoben. Neu hebt sich hier auch die Vierfaktorenlösung von den folgenden Lösungen ab, allerdings auf zu geringem Eigenwertniveau.

Varimax-rotierte Lösungen														
Eigenwert					% Varianz					Kumuliert %				
2004	2003	2002	2001	2000	2004	2003	2002	2001	2000	2004	2003	2002	2001	2000
Zwei-Faktorenlösung														
2.63	2.87	2.66	2.72	2.74	29.26	31.90	29.60	30.19	30.49	29.26	31.90	29.60	30.19	30.49
2.11	2.37	2.49	2.44	2.41	23.48	26.36	27.66	27.16	26.80	52.74	58.26	57.26	57.35	57.30
Drei-Faktorenlösung														
2.56	2.84	2.61	2.71	2.71	28.39	31.57	29.05	30.06	30.12	28.39	31.57	29.05	30.06	30.12
1.93	1.83	1.67	1.82	1.92	21.43	20.34	18.53	20.21	21.29	49.81	51.91	47.57	50.27	51.41
1.17	1.50	1.64	1.48	1.37	12.97	16.623	18.22	16.47	15.20	62.79	68.54	65.79	66.74	66.61

Tabelle 49: Varianzanteile der einzelnen Faktorenlösungen (rotierte Lösungen)

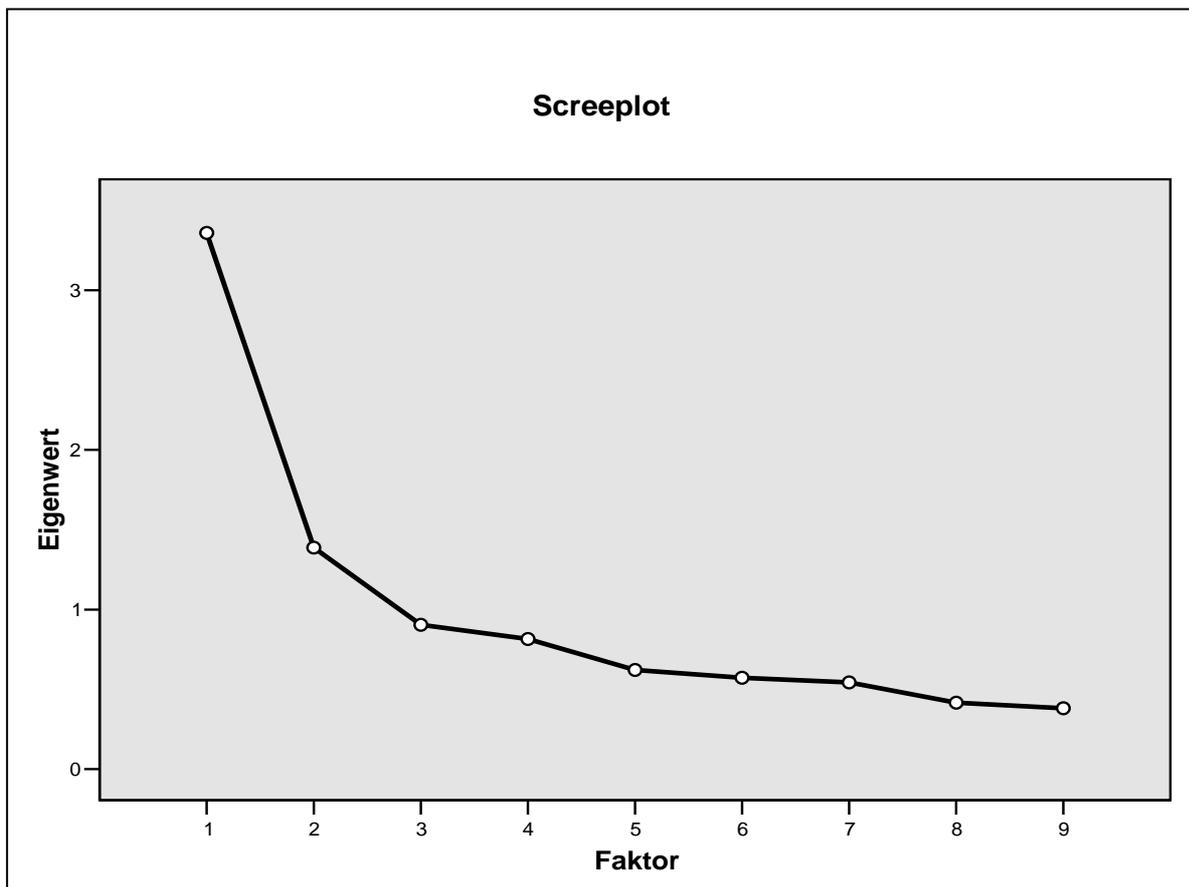


Abbildung 48: Scree-Plot (Eigenwerteverlauf) der Faktoren 1 bis 9 für die Untertests des EMS 2004

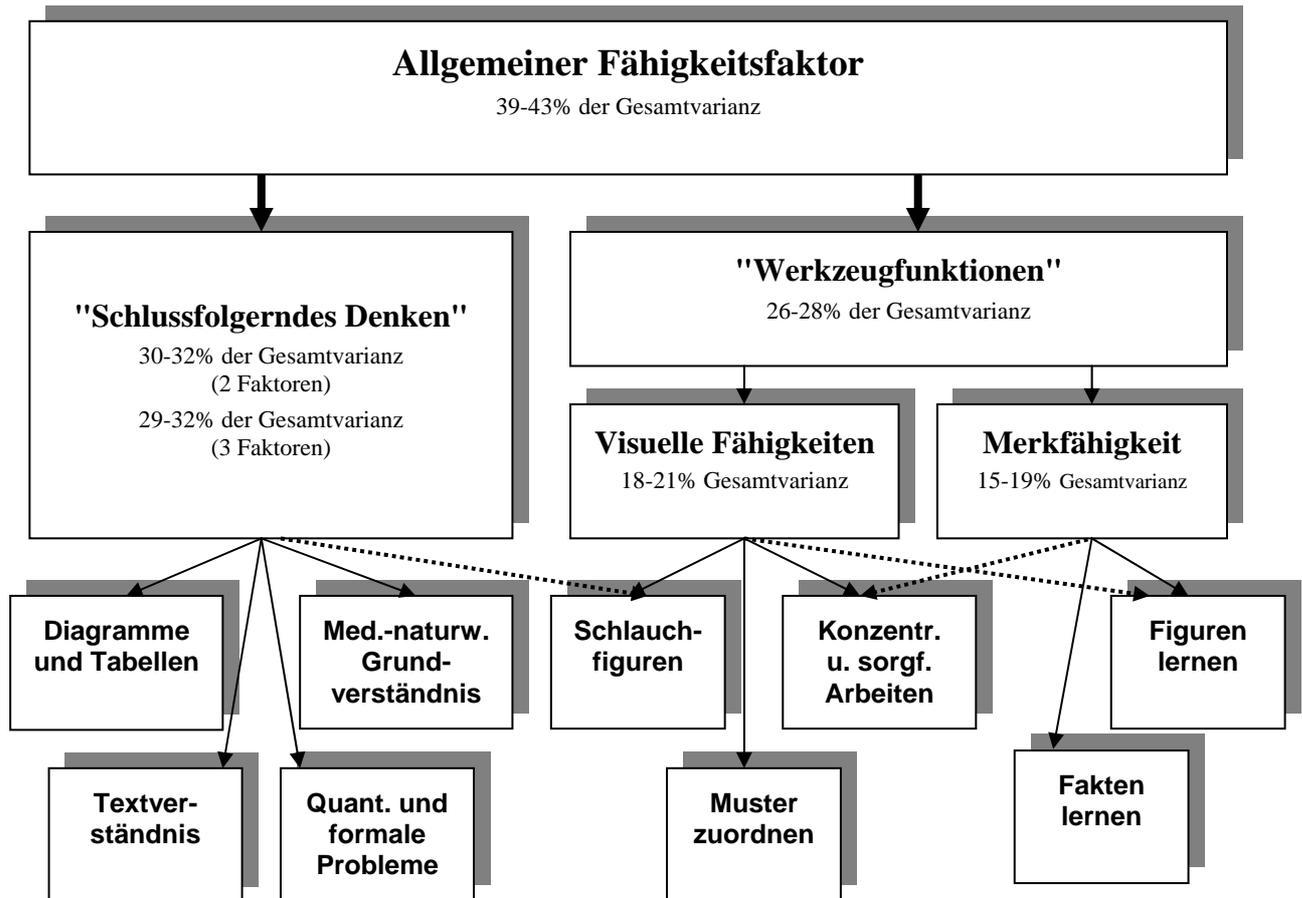


Abbildung 49: Struktur der Untertests des EMS, hierarchische Faktorenstruktur, angegeben auch der Wertebereich der Varianzaufklärung durch den Faktor 1998 bis 2003

Abbildung 49 stellt die Struktur der Untertests der Durchführungen 1998 bis 2003 dar. Der neu konzipierte Untertest „Konzentriertes und Sorgfältiges Arbeiten“ wirkt sich in einer leicht veränderten Struktur aus. In der Vier-Faktoren-Lösung würde die bisherige Struktur erhalten bleiben, „Konzentriertes und Sorgfältiges Arbeiten“ stellt als „Single Faktor“ den neuen vierten Faktor dar. Dies spricht dafür, dass mit der Neukonzipierung und dem damit verbundenen Wegfall der Vorbereitungsmöglichkeiten (im Sinne von exzessivem Training dieses einen Untertests) ein neuer, bisher im Gesamtergebnis nicht berücksichtigter Bereich erfasst wird.

Abbildung 50 stellt die diesjährige Faktoren-Lösung dar. Die Zwei-Faktoren-Lösung entspricht der bisherigen Struktur. In der Drei-Faktoren-Lösung kommt es zu Verschiebungen. „Figuren lernen“ und „Fakten lernen“ als typische Anforderungen an das Gedächtnis verbleiben im Faktor „Merkfähigkeit“. „Konzentriertes und Sorgfältiges Arbeiten“ bildet mit dem Untertest „Muster erkennen“ den Faktor „visuelle Fähigkeiten“. Die sehr wahrscheinlich aufgrund des Trainings früher im Untertest „Konzentriertes und Sorgfältiges Arbeiten“ enthaltene Gedächtniskomponente entfällt vielleicht auch, weil eine Vorbereitung nicht möglich war.

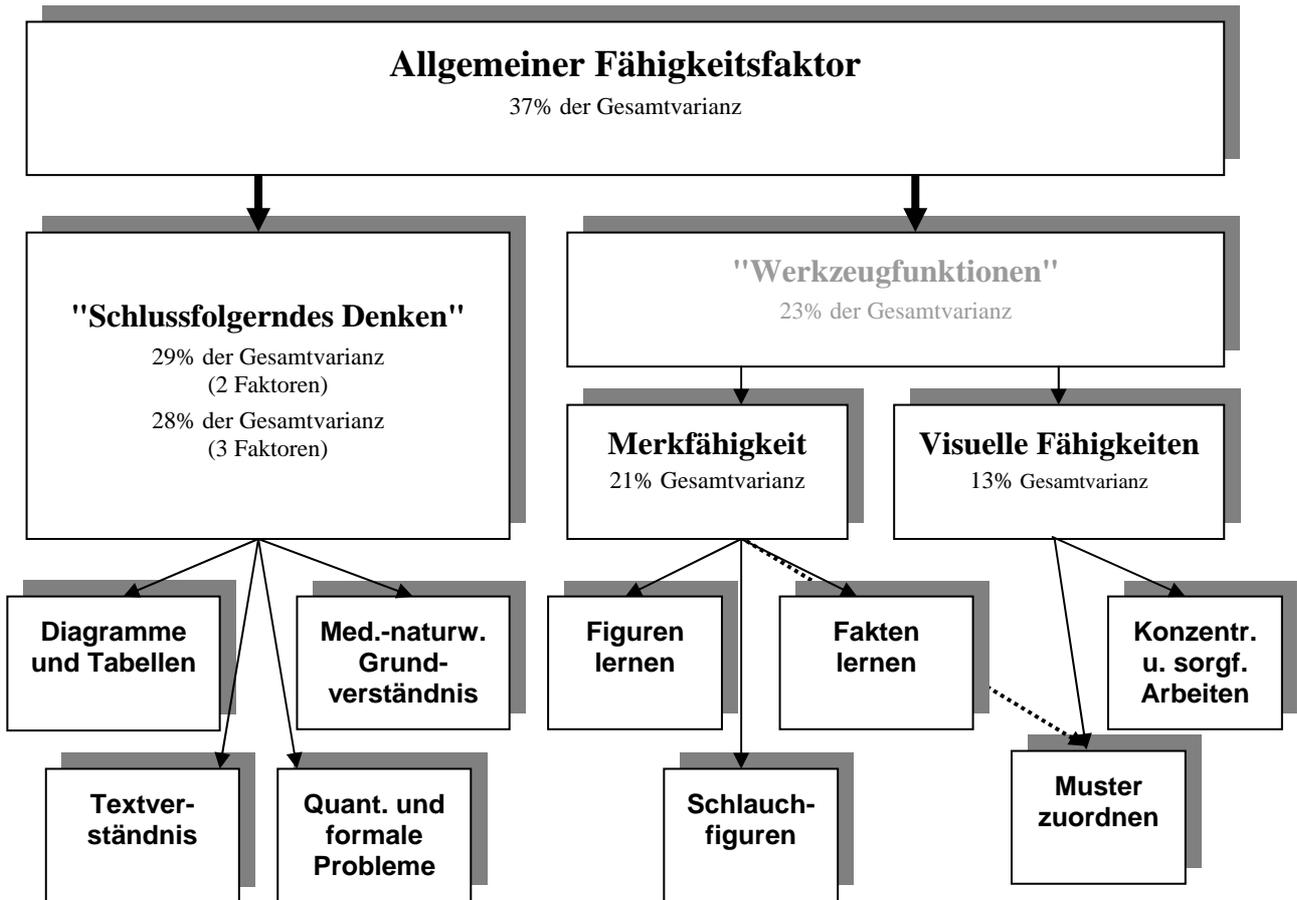


Abbildung 50: Struktur der Untertests des EMS, hierarchische Faktorenstruktur 2004

Der Untertest „Schlauchfiguren“ kommt in der Drei-Faktoren-Lösung neu ebenfalls im Bereich „Merkfähigkeit“ zu liegen. Es scheint sich hier nicht mehr nur um die Speicherung von Informationen zu handeln, sondern auch um die Fähigkeit, einmal gespeicherte Informationen zu verarbeiten und neu zu repräsentieren (Mnemotechniken, räumliches Vorstellungsvermögen).

