

# **EMS • Eignungstest für das Medizinstudium in der Schweiz • 2003**

*Bericht 9 über die Durchführung und Ergebnisse 2003  
Trends und Schlussfolgerungen aus sechs Jahren Testanwendung*

*avec un résumé en français*

**K.-D. Hänsgen und B. Spicher**

**Bericht 9 (2003)**

Hänsgen, Klaus-Dieter; Spicher, Benjamin (2003).

*EMS Eignungstest für das Medizinstudium in der Schweiz 2003*

Berichte des Zentrums für Testentwicklung, Nummer 9, im Auftrag der Schweizerischen Universitätskonferenz (SUK)

Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik, 2003

Unter Mitarbeit von Katharina Stress (Lektorat),

Mauro Bernasconi, Steve Fürst, Sébastien Simonet, Maurizio Strazzeri

Redaktion: Tanja Läng

*Die Mitglieder des Beirates „Eignungstest“ der Schweizerischen Universitätskonferenz zum Zeitpunkt der Testabnahme 2003:*

*Dr. N. Ischi (Vorsitzender, Bern); Prof. Dr. R. Bloch (Bern); Prof. Dr. P. Groscurth (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. Schallberger (Zürich); PD. Dr. Th. Schläpfer (Bern); Prof. Dr. U. M. Spornitz (Basel); Prof. Dr. G. Trost (Bonn/Deutschland); K. Wechsler (Bern).*

---

© ZTD Zentrum für Testentwicklung und Diagnostik 2003

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](mailto:ztd@unifr.ch)

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

Druck: Cric Print, Freiburg/Schweiz

## Inhaltsverzeichnis

<b>1</b>	<b>ZUSAMMENFASSUNG .....</b>	<b>5</b>
<b>2</b>	<b>RESUME .....</b>	<b>8</b>
<b>3</b>	<b>AKTUELLE DISKUSSIONEN ZUM EMS.....</b>	<b>11</b>
<b>4</b>	<b>NUMERUS CLAUSUS (NC) UND MEDIZINSTUDIUM.....</b>	<b>13</b>
<b>5</b>	<b>ANMELDUNG ZU MEDIZINSTUDIUM UND TEST 2003.....</b>	<b>15</b>
5.1	Anmeldungen, Testantritte und Zulassungen.....	15
5.2	Grösse der Testlokale .....	19
5.3	Testorte und Wunschuniversitäten.....	19
5.4	Testorte und Wohnkantone .....	21
5.5	Testabsolvierung nach Alter und Geschlecht .....	24
5.6	Übernahme des Testergebnisses aus 2002 .....	26
5.7	Erneute Testteilnahme 2003 .....	27
5.8	Sprachgruppen.....	29
5.9	Alter und Maturitätsjahr .....	31
<b>6</b>	<b>BESCHREIBUNG DES VERWENDETEN EIGNUNGSTESTS.....</b>	<b>32</b>
6.1	Aufbau des Tests.....	32
6.2	Berechnung der Werte .....	34
<b>7</b>	<b>TESTANWENDUNG IN DER SCHWEIZ 2003.....</b>	<b>35</b>
7.1	Organisation der Testabnahme.....	35
7.2	Verteilungsprüfung .....	35
7.3	Vergleich der Testfassungen 1998 bis 2003.....	39
7.4	Äquivalenz der Sprachversionen .....	42
7.4.1	Sprachvergleich für die Untertests .....	44
7.4.2	Darstellung des Korrekturverfahrens .....	47
7.4.3	Leistungen in identischen Testaufgaben 1998 – 2003 .....	61
7.5	Vergleichbarkeit der Testlokale .....	62
7.6	Vergleich der Geschlechter.....	64
7.7	Vergleich der Kantone .....	69
7.7.1	Vergleich 2003 innerhalb der Deutschschweiz .....	70
7.8	Vergleiche für Altersgruppen.....	73
7.9	Vergleiche nach Wunschuniversitäten .....	76