	Ladungen 1998	Ladungen 1999	Ladungen 2000	Ladungen 2001	Ladungen 2002	Ladungen 2003	Ladungen 2004
Muster zuordnen	.56	.53	.57	.57	.71	.48	.45
Med.-naturwiss. Grundverst.	.76	.76	.74	.74	.70	.74	.71
Schlauchfiguren	.69	.64	.65	.62	.69	.61	.61
Quant. und formale Probleme	.70	.70	.72	.66	.68	.73	.70
Textverständnis	.72	.70	.71	.72	.62	.77	.70
Figuren lernen	.62	.59	.53	.56	.62	.55	.54
Fakten lernen	.53	.52	.55	.52	.57	.51	.51
Diagramme und Tabellen	.72	.75	.71	.72	.56	.76	.74
Konzentr. u. sorgf. Arbeiten	.47	.51	.55	.49	.54	.48	.44

Tabelle 50: Ladungen im ersten Faktor der unrotierten Lösung 1998 bis 2004 („Generalfaktor“ bzw. allgemeiner Fähigkeitsfaktor)

	Faktor 1					Faktor 2					Kommunalitäten (h^2)				
	00	01	02	03	04	00	01	02	03	04	00	01	02	03	04
Diagramme und Tabellen	.79	.77	.77	.83	.79		.21	(.17)	(.17)		.66	.63	.63	.73	.66
Med.-naturwiss. Grund- verständnis	.78	.82	.77	.86	.80	.22	(.18)	.21	(.11)		.66	.70	.63	.76	.65
Textverständnis	.80	.82	.82	.83	.70		(.16)	(.11)	(.18)		.67	.70	.68	.74	.63
Quant. und formale Probleme	.80	.81	.79	.77	.70			(.18)	.20		.68	.66	.65	.65	.60
Konzentr. u. sorgf. Arbeiten						.81	.79	.82	.79	.33	.66	.62	.68	.62	.20
Figuren lernen	(.15)	(.16)	.20			.64	.66	.70	.78	.74	.44	.46	.52	.65	.56
Muster zuordnen		(.15)	.20			.73	.69	.63	.58	.70	.54	.50	.43	.62	.47
Fakten lernen	.27	.15	(.11)	.26		.51	.62	.67	.50	.67	.34	.41	.47	.78	.47
Schlauchfiguren	.30	.25	.29	.26		.65	.66	.60	.66	.67	.51	.50	.45	.60	.52

Tabelle 51: Ladungen und Kommunalitäten der Zwei-Faktorenlösung, varimaxrotiert

In der Zwei-Faktoren-Lösung (Tabelle 51) bleibt die Anordnung der Untertests gegenüber früheren Strukturen unverändert. Bezüglich des Faktors „Schlussfolgerndes Denken“ (Faktor 1) bleibt der Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“ der ladungsstärkste. Der Faktor „Werkzeugfunktionen“ wird nun allerdings nicht mehr in erster Linie durch „Konzentriertes und Sorgfältiges Arbeiten“ repräsentiert, sondern der ladungsstärkste Untertest ist nun „Figuren lernen“. Die geringe Ladung (und Kommunalität) des erstgenannten Untertests zeigt erneut, dass hier offenbar neue Leistungskomponenten erfasst wurden.

Faktor	Schweiz Gesamt								Deutschland		
	1		2		3		h ²		1	2	3
	2003	2004	2003	2004	2003	2004	2003	2004			
Diagramme und Tabellen	.83	.77	.22	.24			.74	.67	.82		
Medizin.-naturwiss. Grundverständnis	.86	.81			.16		.76	.68	.81		.20
Quantitat. u. formale Probleme	.77	.73	.25	.23			.65	.60	.80	.18	
Textverständnis	.82	.80			.24		.74	.66	.79		.20
Muster zuordnen	.15		.77	.54		.52	.62	.57		.81	
Konzentr. u. sorgfält. Arbeiten		.20	.61	.85	.50		.63	.76		.70	.40
Schlauchfiguren	.26	.23	.71		.17	.66	.60	.53	.35	.71	
Fakten lernen	.24				.85	.74	.78	.58	.21		.87
Figuren lernen			.49		.64	.76	.65	.60	.13	.47	.64

Tabelle 52: Faktorenanalyse: Varimaxrotierte Drei-Faktorenlösung, Schweiz und Deutschland (deutsche Daten nach Blum, 1996, in Trost et al., 1998, S. 42)

In der Drei-Faktoren-Lösung (Tabelle 52) werden die Faktoren neu durch die Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Konzentriertes und Sorgfältiges Arbeiten“ und „Figuren lernen“ charakterisiert.

8.3 Item-Trennschärfen

Die Item-Trennschärfen (korrigierte Item-Total-Korrelation) folgen den deutschen Kennwerten. Massive Abweichungen und Werte nahe Null bzw. im Minusbereich könnten z.B. auf fehlerhafte Itemlösungen hinweisen.

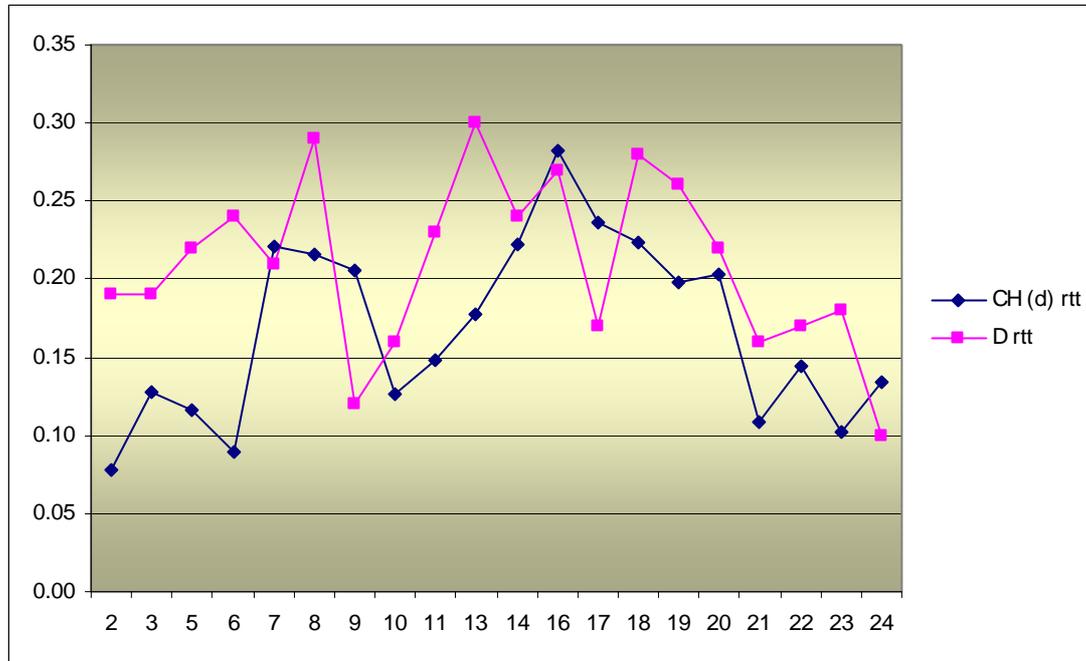


Abbildung 51: Trennschärfen für den Untertest „Muster zuordnen“

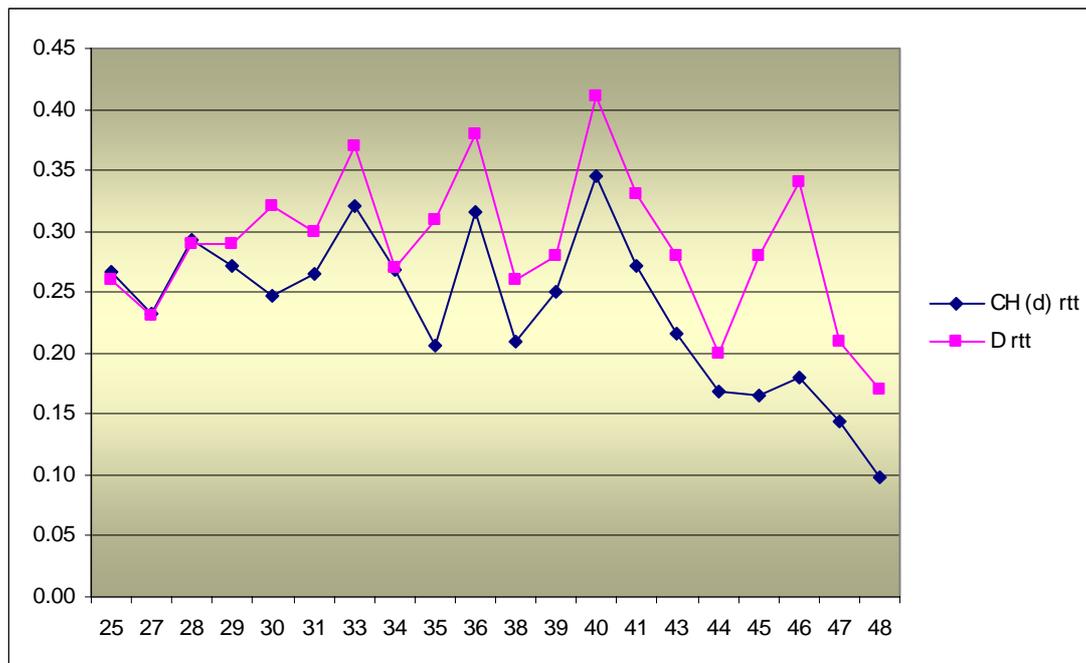


Abbildung 52: Trennschärfen für den Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“

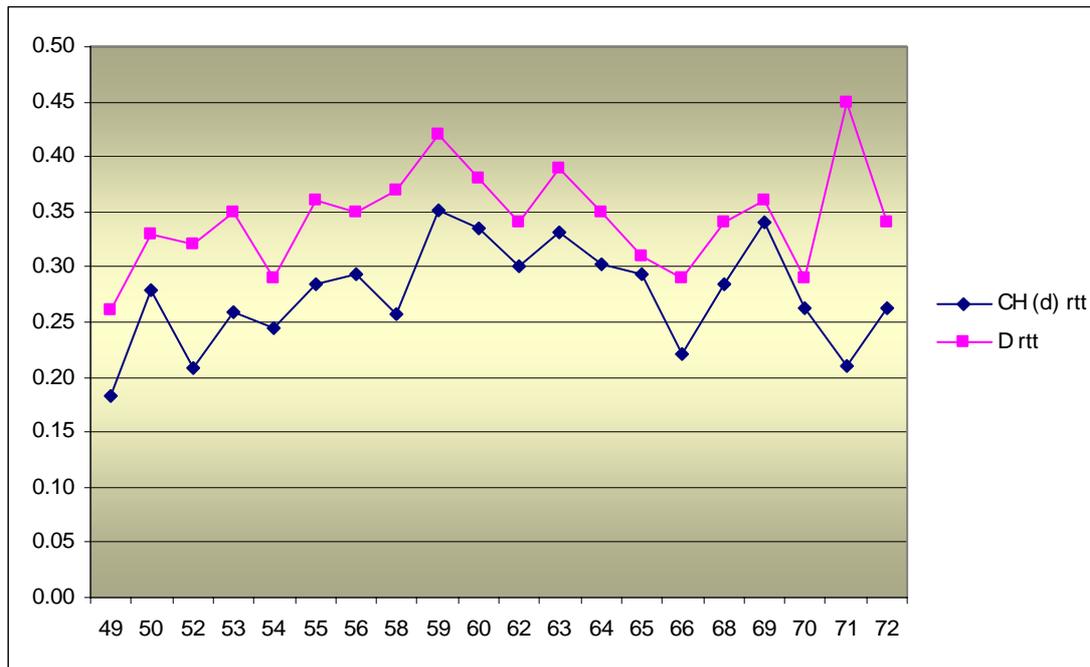


Abbildung 53: Trennschärpen für den Untertest „Schlauchfiguren“

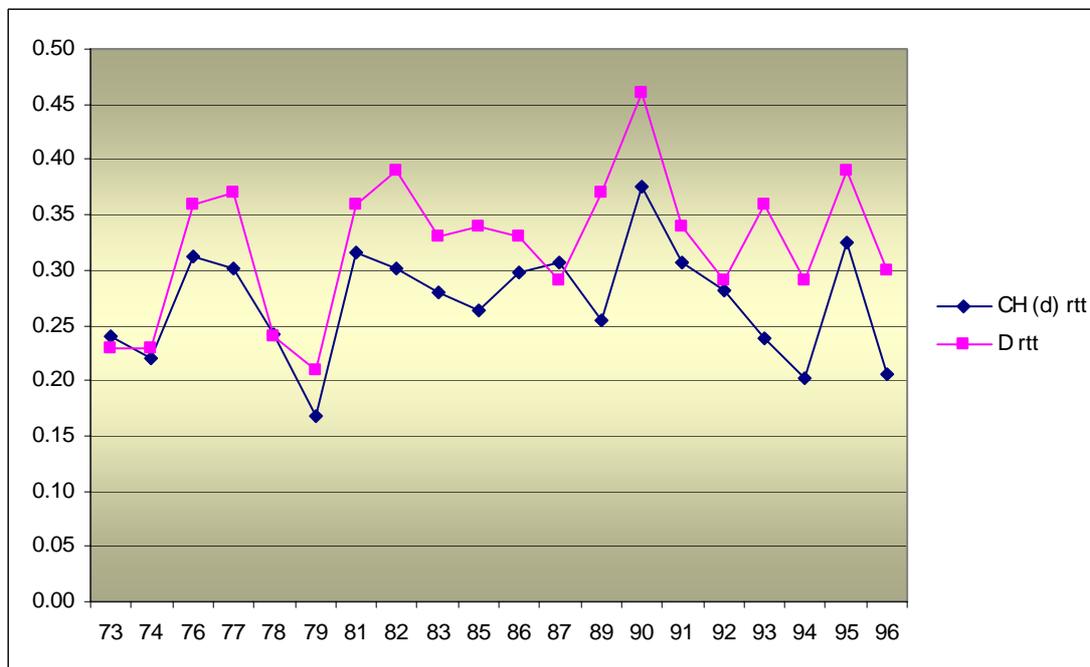


Abbildung 54: Trennschärpen für den Untertest „Quantitative und formale Probleme“

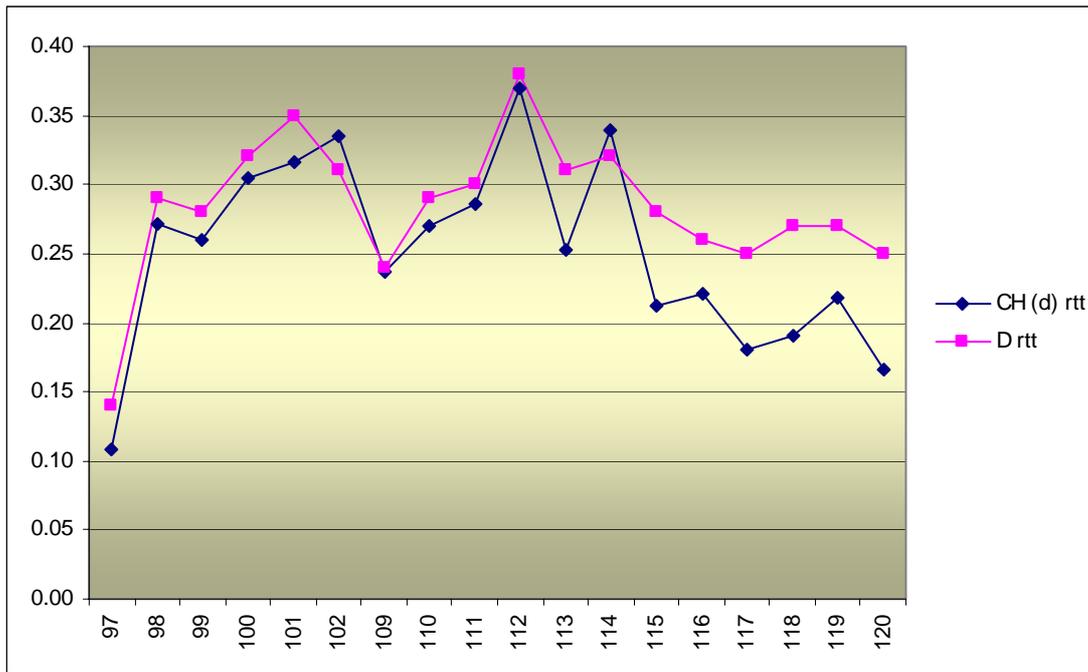


Abbildung 55: Trennschärfen für den Untertest „Textverständnis“

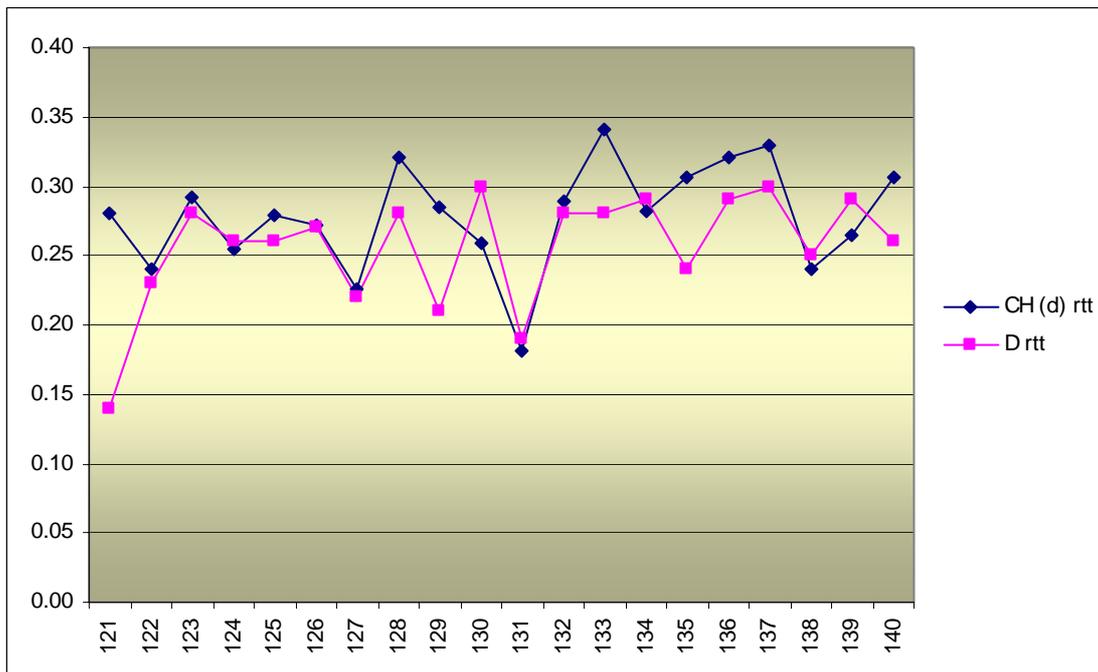


Abbildung 56: Trennschärfen für den Untertest „Figuren lernen“

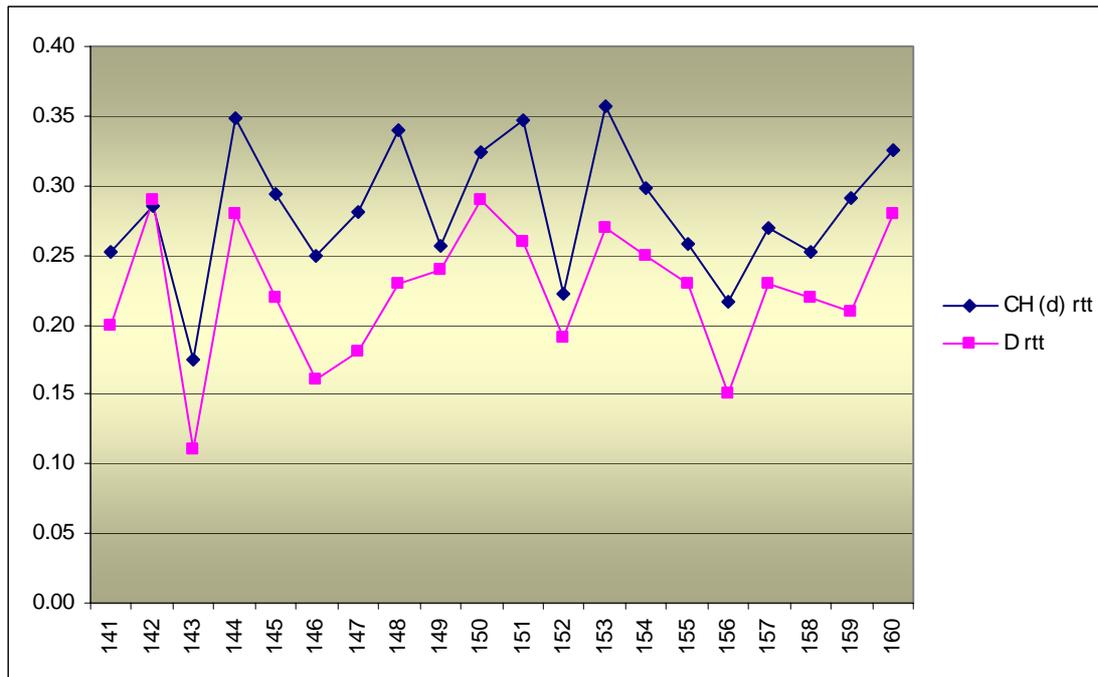


Abbildung 57: Trennschärfen für den Untertest „Fakten lernen“

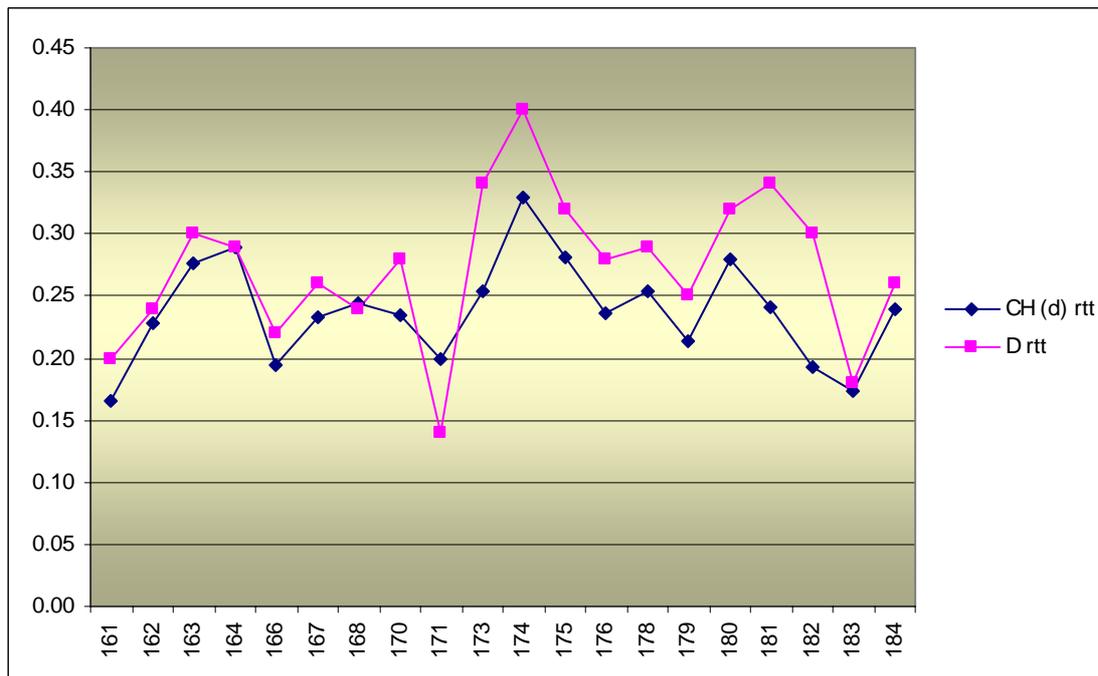


Abbildung 58: Trennschärfen für den Untertest „Diagramme und Tabellen“

8.4 Konzentriertes und Sorgfältiges Arbeiten

Dieser Untertest wurde 2004 revidiert und durch eine andere Anforderung (Zeichen sowie Durchstreichregel) ersetzt. Es hatte sich früher gezeigt, dass durch exzessives Üben (40 bis teilweise 80mal) bei diesem Test ein so hoher Automatisierungsgrad erreicht wurde, dass nahezu alle Zeichen bearbeitet werden konnten (vgl. Bericht 9 des ZTD und Vorbereitungsreport 2003). Dieses Ergebnis ist nicht im Sinne der Messintention dieses Tests: Es soll kein „Fleisstest“ sein, sondern tatsächlich Konzentration und Sorgfalt gemessen werden. Aus der Literatur ist bekannt, dass solche hochspezialisierten Automatisierungen stark bedingungsabhängig sind – ändert sich etwas, müssen diese Automatismen neu antrainiert werden (etwa bei der Fließbandarbeit).

Indem nun pro Jahr Zeichen und Regel vorher nicht bekannt sein werden, wirken sich Lerneffekte durch exzessives Üben nicht mehr so aus. Die weiter empfohlene Vorbereitung mit der bisher verwendeten Testform dient dem Zweck, Erfahrungen zu sammeln, wie man 8 Minuten zusammenhängend konzentriert arbeiten kann und wie man die Markierungen vornehmen muss, dass sie regelkonform sind.

Die Zahl der Zeichen auf dem Arbeitsbogen wurde von 1200 auf 1600 erhöht. Dies war möglich, weil die neue Einlesetechnik keine Markierungszonen um das eigentliche Zeichen mehr verlangt – dadurch wurde der Bogen insgesamt sogar übersichtlicher. „Deckeneffekte“, d.h. dass alle Zeichen richtig gelöst werden, sind dadurch unwahrscheinlich und der Test differenziert in allen Leistungsbereichen gut.

2004 wurde die nachfolgende Anforderung verwendet: In einer Zeilenfolge von p und q waren diejenigen p zu markieren, vor denen ein q steht. Das Testblatt besteht aus 40 Zeilen mit je 40 Zeichen (1600 mögliche zu bearbeitende Zeichen). Davon sind 800 Zeichen p, 400 davon sind durchzustreichen (Target), 400 nicht (Non-Target).

Fehler werden nur vor dem letzten bearbeiteten Zeichen gewertet. 800 Zeichen sind q (Distraktor).

Als Parameter ergeben sich:

	Beschreibung	Wertebereich
Richtige	Target P nach Q markiert	0...400
Fehler I	Target P nach Q nicht markiert, „Übersehen“	0...400
Fehler IIa	Nontarget P markiert, „falscher Alarm“	0...400
Fehler IIb	Distraktor Q markiert, „falscher Alarm“	0...800

Die Auswertung wird mit der bisher üblichen vergleichbar gehalten. Der Punktwert PW berechnet sich wie folgt:

$$PW = (\text{Richtige} - \text{Fehler I} - \text{Fehler IIa} - \text{Fehler IIb}) / K \quad \rightarrow \text{im Intervall 0 bis 20}$$

$$K = \text{Maximum für Richtige} / \text{Maximum Punktwert (20)}$$

Das Verhalten hinsichtlich extremen Antwortverhaltens ist bei bisherigem und neuem Untertest vergleichbar. Theoretisch mögliche Punktwerte kleiner als Null werden auf Null gesetzt.

Bisherige Version (b mit 2 Strichen)	Neue Version (PQ-Test)
--------------------------------------	------------------------

(K = 30)	(K = 20)
Alles richtig	
PW = (600 – 0 – 0) / 30 = 20	PW = (400 – 0 – 0 - 0) / 20 = 20
Nur letztes Zeichen angestrichen - Maximum „Übersehen“	
PW = (1 – 600 – 0) / 30 = -20 gesetzt 0	PW = (1 – 400 – 0 – 0) / 20 = -20 gesetzt 0
Alle Zeichen angestrichen Fehler Maximum: „Fälschlich angestrichen“	
PW = (600 – 0 – 600) / 30 = 0	PW = (400 – 0 – 1200 - 0) / 20 = -40 ->0
Alle p's angestrichen - Fehler Maximum: Non-Target fälschlich angestrichen“	
Entfällt	PW = (400 – 0 – 400 - 0) / 20 = 0
Teile werden übersprungen, Beitrag zum Fehler „Übersehen“	
PW = (0 – Übersehen - 0) / 30: negativ Nicht „neutral“, sondern mindert PW	P (0 – Übersehen – 0 - 0) / 20 = negativ Nicht „neutral“, sondern mindert PW
Leeren Bogen abgegeben	
PW = (0 – 0 - 0) / 30 = 0	PW = (0 - 0 - 0 – 0) / 20 = 0

Alle Extremfälle werden gleich behandelt, das Auswerteprinzip ist damit vergleichbar.

	richtig	Fehler	Fehler I	Fehler IIa	Fehler IIb
Mittelwert	232.4	20.3	10.5	4.9	4.85
Median	230.0	12.00	5.0	1.0	3.00
Standardabweichung	46.7	33.5	16.6	22.4	6.6
Schiefe	.054	5.739	4.402	10.880	4.838
Kurtosis	.506	41.755	27.283	139.213	36.680

Tabelle 53: Statistiken für Parameter des Untertests „Konzentriertes und Sorgfältiges Arbeiten“

	n	richtig		Fehler		Fehler I		Fehler IIa		Fehler IIb	
		m	s	m	s	m	s	m	s	m	s
deutsch	1255	233.4	46.7	20.9	34.6	10.7	17.0	5.2	23.3	5.0	6.9
französisch	68	217.6	48.4	12.3	20.3	6.7	11.1	2.6	9.9	3.0	2.5
italienisch	50	228.5	43.1	17.9	14.8	11.6	12.0	2.3	3.7	4.0	4.7

Tabelle 54: Sprachspezifische Statistiken für Parameter des Untertests „Konzentriertes und Sorgfältiges Arbeiten“

Insgesamt scheint die Anforderung etwas schwerer als beim bisherigen Untertest zu sein: Es werden im Mittel nur ca. die Hälfte der Zeichen bearbeitet wie in der Vorgängerversion. Beim bisherigen Konzentrationstest wurden zwischen 4 und 5% Fehler gemacht, jetzt sind es 8 bis 9%. Die Genauigkeit hat dadurch etwas an Bedeutung für den Punktwert etwas zugenommen. Auch das ist gewünscht, weil Sorgfalt und Arbeitsgeschwindigkeit gleichgewichtig sein sollen.