<b>8</b>	<b>ERGEBNISSE ZUR TESTGÜTE .....</b>	<b>78</b>
8.1	Zuverlässigkeit .....	78
8.2	Binnenstruktur .....	80
8.3	Item-Trennschärfen .....	86
8.4	Konzentriertes und sorgfältiges Arbeiten.....	90
<b>9</b>	<b>ZUSAMMENFASSUNG KANTONE 1998 BIS 2003 .....</b>	<b>94</b>
<b>10</b>	<b>LITERATUR .....</b>	<b>99</b>
10.1	Originaltest zur Information und Vorbereitung.....	100
10.2	Frühere Berichte des ZTD.....	101
<b>11</b>	<b>ANHANG ZUM BERICHT 2003 .....</b>	<b>103</b>
<b>11.1</b>	<b>Beispielaufgaben für die Untertests.....</b>	<b>104</b>
11.1.1	Untertest: Muster zuordnen.....	104
11.1.2	Untertest: Medizinisch-naturwissenschaftliches Grundverständnis .....	104
11.1.3	Untertest: Schlauchfiguren.....	106
11.1.4	Untertest: Quantitative und formale Probleme .....	106
11.1.5	Untertest: Konzentriertes und sorgfältiges Arbeiten.....	107
11.1.6	Untertest: Figuren lernen .....	107
11.1.7	Untertest: Fakten lernen .....	108
11.1.8	Untertest: Textverständnis .....	108
11.1.9	Untertest: Diagramme und Tabellen .....	110
<b>11.2</b>	<b>Weitere Detailergebnisse .....</b>	<b>111</b>

# 1 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2003 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Human- und Veterinärmedizin ab Wintersemester 2003/2004. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2002 herangezogen. **Die Ergebnisse sind nicht repräsentativ für Schlussfolgerungen hinsichtlich einer gesamthaften oder regionalen Evaluation von Bildungseinrichtungen und Bildungswegen.**

## Statistik

Wie seit 1998 in der Humanmedizin und seit 1999 auch in der Veterinärmedizin ist ein Numerus Clausus 2003 für Personen notwendig, die sich an den Universitäten Basel, Bern, Freiburg oder Zürich angemeldet haben. Für die Zahnmedizin ist kein Numerus Clausus notwendig, weil die Anfängerkapazität dort ausreicht.

Bei einer Anfängerkapazität der vier Universitäten von insgesamt 768 Studienplätzen (Humanmedizin: 598, Veterinärmedizin: 170) lagen der CRUS im Februar 2003 Anmeldungen von **1448** Bewerberinnen und Bewerbern (Humanmedizin: 1177, Veterinärmedizin 271) für beide Disziplinen vor.

Insgesamt 1143 Personen haben sich bis zur gesetzten Frist im Mai 2003 bei der CRUS zum Test angemeldet, tatsächlich angetreten sind am Testtag 1089 Personen. 54 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen. **1088 Personen (871 Humanmedizin, 217 Veterinärmedizin) haben den EMS vollständig und mit gültigem Ergebnis absolviert.** Eine weitere Person hat die Testbearbeitung vorzeitig abgebrochen.

58 Personen (47 Humanmedizin, 11 Veterinärmedizin) übernehmen ihr Testergebnis aus dem Vorjahr – diese Zahl ist im Vergleich zu den Vorjahren bisher am grössten.

## Zunehmende Nachfrage nach Studienplätzen in der Medizin

Durch die verstärkte Nachfrage konnten in diesem Jahr von allen Bewerberinnen und Bewerbern **nur rund 66%** (gegenüber rund 80% in den Vorjahren) in Humanmedizin **und 75%** (gegenüber rund 90% in den Vorjahren) in Veterinärmedizin einen Studienplatz erhalten.

Eine Erklärung ist das Vorhandensein von „Doppelmaturitäten“ (zwei Abschlussjahrgänge in einem Jahr durch Verkürzung der Maturitätsdauer) in den einzelnen Kantonen, deren Auswirkungen sich offenbar über mehrere nachfolgende Jahre verteilen. Die Bewerberzahlen in einzelnen Kantonen nehmen allerdings auch ohne diese Doppelmaturitäten zu, so dass eine vermehrte Nachfrage nach Hochschulbildung auch für den Bereich der Medizin feststellbar ist. Dissuasionseffekte durch den Test oder die Wirkung der Information über erhöhte Anforderungen im Medizinstudium lassen nach. Indem ein Ärztemangel im Krankenhaussektor (gegenüber einer weiterhin vorhandenen Überversorgung in eigenen Praxen) erwartet wird, wird sich der Arbeitsmarkt weiterhin günstig entwickeln, wenn auch die Aussicht auf eine selbständige Praxiseröffnung abnimmt.