Bei beiden Unterschieden spielt die geringere Trainierbarkeit aber eine entscheidende Rolle: Die Punktwerteverteilung ist nicht mehr rechtsschief, sondern entspricht mehr einer Normalverteilung.

		Fehler	Fehler_1	Fehler_2a	Fehler_2b
Richtig	Korrelation nach Pearson	-.10(**)	-.24(**)	.02	.04
	Signifikanz (2-seitig)	.000	.000	.477	.198
Fehler	Korrelation nach Pearson		.72(**)	.84(**)	.41(**)
	Signifikanz (2-seitig)		.000	.000	.000
Fehler_1	Korrelation nach Pearson			.28(**)	.19(**)
	Signifikanz (2-seitig)			.000	.000
Fehler_2a	Korrelation nach Pearson				.17(**)
	Signifikanz (2-seitig)				.000

Tabelle 55: Korrelationen der Parameter im Konzentrationstest; kursiv: Signifikanzniveau
 ** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant

Die Beziehung zwischen Mengen- und Sorgfalsleistung zeigt keine lineare Abhängigkeit, die Korrelation mit -0.10 ist angesichts der hohen Personenzahl vergleichsweise gering. Extreme Fehlerzahlen resultieren vor allem durch Übersehen von Zeichen, nicht instruktionsgemäßes Vorgehen (alle p durchgestrichen) und in einem Falle nicht regelkonformer Markierung der Zeichen.

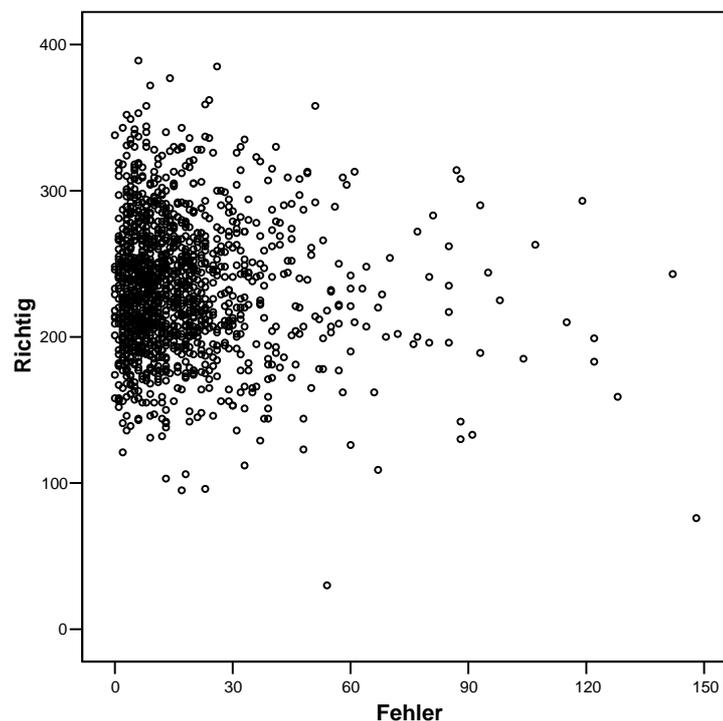


Abbildung 59: Scatterplot für alle Personen - Mengenleistung (Richtige) und Genauigkeit (niedrige Fehlerzahl)

9 Zusammenfassung Kantone 1998 bis 2004

Seit 1998 haben bisher 6720 Personen am EMS teilgenommen. In der Tabelle 56 sind alle Teilnehmer am EMS 1998 bis 2004 nach Kantonen aufgeführt. Zu beachten ist, dass 1998 nur Humanmedizin, seit 1999 Human- und Veterinärmedizin und 2004 zusätzlich Zahnmedizin einem NC unterlagen.

Wohnkanton/ Wohnort	Disziplin			Geschlecht		Sprache			Total
	Human- medizin	Veterinär- medizin	Zahn- medizin	männlich	weiblich	deutsch	franzö- sisch	italie- nisch	
AG	431	71	14	178	338	516			516
AI	12	5		6	11	17			17
AR	42	6	4	19	33	52			52
BE	877	182	31	377	713	1071	18	1	1090
BL	316	41	10	139	228	365	2		367
BS	258	21	4	101	182	282		1	283
FR	244	46	1	109	182	107	184		291
GE	4	45		13	36	2	47		49
GL	24	4		6	22	28			28
GR	192	43	8	86	157	235		8	243
JU	5	13		6	12	1	17		18
LU	315	58	2	130	245	375			375
NE	6	30		5	31	3	33		36
NW	32	5	1	10	28	38			38
OW	31		2	12	21	33			33
SG	360	57	4	149	272	421			421
SH	57	9	3	24	45	69			69
SO	197	39	9	100	145	244	1		245
SZ	108	14	2	40	84	124			124
TG	162	37	6	60	145	205			205
TI	187	43	8	103	135	4		234	238
UR	18	5		11	12	23			23
VD	11	90		35	66	5	96		101
VS	82	35	4	45	76	81	40		121
ZG	80	26	1	42	65	106	1		107
ZH	1298	193	33	536	988	1511	4	9	1524
FL	25	4	2	11	20	31			31
Ausland	62	11	2	31	44	73	2		75
Total	5436	1133	151	2384	4336	6022	445	253	6720

Tabelle 56: Gesamtzahl der Teilnehmer am EMS 1998 bis 2004. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die Testwerte sind durch die Standardisierung auf den jeweiligen Mittelwert und die Streuung des Testjahrganges zwischen den Jahren direkt vergleichbar. Dies lässt, wie mehrfach betont, keine Bewertung von Schulbildung zu, weil der Test eher bildungsunabhängig ist und die Bewerber nicht repräsentativ für die Maturanden sind.

Wohnkanton/ Wohnort	Humanmedizin					Veterinärmedizin				
	n	Mittelwert	Std. Abw.	Minimum	Maximum	n	Mittelwert	Std. Abw.	Minimum	Maximum
AG	431	102.4	9.1	74	129	71	100.2	8.7	78	118
AI	12	104.7	8.8	88	116	5	100.6	7.1	95	111
AR	42	100.1	9.5	81	127	6	102.0	10.5	91	118
BE	877	100.4	9.9	72	125	182	97.4	8.8	76	121
BL	316	100.4	10.1	70	125	41	99.7	9.7	74	118
BS	258	97.0	10.2	71	122	21	96.8	10.3	76	117
FR	244	95.2	8.8	70	121	46	93.0	8.0	73	106
GE	4	100.3	10.8	85	110	45	93.2	8.4	76	117
GL	24	102.6	9.3	90	130	4	97.8	6.0	89	102
GR	192	99.4	10.0	76	128	43	97.8	8.1	77	113
JU	5	90.0	4.2	85	96	13	93.8	7.5	85	109
LU	315	100.9	9.6	71	126	58	98.1	8.7	76	118
NE	6	102.3	10.4	91	116	30	91.7	8.3	76	106
NW	32	101.2	11.2	78	121	5	98.6	9.0	88	109
OW	31	100.4	8.2	89	126	
SG	360	101.9	10.0	70	128	57	100.5	8.0	79	118
SH	57	101.2	9.1	82	123	9	98.0	10.4	86	114
SO	197	101.0	9.6	70	127	39	98.1	8.4	80	119
SZ	108	98.5	9.5	72	127	14	98.4	10.2	80	121
TG	162	101.0	9.8	77	123	37	102.2	8.8	82	122
TI	187	96.8	9.3	73	120	43	94.2	7.4	82	112
UR	18	102.1	7.7	87	120	5	98.4	8.8	91	112
VD	11	103.2	9.1	91	119	90	93.5	7.7	78	110
VS	82	98.5	10.1	70	126	35	99.0	10.5	75	120
ZG	80	101.8	10.2	73	128	26	98.5	7.0	84	109
ZH	1298	102.5	10.2	70	130	193	99.4	10.2	76	126
FL	25	100.0	10.1	82	120	4	105.0	12.7	92	122
Ausland	62	97.3	9.9	81	119	11	87.7	7.7	76	99
Total	5436	100.6	10.0	70	130	1133	97.5	9.2	73	126

Tabelle 57: Kennwerte für den Testwert von Kantonen 2004 nach Disziplinen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt. Zahnmedizin wird wegen der geringen Personenzahl hier nicht aufgeführt.

Nachfolgend werden in diesem Jahr erstmals die mittleren Punktwerte der einzelnen Untertests für alle Teilnehmenden am EMS aufgeführt. Diese sind nur innerhalb eines Jahres numerisch vergleichbar – durch die Mittelung über alle Jahre sind dennoch Aussagen möglich, aus welchen Teilergebnissen die Testwertunterschiede pro Kanton resultieren.

WK	Muster zuordnen		Med.-nat. Grundverst.		Schlauchfiguren		Quant-formale Probl.		Text-verständ-nis		Figuren lernen		Fakten lernen		Diagr. u. Tabellen		Konz. u. sorgfält. Arbeit.	
	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S	M	S
AG	11.1	2.9	11.4	3.4	12.9	3.5	11.9	3.7	9.9	3.6	12.8	3.8	11.6	3.6	11.2	3.2	12.9	3.4
AI	11.6	3.4	11.1	2.8	12.2	3.3	13.0	2.6	8.1	2.7	14.6	4.1	12.1	3.8	11.5	2.6	13.4	3.7
AR	10.8	3.0	10.3	4.0	13.0	4.2	10.8	3.9	8.8	3.6	12.2	4.0	9.6	3.9	10.6	3.5	13.3	3.7
BE	10.8	2.9	10.7	3.5	12.6	3.7	10.6	3.9	9.3	3.5	12.6	3.9	11.1	3.8	10.8	3.3	12.3	3.6
BL	11.1	3.2	10.8	3.4	13.0	3.7	10.5	3.7	9.2	3.4	12.9	4.0	11.2	4.0	10.9	3.2	12.7	3.8
BS	10.6	3.0	9.8	3.7	12.5	3.5	9.5	3.7	8.2	3.6	12.2	4.1	10.5	3.7	10.1	3.3	12.4	3.7
FR	9.8	2.9	9.0	3.2	11.1	3.4	10.3	3.8	8.2	3.2	11.2	3.8	10.1	3.3	9.6	3.1	10.7	3.3
GE	9.6	2.9	8.3	3.4	11.9	3.4	10.1	3.7	6.8	3.2	10.8	4.0	9.0	3.2	9.0	3.0	10.8	3.5
GL	11.3	3.1	11.5	2.8	12.9	3.7	11.0	3.7	9.9	3.7	12.7	3.6	11.5	2.9	11.9	3.0	13.6	3.6
GR	10.7	3.0	10.2	3.4	12.2	3.6	10.8	3.6	8.7	3.6	12.5	3.8	11.3	3.7	10.5	3.1	12.7	3.8
JU	9.7	2.7	7.8	3.7	9.4	3.1	9.5	3.2	6.7	2.2	12.1	2.9	10.4	2.8	9.2	3.0	10.6	3.5
LU	11.0	3.2	10.8	3.1	12.4	3.4	11.4	3.9	9.4	3.6	13.1	3.9	11.2	3.5	11.0	3.3	12.6	3.4
NE	10.4	3.4	8.8	3.4	11.1	3.8	8.9	4.2	6.9	3.5	11.9	4.4	9.9	3.1	8.7	3.2	10.4	2.8
NW	11.3	3.4	11.0	3.3	12.1	4.4	11.3	3.6	9.3	3.3	14.2	4.1	12.6	3.5	10.9	3.7	12.3	3.9
OW	11.5	3.0	10.5	3.4	12.6	2.7	10.7	4.4	8.6	3.3	12.5	3.5	10.9	3.0	10.6	3.3	13.1	3.4
SG	11.6	3.1	10.8	3.3	13.3	3.7	11.4	3.6	9.3	3.5	13.0	4.1	11.6	3.7	11.1	3.3	13.3	3.5
SH	11.0	2.6	10.4	3.1	12.6	3.6	11.3	3.7	9.6	3.5	12.0	3.7	10.9	3.6	11.1	3.3	13.7	3.5
SO	11.2	3.0	10.6	3.3	12.9	3.7	10.8	3.5	8.9	3.6	13.4	3.8	11.0	3.7	10.9	3.1	13.0	4.1
SZ	11.1	3.2	9.9	3.8	12.0	3.9	10.6	3.8	8.8	3.7	12.4	3.7	10.6	3.5	10.3	3.2	12.8	3.5
TG	11.2	3.0	10.5	3.5	12.9	3.8	11.1	3.7	9.1	3.5	13.0	4.1	11.4	3.6	10.8	3.3	13.4	3.5
TI	10.3	2.9	9.3	3.4	11.0	3.3	10.6	3.9	8.2	3.3	11.0	3.6	10.3	3.6	9.2	3.2	11.6	3.2
UR	10.9	2.4	10.7	4.1	13.4	3.1	11.6	3.7	9.1	3.8	12.7	4.2	11.3	3.3	11.8	2.9	13.1	3.3
VD	9.8	2.8	9.1	3.0	11.1	3.2	9.0	4.1	8.0	3.1	10.9	3.5	10.0	3.5	9.5	3.1	10.8	2.6
VS	10.7	2.9	10.1	3.5	11.8	4.0	11.1	3.8	9.0	3.2	12.1	4.1	11.1	3.9	10.4	3.2	11.7	3.4
ZG	11.3	3.0	10.5	3.3	13.3	3.6	10.8	3.8	9.0	3.7	13.1	4.0	11.1	3.6	10.7	3.6	13.0	4.0
ZH	11.3	3.0	11.0	3.5	13.2	3.8	11.4	3.8	9.7	3.7	13.0	4.2	11.8	3.9	11.1	3.3	13.2	4.0
FL	11.2	2.9	10.6	2.9	12.4	3.9	10.3	3.2	8.9	3.8	11.8	4.4	11.8	3.7	11.5	3.3	13.4	4.0
Aus-land	9.8	2.9	9.3	3.9	11.7	3.7	9.4	4.3	8.5	3.8	11.6	3.9	10.8	3.1	9.8	3.8	12.0	4.0
To-tal	11.0	3.0	10.6	3.5	12.6	3.7	10.9	3.8	9.2	3.6	12.6	4.0	11.2	3.8	10.7	3.3	12.6	3.8

Tabelle 58: Mittelwerte (M) und Standardabweichungen (S) für die Punktwerte der Untertests nach Wohnkantonen (WK) für alle EMS-Teilnehmenden.

Im Folgenden werden getrennt für Human- und Veterinärmedizin Testwertverteilungen und theoretische (!) Zulassungsquoten verglichen. Letztere sind nach der Kapazität festgelegt und berücksichtigen nicht die Rückzüge und erneute Vergabe dieser Plätze. Personenzahlen kleiner 20 sind in Klammern beim Kanton aufgeführt. Zahnmedizin wird wegen der geringen Personenzahl noch nicht berücksichtigt.

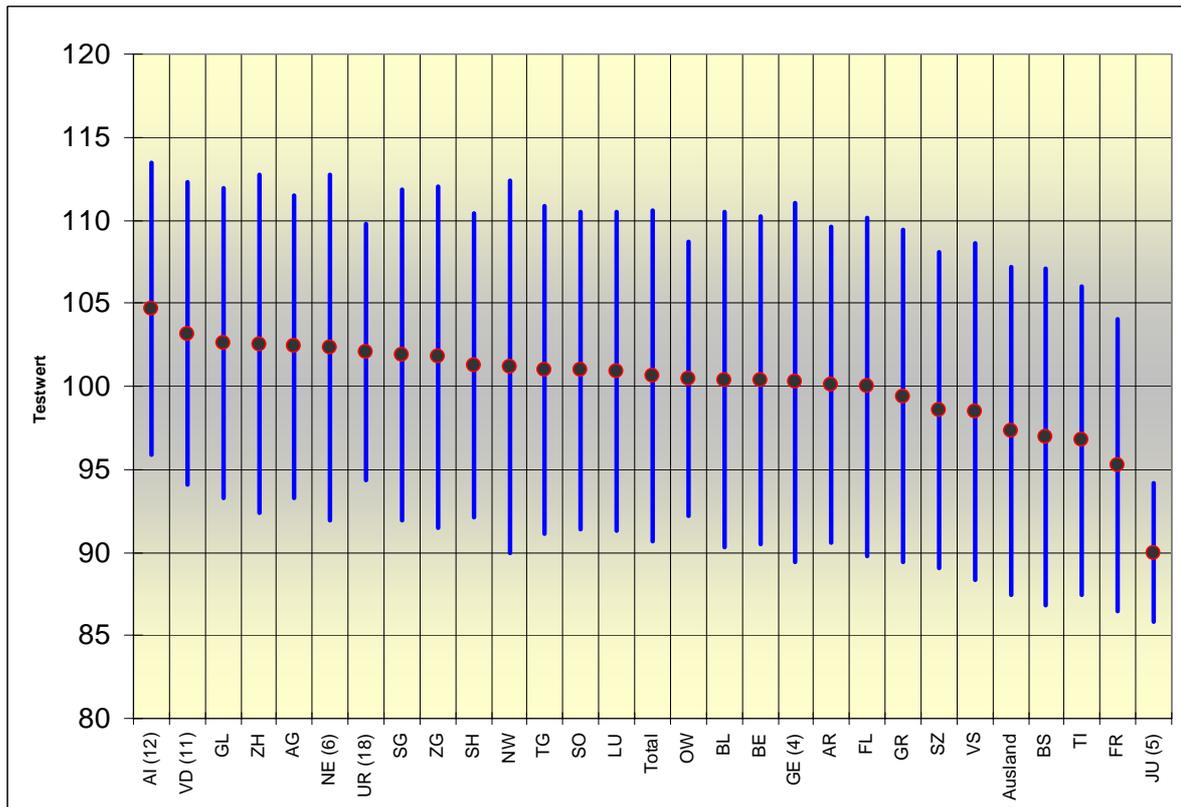


Abbildung 60: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1998 bis 2004 zusammengefasst; **Humanmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

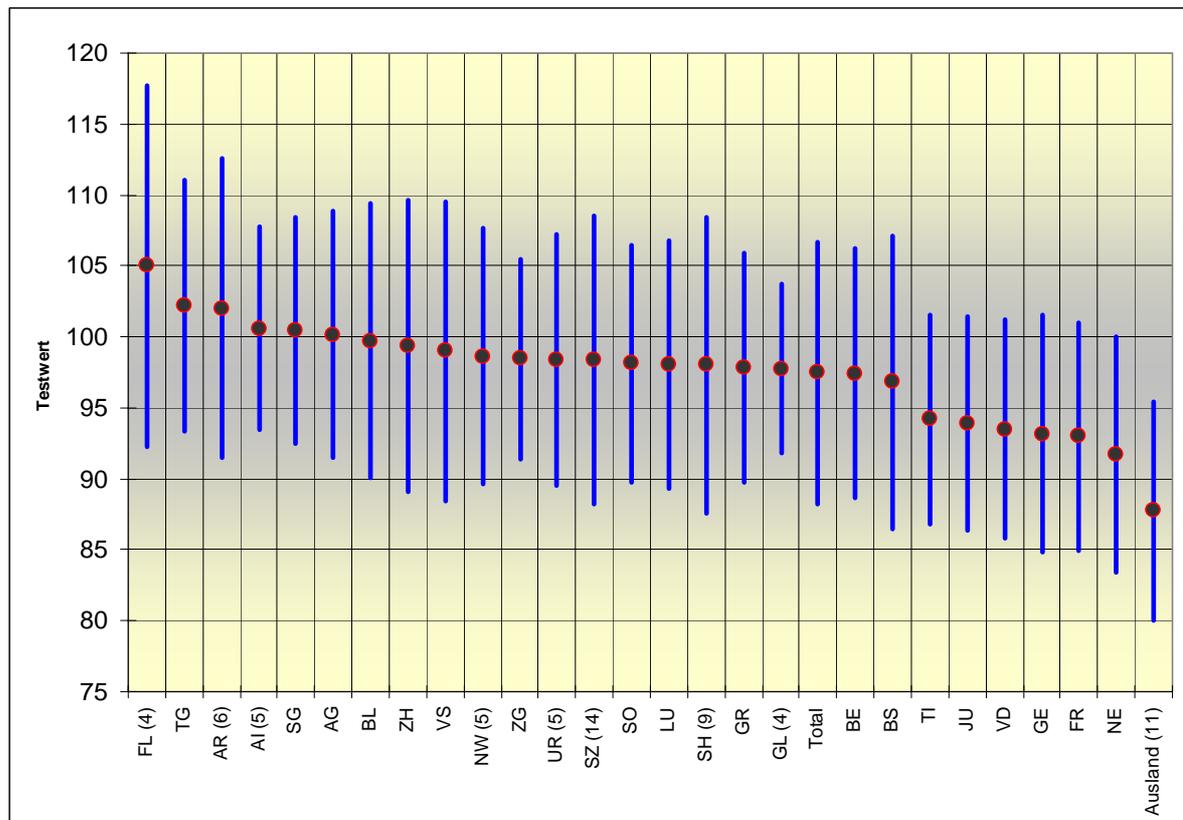


Abbildung 61: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1999 bis 2004 zusammengefasst; **Veterinärmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

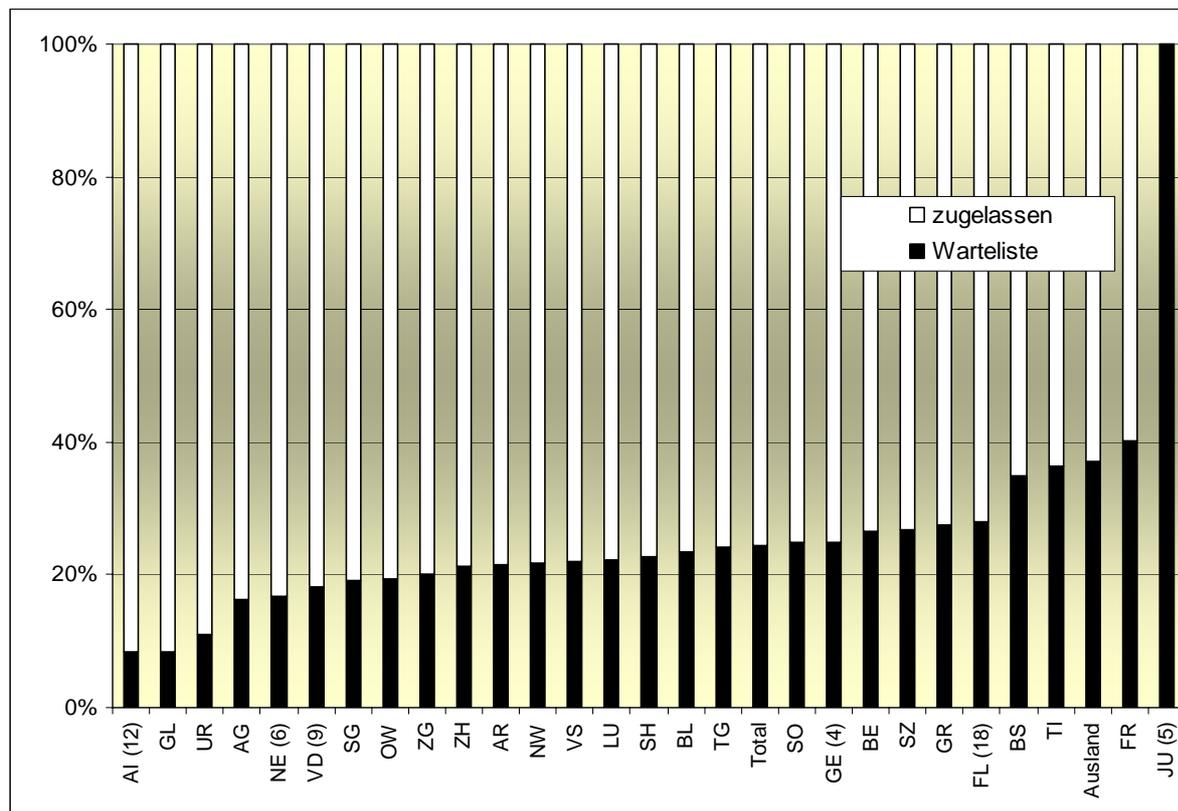


Abbildung 62: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1998 bis 2004; **Humanmedizin**

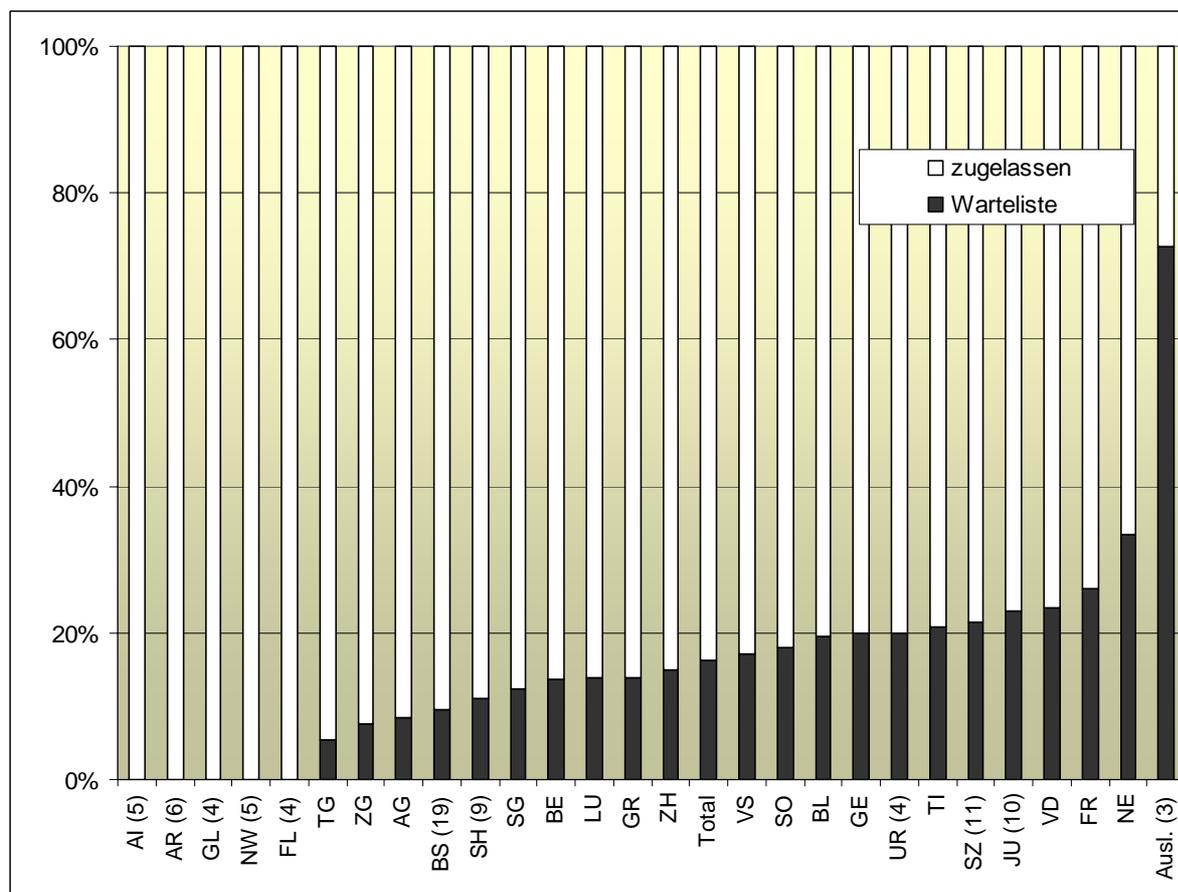


Abbildung 63: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1999 bis 2004; **Veterinärmedizin**

10 Literatur

- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In Holland, P.W., Wainer, H. (Eds.), Differential Item Functioning. Hillsdale, NJ: Erlbaum.
- Beller, M., Gafni, N. (1995). Translated Versions of Israel's interuniversity Psychometric Entrance Test (PET). In T. Oakland & R.K. Hambleton (Eds.), International Perspectives of Academic Assessment, S.207-218. Boston: Kluwer.
- Beller, M. (1996). Translating, equating and validating Scholastic Aptitude Tests: The Israeli Case. In Hänsgen, K.-D., Hofer, R. & Ruefli, D. (Hrsg). Eignungsdiagnostik und Medizinstudium, (S. 14-29), Bericht 2. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.
- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (Second Edition). Hillsdale, NJ: Erlbaum.
- Cook, L.L. (1998). Can Scores Obtained on Test Given in Different Languages to Examinees of Different Cultures be Equally Valid? ICAP San Francisco.
- Deidesheimer Kreis (1997). Hochschulzulassung und Studieneignungstests: studienfeldbezogene Verfahren zur Feststellung der Eignung für Numerus Clausus und andere Studiengänge. Göttingen, Zürich: Vandenhoeck und Ruprecht.
- Ebach, J., Trost, G. (1997). Admission to Medical Schools in Europe. Lengerich: Pabst.
- Hänsgen K-D, Spicher B. (2002). Numerus Clausus: Finden wir mit dem «Eignungstest für das Medizinstudium» die Geeigneten? Schweizerische Ärztezeitung / Bulletin des médecins suisses / Bollettino dei medici svizzeri - 2002; 83(31):1653-1660. <http://www.saez.ch/pdf/2002/2002-31/2002-31-842.PDF>
- Hänsgen K-D, Spicher B. (2002). Numerus Clausus: Numerus Clausus: le « test d'aptitudes pour les études de médecine » (AMS) permet-il de trouver les personnes les plus aptes? Schweizerische Ärztezeitung / Bulletin des médecins suisses / Bollettino dei medici svizzeri - 2002; 83 (47) 2562 – 2569. <http://www.saez.ch/pdf/2002/2002-47/2002-47-1144.PDF>
- Hänsgen, K.-D., Spicher, B. (2000). Zwei Jahre Numerus Clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 1: Erfahrungen mit dem EMS als Zulassungskriterium. Schweizerische Ärztezeitung Heft 12 S. 666 – 672.
- Hänsgen, K.-D., Spicher, B. (2000). Zwei Jahre Numerus Clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 2: EMS und Chancengleichheit. Schweizerische Ärztezeitung Heft 13 S. 723-730.

- Hänsgen, K.-D., Hofer, R. & Ruefli, D. (1995a). Un test d'aptitudes aux études de médecine est-il faisable en Suisse? Bulletin des médecins suisses, 7, S. 267 - 274.
- Hänsgen, K.-D., Hofer, R. & Ruefli, D. (1995b). Der Eignungstest für das Medizinstudium in der Schweiz. Schweizerische Ärztezeitung, 37, S. 1476 - 1496.
- Longford, Holland & Thayer, (1993). Stability of the MH D-DIF Statistics Across Populations. In P. W. Holland & H. Wainer (Ed.) Differential Item Functioning, S. 171 – 196. Lawrence Erlbaum Associates, New Jersey.
- Oswald, U. (1999). Der Eignungstest 1998 für das Medizinstudium. Schweizerische Ärztezeitung 80, S. 1313 – 1317.
- Trost, G. (Hrsg.) (1994). Test für Medizinische Studiengänge (TMS): Studien zur Evaluation (18. Arbeitsbericht). Bonn: ITB.
- Trost, G. (Hrsg.) (1995). Test für Medizinische Studiengänge (TMS): Studien zur Evaluation (19. Arbeitsbericht). Bonn: ITB.
- Trost, G. (Hrsg.) (1996). Test für Medizinische Studiengänge (TMS): Studien zur Evaluation (20. Arbeitsbericht). Bonn: ITB.
- Trost, G. (Hrsg.) (1997). Test für Medizinische Studiengänge (TMS): Studien zur Evaluation (21. Arbeitsbericht). Bonn: ITB.
- Trost, G., Blum, F., Fay, E., Klieme, E., Maichle, U., Meyer, M. & Nauels, H.-U. (1998). Evaluation des Tests für Medizinische Studiengänge (TMS): Synopse der Ergebnisse. Bonn: ITB.
- ZVS Press: Pressedienst der Zentralstelle für die Vergabe von Studienplätzen (1998). Deutlicher Anstieg der Bewerbungen für das Medizin-Studium. Dortmund: ZVS.

10.1 Originaltest zur Information und Vorbereitung

- Institut für Test- und Begabungsforschung (Hrsg.). (1995). Test für medizinische Studiengänge (Aktualisierte Originalversion 2). Herausgegeben im Auftrag der Kultusminister der Länder der BRD. 4. Auflage. Göttingen: Hogrefe.
- Centre pour le développement de tests et le diagnostic, Université de Fribourg (Suisse) en collaboration avec l'Institut für Test- und Begabungsforschung, Bonn, Allemagne (Editeur). (1996). Le test d'aptitudes pour les études de médecine. Adaptation française de la version originale dans son intégralité. Göttingen: Hogrefe.
- Centre pour le développement de tests et le diagnostic, Università di Friborgo (Svizzera) in collaborazione con l'Institut für Test- und Begabungsforschung, Bonn, Germania (Editore). (1996). Il test attitudinale per lo studio della medicina. Adattamento italiano di una versione originale completa. Göttingen: Hogrefe.

10.2 Frühere Berichte des ZTD

Alle Berichte auch im Internet: <http://www.unifr.ch/ztd/ems>

Bericht 9 (2003)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2003 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 8 (2002)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2002 <i>Bericht über die Durchführung und Ergebnisse, Zusammenfassender Vergleich 1998 bis 2002 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 7 (2001)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2001 <i>Bericht über die Durchführung und Ergebnisse, Vorhersage von Prüfungserfolg Zusammenfassender Vergleich 1998 bis 2001 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 6 (2000)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2000 <i>Bericht über die Durchführung und Ergebnisse, Evaluation der Vorhersage von Prüfungserfolg Zusammenfassender Vergleich 1998 bis 2000 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 5 (1999)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 1999 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 4 (1998)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 1998 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>

Bericht 3 (1997) - Rapport 3 (1997)	
<p>Eignungsprüfung für das Medizinstudium - Kriterien und Testverfahren <i>Bericht über das Internationale Symposium in Bern am 8. November 1996</i> Hrsg. von K.-D. Hänsgen und N. Ischi</p>	<p>L'examen d'aptitude aux études de médecine en Suisse - Critères et procédés d'application du test <i>Rapport sur le Symposium international à Berne</i> Editeurs: K.-D. Hänsgen; N. Ischi</p>
<p>Günter Trost Eignungskriterien bei der Zulassung zum Medizinstudium in Europa: Ergebnisse einer Erhebung Ingemar Wedman & Widar Henriksson The Swedish Scholastic Aptitude Test. Research and main findings John L. Hackett The Medical College Admission Test (MCAT) - its use in U.S. and Canada and some results of validation Piet J. Janssen Admission to the study in medicine in Belgium: two 'different' solutions to the 'same' problem; reflections of a Flemish school psychologist</p>	<p>Klaus-Dieter Hänsgen Der Eignungstest für das Medizinstudium in der Schweiz als Instrument für die Studienzulassung Rainer Hofer & Klaus-Dieter Hänsgen Die Trainierbarkeit von Testleistungen im Zusammenhang mit einem Eignungstest für das Medizinstudium Urs Schallberger Anforderungen an das Zulassungsverfahren für das Medizinstudium in der Schweiz: Leitlinien</p>
Bericht 2 (1996) - Rapport 2 (1996)	
<p>Eignungsdiagnostik und Medizinstudium <i>Symposiumsbericht</i> Hrsg. von K.-D. Hänsgen, R. Hofer und D. Ruefli</p>	<p>Diagnostic d'aptitudes et études de médecine <i>Rapport d'un symposium</i> Edité par K.-D. Hänsgen, R. Hofer et D. Ruefli</p>
<p>Klaus-D. Hänsgen Vorwort: Eignungstests und Medizinstudium Christina Stage Experiences with the Swedish Scholastic Aptitude Test Michal Beller Translating, equating and validating Scholastic Aptitude Tests: The Israeli Case Eckhard Klieme Erfolgsprognose in medizinischen Studiengängen - Zur Validität des Tests für medizinische Studiengänge und anderer Auswahlinstrumente</p>	<p>Günter Trost Testergebnisse versus Schulnoten als Auswahlkriterien: Paternoster-Effekt, Filter-Effekt, Kosten-Nutzen-Effekte und Auswirkungen auf die Fairness der Zulassung Urs Schallberger Nutzen, Fairness, Validität und Akzeptanz von Selektionsverfahren Rainer Hofer, Daniel Ruefli & Klaus-D. Hänsgen Der „Test des Tests“ - Ergebnisse eines Probelaufs des Eignungstests in der Schweiz in deutscher und französischer Sprache</p>
Bericht 1 (1995) - Rapport 1 (1995)	
<p>Der Eignungstest für das Medizinstudium in der Schweiz - ein Probelauf Rainer Hofer, Daniel Ruefli & Klaus-D. Hänsgen</p>	<p>Le test d'aptitudes pour les études de médecine en Suisse - Epreuve d'essai Rainer Hofer, Daniel Ruefli & Klaus-D. Hänsgen</p>

11 Anhang zum Bericht 2004

11.1 Beispielaufgaben für die Untertests

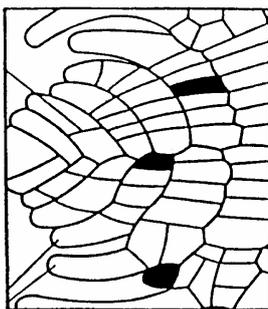
Nachfolgend wird pro Untertest eine Beispielaufgabe dargestellt. So können lediglich die Prinzipien der Aufgabenstruktur verdeutlicht werden – die Aufgaben unterscheiden sich innerhalb jedes Untertests bezüglich des Schwierigkeitsgrades und der Anforderung.

11.1.1 Untertest: Muster zuordnen

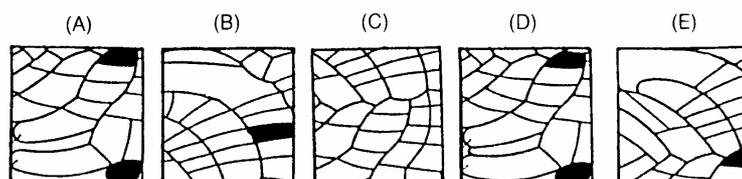
In diesem Untertest wird die Fähigkeit geprüft, Ausschnitte in einem komplexen Bild wiederzuerkennen. Dazu werden pro Aufgabe ein Muster und je fünf Musterausschnitte (A) bis (E) vorgegeben. Die Testteilnehmerin oder der Testteilnehmer soll herausfinden, welcher dieser fünf Musterausschnitte an irgendeiner beliebigen Stelle deckungsgleich und vollständig auf das Muster gelegt werden kann.

Ein Beispiel dazu:

Muster



Musterausschnitte



In den meisten Aufgaben dieser Art heben sich die vier nicht deckungsgleichen Musterausschnitte dadurch vom Muster ab, dass Details entweder hinzugefügt oder weggelassen sind. Zugleich stellt dieser Untertest Anforderungen an die Schnelligkeit der Bearbeitung.

In durchschnittlich 55 Sekunden je Aufgabe muss die Testperson die richtige Lösung herausgefunden haben, dass beispielsweise in der obigen Aufgabe nur der Musterausschnitt (A) deckungsgleich mit einem Teil des Musters ist, und zwar in dessen unterem Bereich, etwa in der Mitte.

11.1.2 Untertest: Medizinisch-naturwissenschaftliches Grundverständnis

Hier wird das Verständnis für Fragen der Medizin und der Naturwissenschaften geprüft. Der Text könnte so in einem Lehrbuch stehen. Wichtig für das Verständnis dieser Textpassage ist, ob daraus bestimmte logische Schlüsse gezogen werden können. Alle Fakten, die für die Beantwortung der Aufgabe notwendig sind, stehen im Text – spezielles medizinisches Vorwissen ist nicht erforderlich. Dieses wichtige Prinzip findet sich bei allen Untertests und ist verantwortlich für die geringe Trainierbarkeit der Aufgabenlösung.

Im Kindesalter kann das Zentrum für Sprache, Spracherwerb und Sprachverständnis noch in der linken oder in der rechten Hälfte (Hemisphäre) des Gehirns in einem umschriebenen Hirnrindengebiet (sog. Sprachregion) angelegt werden. Spätestens im zwölften Lebensjahr sind die sprachlichen Fähigkeiten jedoch fest in einer der beiden Hemisphären verankert, und zwar bei den Rechtshändern in der Regel links, bei den Linkshändern in der Mehrzahl ebenfalls links, zum Teil aber auch rechts; die korrespondierende Region der Gegenseite hat zu diesem Zeitpunkt bereits andere Funktionen fest übernommen. Welche der nachfolgenden Aussagen lässt bzw. lassen sich aus diesen Informationen ableiten?

Bei irreversiblen Hirnrindenverletzungen im Bereich der sogenannten Sprachregion der linken Hemisphäre ...

- I. kommt es bei erwachsenen Linkshändern in der Regel zu keinen wesentlichen Sprachstörungen.
 - II. kommt es bei einem Vorschulkind in der Regel zu einer bleibenden Unfähigkeit, die Muttersprache wieder zu erlernen.
 - III. ist bei zwanzigjährigen Rechtshändern die Fähigkeit, eine Sprache zu erlernen, in der Regel verloren gegangen.
- (A) Nur Ausfall I ist zu erwarten.
 - (B) Nur Ausfall II ist zu erwarten.
 - (C) Nur Ausfall III ist zu erwarten.
 - (D) Nur die Ausfälle I und III sind zu erwarten.
 - (E) Nur die Ausfälle II und III sind zu erwarten.

Bei diesem Aufgabentyp folgen nach der Schilderung des Sachverhalts in der Regel drei oder fünf Aussagen in Form von Behauptungen. Die Testperson muss sich dabei entscheiden, ob sich die Aussagen aus den im Aufgabentext enthaltenen Informationen ableiten lassen. Dazu sind keine speziellen Sachkenntnisse erforderlich. Die korrekte Beurteilung der einzelnen Aussagen setzt das Verstehen des Sachverhalts voraus sowie die Fähigkeit, Schlussfolgerungen aus den im Text enthaltenen Informationen zu ziehen. Konkret lässt sich die Aufgabe, unter Berücksichtigung des unterstrichenen Textes, folgendermassen lösen:

- I. Da bei der Mehrzahl der Linkshänder die Sprachregion in der linken Hemisphäre liegt, müssen sie also mit einer Sprachstörung rechnen, weshalb Aussage I falsch ist.
- II. Da es im Kindesalter noch offen ist, in welcher Hälfte des Gehirns die Sprachregion angelegt wird, besteht für ein Vorschulkind immer noch die Möglichkeit, die Muttersprache wieder zu erlernen. Die Sprachregion wird dann in der rechten Hälfte der Hemisphäre angelegt. Somit ist Aussage II ebenfalls falsch.

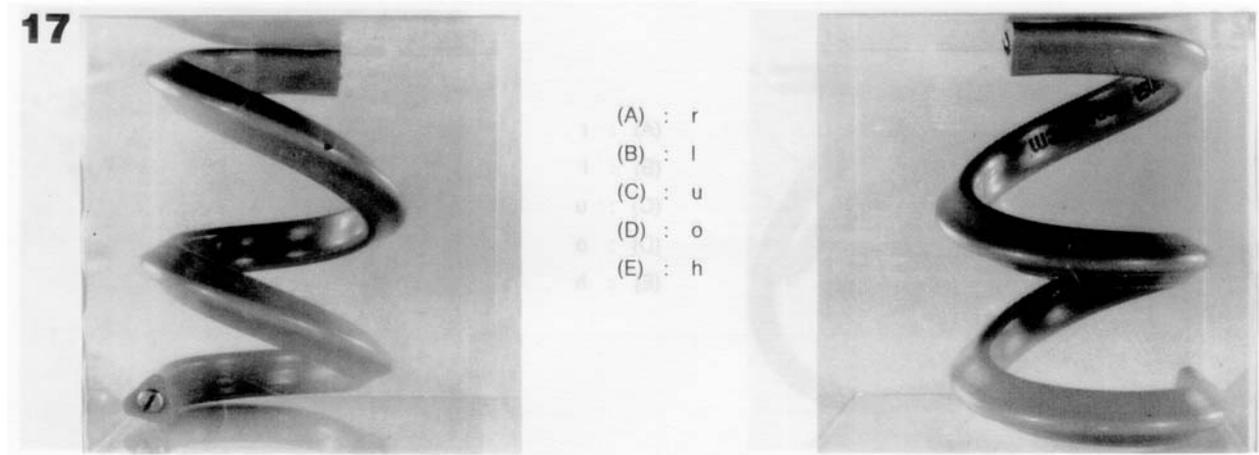
- III. Da spätestens im zwölften Lebensjahr die Sprachregion bei Rechtshändern in der Regel fest in der linken Hälfte des Gehirns liegt, ist bei zwanzigjährigen Rechtshändern zu erwarten, dass sie die Fähigkeit, eine Sprache zu erlernen, verloren haben. Die Aussage III ist darum richtig.

Nach dieser Analyse des Textes ist es offensichtlich, dass die Antwort (C) richtig ist.

11.1.3 Untertest: Schlauchfiguren

Die folgenden Aufgaben prüfen das räumliche Vorstellungsvermögen – eine Funktion, die beispielsweise für das Verständnis von Röntgenbildern wichtig ist. Während des Studiums werden zahlreiche eigentlich dreidimensional zu betrachtende Strukturen und Vorgänge in zweidimensionalen Abbildungen vermittelt.

Jede Aufgabe besteht aus zwei Abbildungen eines durchsichtigen Würfels, in dem sich ein, zwei oder drei Kabel befinden. Die erste Abbildung (links) zeigt stets die Vorderansicht des Würfels; auf dem rechten Bild daneben, in welchem derselbe Würfel noch einmal abgebildet ist, soll die Testteilnehmerin oder der Testteilnehmer herausfinden, ob die Abbildung die Ansicht von rechts (r), links (l), unten (u), oben (o) oder von hinten (h) zeigt.



Hier sehen Sie den Würfel von vorne!

Hier sehen Sie den Würfel von ? (hinten!)

11.1.4 Untertest: Quantitative und formale Probleme

Mit Hilfe dieses Untertests wird die Fähigkeit überprüft, im Rahmen medizinischer und naturwissenschaftlicher Fragestellungen mit Zahlen, Grössen, Einheiten und Formeln richtig umzugehen. Diese Anforderung dürfte für mehrere Fächer des Grundlagenstudiums der Medizin bedeutsam sein.

Zum Beispiel:

Eine Proteineinheit (BE) ist definiert als diejenige Nahrungsmenge in Gramm, die 12 Gramm Kohlenhydrate enthält. Bei der Verbrennung von 1 g Kohlenhydraten im Organismus werden 16 Kilojoule (kJ) an Energie frei. Ein Patient, der auf Diät gesetzt ist, soll pro Tag 4800 kJ zu sich nehmen, ein Fünftel davon in Kohlehydraten.

Wie viele sind dies täglich?

- (A) 60 BE
- (B) 25 BE
- (C) 6 BE
- (D) 5 BE
- (E) 0,5 BE

Bei solchen Fragen werden die Kenntnisse der Mittelstufen-Mathematik, nicht jedoch Lerninhalte vorausgesetzt. Der Patient soll ein Fünftel von 4800 kJ in Kohlehydraten zu sich nehmen, das sind also 960 kJ. Dividiert man diese Zahl durch 16, so erhält man die Anzahl g Kohlehydrate, nämlich 60 g, die es braucht, damit 960 kJ an Energie frei werden. Umgerechnet in Broteinheiten müssen die 60 g Kohlehydrate noch einmal durch 12 dividiert werden und das gibt 5 BE. Somit ist bei dieser Frage die Antwort (D) richtig.

11.1.5 Untertest: Konzentriertes und Sorgfältiges Arbeiten

Bei diesem Untertest soll die Fähigkeit, rasch, sorgfältig und konzentriert zu arbeiten, gemessen werden. Dabei sollen möglichst alle b, die mit zwei Querstrichen versehen sind, die entweder beide unten, beide oben oder je einer unten und oben angebracht sind, markiert werden. Die Lösungsmenge ist ebenso wichtig wie die Fehlerfreiheit der Bearbeitung. Dieser Test ist trainierbar – in der Test-Info wird darauf hingewiesen, diesen Untertest vor der Testabnahme mehrfach zu üben.

b b b

Diese Buchstaben b mit zwei Querstrichen sind eingestreut unter b mit einem, drei oder vier Querstrichen sowie unter q mit einem oder mehreren Querstrichen. Im folgenden Beispiel wären also das 1., 4., 6., 8., 9. und 13. Zeichen zu markieren.

b q b b b b b b b q b b b q

Hinweis: Dieser Untertest ist nur ein Beispiel für eine mögliche Anforderung. Es werden jährlich wechselnde Zeichen und Anforderungen verwendet, die vorher nicht bekannt gegeben werden. So sind Lerneffekte deutlich begrenzt.

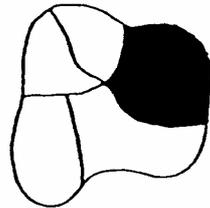
11.1.6 Untertest: Figuren lernen

Für beide nachfolgende Gedächtnistests wird nach der Mittagspause das Material zum Einprägen ausgeteilt. Vor der Abfrage des Gelernten wird der Untertest „Textverständnis“ bearbeitet, damit liegt die Zeit des Behaltens der gelernten Inhalte über einer Stunde. Gedächtnisleistungen sind wichtige Voraussetzungen für Studienerfolg.

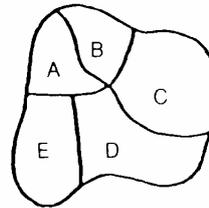
Der Untertest „Figuren lernen“ prüft, wie gut man sich Einzelheiten von Gegenständen einprägen und merken kann.

Ein Beispiel:

Gezeigte Figur zum Einprägen



Gezeigte Figur beim Abfragen



Die Testperson hat vier Minuten Zeit, um sich 20 solcher Figuren einschliesslich der Lage der schwarzen Flächen einzuprägen. Nach ca. einer Stunde muss sie angeben können, welcher Teil der Abbildung geschwärzt war, und dies direkt auf dem Antwortbogen eintragen. Die Lösung ist natürlich C.

11.1.7 Untertest: Fakten lernen

Analog dem Prinzip beim „Figuren lernen“ sollen hier Fakten eingepägt und behalten werden, die ebenfalls nach der gleichen Zwischenzeit abgefragt werden. Dabei werden 15 Patienten vorgestellt, von denen jeweils der Name, die Altersgruppe, Beruf und Geschlecht, ein weiteres Beschreibungsmerkmal (z.B. Familienstand) sowie die Diagnose erfahren wird. Ein Beispiel für eine derartige Fallbeschreibung ist:

Lemke, 30 Jahre, Dachdecker, ledig, Schädelbasisbruch

Eine Frage zum obigen Beispiel könnte z.B. lauten:

Der Patient mit dem Schädelbasisbruch ist von Beruf ...

- (A) Installateur
- (B) Lehrer
- (C) Dachdecker
- (D) Handelsvertreter
- (E) Physiker

11.1.8 Untertest: Textverständnis

Mit Hilfe dieses Untertests wird die Fähigkeit geprüft, umfangreiches und komplexes Textmaterial aufzunehmen und zu verarbeiten. Die Texte sind inhaltlich und grammatikalisch anspruchsvoll – sie können unter Nutzung von Notizen und Unterstreichungen erarbeitet werden. Die Abfrage erfolgt wiederum über die Auswahl einer richtigen oder falschen Aussage aus fünf vorgegebenen Aussagen. Diese Texte waren vor allem beim Übersetzen anspruchsvoll – zur Schwierigkeit gehören nicht nur die Inhalte, sondern auch die Satzstruktur.

Ein Beispiel:

Zu den Aufgaben der Schilddrüse gehören Bildung, Speicherung und Freisetzung der jodhaltigen Hormone Trijodthyronin (T_3) und Thyroxin (T_4). In der Schilddrüse befinden sich zahlreiche Hohlräume, Follikel genannt, deren Wände von einer Schicht sogenannter Epithelzellen gebildet werden. Diese Follikel sind mit einer Substanz gefüllt, in der die Hormone T_3 und T_4 als inaktive Speicherformen enthalten sind. Beim Menschen ist in den Follikeln so viel T_3 und T_4 gespeichert, dass der Organismus damit für etwa 10 Monate versorgt werden kann.

Das für die Hormonbildung erforderliche Jod entstammt der Nahrung und wird von den Epithelzellen als Jodid aus dem Blut aufgenommen. Die Jodidaufnahme erfolgt an der äusseren Zellmembran der Epithelzellen durch eine sogenannte Jodpumpe. Diese wird durch ein Hormon aus der Hirnanhangsdrüse, das TSH, stimuliert und kann pharmakologisch durch die Gabe von Perchlorat gehemmt werden. Ferner gibt es erbliche Schilddrüsenerkrankungen, bei deren Vorliegen die Jodpumpe nicht funktioniert.

Bei Gesunden wird das in die Epithelzellen aufgenommene Jodid im nächsten Schritt unter dem Einfluss eines Enzyms in freies Jod umgewandelt und in die Follikel abgegeben. Die Aktivität dieses Enzyms kann ebenfalls pharmakologisch gehemmt werden.

Die letzten Schritte der Hormonbildung finden in den Follikeln, also ausserhalb der einzelnen Epithelzellen, statt. In dort vorhandene sogenannte Tyrosin-Reste (des Thyreoglobulins) wird zunächst ein Jodatomb eingebaut. So entstehen Monojodtyrosin-Reste (MIT), von denen ein Teil durch die Bindung je eines weiteren Jodatoms in Dijodtyrosin-Reste (DIT) umgewandelt wird. Durch die Verknüpfung von je zwei DIT-Resten entsteht schliesslich T_4 , während aus der Verbindung je eines MIT-Restes mit einem DIT-Rest T_3 hervorgeht. T_3 und T_4 werden dann in den Follikeln gespeichert und bei Bedarf über die Epithelzellen ins Blut freigesetzt.

Diese Freisetzung von T_3 und T_4 ins Blut (Sekretion) wird über die Hirnanhangsdrüse und den Hypothalamus, einen Teil des Zwischenhirns, gesteuert: Das erwähnte Hormon TSH stimuliert ausser der Bildung auch die Sekretion von T_3 und T_4 ; es ist hinsichtlich seiner eigenen Sekretionsrate jedoch abhängig von der Stimulation durch das hypothalamische Hormon TRH. Die TRH-Sekretion wiederum wird z.B. durch Kälte stimuliert, während Wärme hemmend wirken kann. Neben diesen übergeordneten Steuerungsmechanismen existiert noch ein sogenannter Rückkopplungsmechanismus: Eine hohe Konzentration von T_3 und T_4 im Blut hemmt die TSH- und die TRH-Sekretion, eine niedrige Konzentration stimuliert sie. Bei den an der Steuerung der Schilddrüsenhormon-Sekretion beteiligten Arealen von Hirnanhangsdrüse und Hypothalamus können krankheitsbedingte Störungen auftreten, die zu einer Über- oder Unterfunktion der Schilddrüse führen.

Eine der Hauptwirkungen von T_3 und T_4 ist die Beeinflussung des Energieumsatzes durch eine Steigerung des Sauerstoffverbrauchs in stoffwechselaktiven Organen. Entsprechend senkt eine zu niedrige Konzentration der beiden Hormone im Blut (Hypothyreose) den Energieumsatz bzw. die Stoffwechselaktivität unter den normalen Wert, während bei einer zu hohen Konzentration (Hyperthyreose) die Stoffwechselaktivität gesteigert wird. Die Hormone T_3 und T_4 können ebenso wie TSH und TRH für diagnostische und therapeutische Zwecke synthetisch hergestellt werden.

Auf einen solchen Text folgen Fragen, die sich ausschliesslich auf im Text vorhandene Inhalte beziehen; eine Frage mit niedrigem Schwierigkeitsgrad ist zum Beispiel so formuliert:

Welcher der folgenden Vorgänge gehört nicht zu den im Text beschriebenen Schritten, die zur Bildung von T_3 führen?

- (A) Transport von Jod aus den Epithelzellen in die Follikel
- (B) Umwandlung von Jod in Jodid in den Follikeln
- (C) Transport von Jodid aus dem Blut in die Epithelzellen
- (D) Verknüpfung von MIT- und DIT-Resten in den Follikeln
- (E) Verknüpfung von Jod und Tyrosin-Resten in den Follikeln

Für die Beantwortung dieser Frage ist das Verständnis der im obigen Text unterstrichenen Stellen wichtig (im Original sind selbstverständlich keine Hervorhebungen). Der Text sagt nichts über eine Umwandlung von Jod in Jodid in den Follikeln aus, und auch der umgekehrte Prozess, die Umwandlung von Jodid in Jod, findet nicht in den Follikeln statt, sondern in den Epithelzellen. Somit gehört der Vorgang (B) nicht zu den vom Text beschriebenen Schritten, die zur Bildung von T_3 führen.

11.1.9 Untertest: Diagramme und Tabellen

Mit dieser Aufgabengruppe wird die Fähigkeit geprüft, Diagramme und Tabellen richtig zu analysieren und zu interpretieren. In dieser Form werden während des Studiums zahlreiche Zusammenhänge vermittelt. Eine Aufgabe dazu:

Die folgende Tabelle beschreibt die Zusammensetzung und den Energiegehalt von vier verschiedenen Milcharten. Unter Energiegehalt der Milch verstehen wir dabei die Energiemenge, gemessen in Kilojoule (kJ), welche 100 Gramm (g) Milch dem Organismus ihres Konsumenten liefern können.

Milchart	Eiweiss	Fett	Milchzucker	Salze	Energiegehalt
menschliche Muttermilch	1,2 g	4,0 g	7,0 g	0,25 g	294 kJ
Vollmilch	3,5 g	3,5 g	4,5 g	0,75 g	273 kJ
Magermilch	3,3 g	0,5 g	4,5 g	0,75 g	160 kJ
Buttermilch	3,0 g	0,5 g	3,0 g	0,55 g	110 kJ

Welche Aussage lässt sich aus den gegebenen Informationen nicht ableiten?

- (A) Menschliche Muttermilch enthält mehr als doppelt soviel Milchzucker wie Buttermilch.
- (B) Vollmilch enthält im Vergleich zur menschlichen Muttermilch etwa die dreifache Menge an Salzen und Eiweiss.
- (C) Zur Aufnahme der gleichen Energiemenge muss ein Säugling fast dreimal soviel Buttermilch wie Muttermilch trinken.
- (D) Der Unterschied zwischen Magermilch und Vollmilch ist bei der Mehrzahl der aufgeführten Merkmale geringer als der Unterschied zwischen Magermilch und Buttermilch.
- (E) Der Eiweissgehalt der Milch ist für den Energiegehalt von entscheidender Bedeutung.

Wie bei den Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“ und „Textverständnis“ sind auch hier zur Lösung dieser Aufgabe keine speziellen naturwissenschaftlichen, medizinischen oder statistischen Kenntnisse erforderlich. Die richtige Lösung lässt sich allein aus der jeweils graphisch oder tabellarisch dargebotenen Information und dem zugehörigen Aufgabentext ableiten. Aus den angegebenen Werten ist kein systematischer Zusammenhang zwischen Eiweiss- und Energiegehalt ableitbar, so dass die Aussage (E) nicht abgeleitet werden kann.

Diese Beispielaufgaben aus den neun Untertests zeigen, dass es hier um Problemstellungen geht, die auch aus einem Lehrbuch des Grundstudiums Medizin stammen könnten. In den Aufgabenstellungen sind alle Informationen enthalten, die man zum Lösen benötigt. Das Problem ist zunächst zu erkennen, die Information genau zu analysieren und eine Lösung zu finden.

11.2 Weitere Detailergebnisse

Gesamt	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen	2	1	1	8	15	26	55	98	142	181	197	176	142	136	83	58	34	10	5	3	
Med.-nat. Grundv.		2	5	15	34	49	70	108	128	146	150	161	133	108	107	72	44	21	13	5	2
Schlauchfiguren		2	2	9	19	27	32	49	77	94	111	141	156	138	132	98	88	85	55	39	19
Qua.form Probl.		3	13	27	42	73	96	107	136	151	152	144	115	83	77	57	44	25	14	11	3
Textverständnis		1	8	29	66	103	170	154	149	147	137	114	78	85	53	37	19	17	6		
Figuren lernen	1		3	6	3	17	28	46	67	86	101	125	125	149	135	136	97	96	69	46	37
Fakten lernen				5	9	24	36	60	92	112	133	128	135	115	124	112	87	75	48	39	39
Diagr. und Tabellen	1	6	7	10	31	35	62	100	137	120	168	149	141	124	118	66	60	25	9	4	
Kon.sorgf. Arbeiten	35	7	5	10	30	22	38	68	74	118	127	138	130	134	121	84	72	47	37	29	47

Tabelle A 1: Häufigkeitsverteilung der Punkte Gesamt 2004

Deutsch	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen	2	1	1	5	11	21	48	90	126	162	179	164	131	130	80	53	33	10	5	3	
Med.-nat. Grundv.		2	5	11	31	45	65	94	115	132	138	151	118	102	101	68	40	20	12	4	1
Schlauchfiguren		2	1	8	13	24	26	42	67	78	99	129	141	126	127	94	84	83	55	37	19
Qua.form Probl.		2	9	26	37	66	83	97	127	136	140	136	105	74	71	55	42	24	13	10	2
Textverständnis			8	25	59	90	155	139	139	131	128	106	70	81	52	33	17	16	6		
Figuren lernen	1		2	5	3	16	22	39	56	77	90	118	117	132	126	118	91	94	67	44	37
Fakten lernen				4	7	19	30	51	85	104	125	119	119	106	111	105	79	69	47	38	37
Diagr. und Tabellen	1	5	4	10	28	27	53	88	119	107	155	137	131	120	111	63	59	25	8	4	
Kon.sorgf. Arbeiten	33	5	5	9	27	21	35	59	67	107	116	119	118	128	115	78	65	43	35	26	44

Tabelle A 2: Häufigkeitsverteilung der Punkte deutsche Sprachgruppe 2004

Französisch	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen				2	2	2	4	7	10	10	7	5	7	5	2	4	1				
Med.-nat. Grundv.				1		2	5	10	6	10	8	5	7	3	2	3	3	1		1	1
Schlauchfiguren			1	1	2	3	2	3	6	11	6	8	9	5	3	4	2			2	
Qua.form Probl.		1	2		4	6	7	5	6	10	7	2	4	6	1	2	2	1		1	1
Textverständnis		1		1	5	9	10	9	4	10	5	3	4	1	1	3	1	1			
Figuren lernen						1	1	2	8	5	7	3	3	13	5	12	6	1		1	
Fakten lernen					1	2	3	6	3	4	7	1	10	8	8	4	8	3			
Diagr. und Tabellen			1		2	5	5	7	11	5	10	8	5	3	4	1			1		
Kon.sorgf. Arbeiten	2	1			1		1	7	5	8	8	8	7	5	1	2	4	2	1	3	2

Tabelle A 3: Häufigkeitsverteilung der Punkte französische Sprachgruppe 2004

Italienisch	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen				1	2	3	3	1	6	9	11	7	4	1	1	1					
Med.-nat. Grundv.				3	3	2		4	7	4	4	5	8	3	4	1	1		1		
Schlauchfiguren					4		4	4	4	5	6	4	6	7	2		2	2			
Qua.form Probl.			2	1	1	1	6	5	3	5	5	6	6	3	5				1		
Textverständnis				3	2	4	5	6	6	6	4	5	4	3		1	1				
Figuren lernen			1	1			5	5	3	4	4	4	5	4	4	6		1	2	1	
Fakten lernen				1	1	3	3	3	4	4	1	8	6	1	5	3		3	1	1	2
Diagr. und Tabellen		1	2		1	3	4	5	7	8	3	4	5	1	3	2	1				
Kon.sorgf. Arbeiten		1		1	2	1	2	2	2	3	3	11	5	1	5	4	3	2	1		1

Tabelle A 4: Häufigkeitsverteilung der Punkte italienische Sprachgruppe 2004

Human- medizin	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen	2	1	1	5	9	17	44	77	100	139	146	131	111	106	54	43	27	6	5	2	
Med.-nat. Grundv.		2	4	10	26	34	45	67	87	108	121	130	95	88	79	61	36	17	10	5	1
Schlauch- figuren		1	2	7	13	18	18	33	62	68	80	106	114	100	98	80	70	68	42	33	13
Qua.form Probl.		2	7	18	29	46	72	76	103	117	104	98	91	66	62	49	39	21	13	10	3
Textver- ständnis			6	19	42	70	119	106	118	110	99	91	63	68	46	33	18	14	4		
Figuren lernen	1		2	4	2	11	24	38	49	63	77	85	99	120	96	103	73	70	53	26	30
Fakten lernen				2	6	21	19	41	69	83	93	97	101	88	90	90	63	59	34	35	35
Diagr. und Tabellen	1	4	5	7	19	25	39	68	98	87	126	105	117	103	96	50	47	17	8	4	
Kon.sorgf. Arbeiten	27	5	4	6	20	18	26	50	55	81	93	108	93	104	93	70	54	37	26	20	36

Tabelle A 5: Häufigkeitsverteilung der Punkte Humanmedizin 2004

Veterin.- medizin	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen				2	2	7	8	11	24	26	24	27	16	19	18	7	3	2			
Med.-nat. Grundv.				2	2	6	12	23	22	18	16	20	23	10	21	9	6	3	2		1
Schlauch- figuren				1	5	6	7	10	11	15	22	25	19	24	11	13	7	10	7	1	2
Qua.form Probl.		1	5	4	7	15	13	19	12	16	31	26	16	14	8	3	4	1		1	
Textver- ständnis		1		3	11	15	26	31	20	25	18	15	7	11	5	3	1	2	2		
Figuren lernen				1		3	2	4	8	8	17	26	14	14	24	21	14	16	9	11	4
Fakten lernen				1	1	2	12	11	10	14	22	17	22	21	20	9	15	9	6	2	2
Diagr. und Tabellen			1	2	4	4	14	19	20	15	20	31	16	13	13	10	8	5	1		
Kon.sorgf. Arbeiten	3			2	5	3	6	8	10	13	21	15	25	18	18	10	14	4	8	7	6

Tabelle A 6: Häufigkeitsverteilung der Punkte Veterinärmedizin 2004

Zahnmedizin	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen				1	4	2	3	10	18	16	27	18	15	11	11	8	4	2		1	
Med.-nat. Grundv.			1	3	6	9	13	18	19	20	13	11	15	10	7	2	2	1	1		
Schlauchfiguren		1		1	1	3	7	6	4	11	9	10	23	14	23	5	11	7	6	5	4
Qua.form Probl.			1	5	6	12	11	12	21	18	17	20	8	3	7	5	1	3	1		
Textverständnis			2	7	13	18	25	17	11	12	20	8	8	6	2	1		1			
Figuren lernen			1	1	1	3	2	4	10	15	7	14	12	15	15	12	10	10	7	9	3
Fakten lernen				2	2	1	5	8	13	15	18	14	12	6	14	13	9	7	8	2	2
Diagr. und Tabellen		2	1	1	8	6	9	13	19	18	22	13	8	8	9	6	5	3			
Kon.sorgf. Arbeiten	5	2	1	2	5	1	6	10	9	24	13	15	12	12	10	4	4	6	3	2	5

Tabelle A 7: Häufigkeitsverteilung der Punkte Zahnmedizin 2004

	Humanmedizin		Veterinärmedizin		Zahnmedizin			Humanmedizin		Veterinärmedizin		Zahnmedizin	
	Punkt-wert	Test-wert	Punkt-wert	Test-wert	Punkt-wert	Test-wert		Punkt-wert	Test-wert	Punkt-wert	Test-wert	Punkt-wert	Test-wert
30					1		95	19	45		10	3	5
37					1		96	18	47	4	7		5
38					1		97	22	40	4	3	4	4
41	1				1		98	30	37	3	4	2	3
42	1						99	21	52	6	7	2	6
43	1						100	23	21	5	6	2	2
44	1						101	19	42	7	12	4	6
45	1						102	21	45	3	8	6	10
46	1						103	24	48	5	8	4	4
47	1						104	27	43	4	7	2	2
48	1						105	21	34	4	4	2	3
49	1						106	21	31	5	17	1	4
50	1						107	22	34	2	4	1	5
52	1				2		108	19	30	1	4	2	4
53	2				1		109	15	33	3	3	1	5
54	1						110	16	28	10	3		4
56	1				1		111	15	25	7	5	4	2
57	1				2		112	15	15	3	2	4	
58	2		2				113	19	22	1	3	1	
59	2						114	16	9	1	1	2	1
60	3		1				115	14	18	3	2	2	3

Fortsetzung nächste Seite...

Fortsetzung

	Humanmedizin		Veterinärmedizin		Zahnmedizin			Humanmedizin		Veterinärmedizin		Zahnmedizin	
	Punkt-wert	Test-wert	Punkt-wert	Test-wert	Punkt-wert	Test-wert		Punkt-wert	Test-wert	Punkt-wert	Test-wert	Punkt-wert	Test-wert
61	1				1		116	13	14		2	2	3
62	4						117	20	15	3	1	3	2
63	4		3		1		118	16	8	2	2	2	
64	1		1				119	12	11	1	1	2	
65	3				3		120	13	8	1	1	2	2
66	6				2		121	12	4	4			
67	4		1		1		122	7	3		1		
68	4		2		1		123	8	5	2			2
69	7		1				124	11		1			
70	5	2	1		2	4	125	11		2			
71	5	2					126	9	3	1		1	
72	7	2	3		1		127	9	1	1			
73	6	1	1		4		128	9		1		3	
74	6	2	1		2		129	9	1	2			
75	14	1	2		2		130	5				3	
76	12	3	1			3	131	8				1	
77	8	1	3		2		132	7		1		1	
78	8	2	3		4	3	133	6		2			
79	6	4	3	2	3		134	2					
80	12	4	4	1	4	1	135	7		1			
81	17	8	5	3	5	1	136	4					
82	11	4	6	1	3	3	137	4				2	
83	19	10		1	1	3	138	4		1			
84	11	11	4	3	1	1	139	3					
85	11	10	4	1	2	2	140	1					
86	12	13	6	4	4	5	141	1		1			
87	15	6	6	1	1	2	142	2					
88	12	26	4	3	4	2	143	4				1	
89	22	16	5	6	1	6	144	1				1	
90	23	18	5	7	4	7	149	2					
91	22	28	1	11	1	8	150	1					
92	25	30	6	4	4	2	152	1					
93	23	23	3	10	3	6	155	1					
94	17	27		10	1	5							

Tabelle A 8: Punkt- und Testwertverteilung nach Disziplinen 2004

Das Zentrum für Testentwicklung und Diagnostik im Internet:

www.unifr.ch/ztd

Microsoft Internet Explorer
Datei Bearbeiten Ansicht Favoriten Extras ?
Adresse <http://www.unifr.ch/ztd/ems/welcome.htm> Wechseln zu

Eignungstest für das Medizinstudium in der Schweiz
Test d'aptitudes pour les études de médecine en Suisse
Test attitudinale per lo studio della medicina in Svizzera

EMS
AMS

Kandidatinnen & Kandidaten

- Termine & Ablauf
- Wie vorbereiten?
- Fragen & Antworten
- CRUS (mehr Info)

[Dokumentation "Was ist der EMS?"](#)



Forschung zum EMS

- deutsch
- français

Weitere Informationen der CRUS

CRUS
Rektorenkonferenz der Schweizer Universitäten
Conférence des Recteurs des Universités Suisses
Conferenza dei Rettori delle Università Svizzere

[E-MAIL : ZTD@unifr.ch](mailto:ZTD@unifr.ch)

Information des ZTD für Teilnehmende am EMS

➤ [Rapport statistique en français](#)

Anmeldung zum Studium 2005

Im November 2004 beginnt die Anmeldefrist zur Aufnahme eines Medizinstudiums im Jahre 2005. Informationen zum Anmeldeverfahren finden Sie [bei der CRUS](#). Über die Notwendigkeit eines **Numerus Clausus** (Zulassungsbeschränkung - welche Universitäten, welche Disziplinen) als Notmassnahme und die **Durchführung eines EMS 2005** wird **erst im März 2005 entschieden**. Alle Angemeldeten erhalten rechtzeitig Informationsmaterial zum EMS, um sich auf den EMS vorzubereiten. Wie die Ergebnisse des Vorbereitungsreports (s.u.) zeigen, ist die dann zur Verfügung stehende Zeit bis zum Testtermin noch ausreichend, sich optimal vorzubereiten.

Wichtig: es wird 2005 beim EMS weitere Änderungen geben, über die alle Angemeldeten rechtzeitig und ausführlich informiert werden. Bitte beachten Sie, dass wir für Informationen von Dritten (etwa den professionellen Trainingsanbietern) keine Verantwortung tragen(s.u.). Ausschlaggebend sind NUR die Informationen, die Sie von der CRUS oder dem ZTD auf offiziellem Wege erhalten.

Falls Sie heute schon mehr über den EMS wissen möchten: im linken roten Menüteil sowie nachfolgend finden Sie mehr. ➤ [Hier finden Sie Hinweise zum weiteren Ablauf.](#)

Session 2004 erfolgreich

Am 2.7. 2004 wurde der EMS in den Testorten Basel, Bellinzona (italienisch) Bern, Freiburg (französisch) und Zürich erfolgreich durchgeführt.

Die aktualisierten Zahlen für die Bewerbungssituation 2004 finden Sie im ➤ [Statistischen Bericht 2004](#) als PDF-Dokument.
(version française ➤ [Rapport statistique en français](#))

Internet