Für die Universitäten hat diese Nachfrage eine positive Wirkung. Bekanntlich konnte ein linearer Zusammenhang zwischen Testwert und Studienerfolg (Dauer, Prüfungsergebnisse) in

mehreren Studien nachgewiesen werden. Wenn der „kritische“ Testwert für die Zulassung 2003 höher liegt, sind die Prognosen für den Studienerfolg der Zugelassenen auch höher.

## Organisation und Ablauf

Der Test fand am 4.7.2003 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. In diesem Jahr gab es einen einzigen vorzeitigen Testabbruch, hingegen keine Verwarnungen oder Testausschlüsse. Dies spricht für eine disziplinierte und motivierte Testteilnahme aller Kandidatinnen und Kandidaten – wie bisher in jedem Jahr seit 1998.

## Testgüte

Ein Vorzug des EMS ist es, dass durch ein aufwendiges Konstruktionsverfahren in jedem Jahr Testversionen mit vergleichbaren Güte-Kennwerten zur Verfügung stehen. Die Übertragbarkeit von Testergebnissen zwischen den Jahren wird dadurch ebenfalls legitimiert. Auch im Jahre 2003 liegen die Zuverlässigkeitswerte der Untertests im Wertebereich der Vorjahre. Die Reliabilität des Testwertes liegt mit 0.92 bzw. 0.81 (Konsistenz) ausreichend hoch. Wiederum tragen alle Untertests gleichermaßen zur Testgüte bei – daher würde das Weglassen eines Untertests die Zuverlässigkeit verringern. Auch die Faktorenstruktur der Vorjahre kann als sehr stabil bestätigt werden.

## Gleichbehandlung der Sprachgruppen

Es zeigen sich die wie seit 1998 in jedem Jahr festgestellten Unterschiede in den mittleren Testwerten zwischen den Sprachgruppen, die aufgrund der unterschiedlichen Zusammensetzung der drei Sprachgruppen und der über die Jahre zu beobachtenden Dynamik auch 2003 zu erwarten waren. Von besonderer Bedeutung bei der Bewertung der Unterschiede bleibt der gesicherte Befund, dass in allen drei Sprachgruppen gleiche Testwerte eine gleiche Studieneignung bedeuten – die mittleren Testwerte für die Sprachgruppen nach dem Kriterium des Bestehens oder Nichtbestehens der ersten Vorprüfung unterscheiden sich nicht (siehe Bericht 7 des ZTD).

Ein Ausgleich des Testwertes soll nur für einzelne Aufgaben erfolgen, die sich gegenüber allen anderen Aufgaben besonders zwischen den Sprachgruppen unterscheiden (DIF-Verfahren) – d.h. ein Ausgleich um den Anteil, den diese spezifische Abweichung ausmacht. Dafür gilt weiter der Beschluss der Schweizerischen Hochschulkonferenz aus dem Jahre 1999: (1) Ausgleich nur bei sprachabhängigen Tests – dies sind fünf von neun Untertests. Der Sprachausgleich beschränkt sich auf die Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ sowie „Diagramme und Tabellen“; (2) den Ausgleich nur vorzunehmen, wenn Mittelwertunterschiede zwischen den Sprachgruppen im entsprechenden Untertest vorhanden sind.

Mittelwertunterschiede ergeben sich 2003 in fünf (französisch) bzw. zwei (italienisch) der möglichen fünf Untertests. In der italienischsprachigen Version zeigten sich nur geringe

Abweichungen weniger Aufgaben vom Trend aller Items (DIF), sodass dort, nach der Rundung, keine Korrekturen notwendig waren. In der französischsprachigen Version waren Korrekturen notwendig – maximal wurden pro Person nur 2 Punkte gutgeschrieben. Wiederum zeigt sich, dass die Unterschiede in den nicht sprachabhängigen Tests deutlich grösser ausfallen als in den sprachabhängigen Tests.

Es kann auch 2003 als gesichert gelten, dass es aufgrund des Ausgleichs keine testbedingten Unterschiede in den Ergebnissen mehr gibt, welche die französisch- oder italienischsprachige Gruppe benachteiligen würden.

### **Differenzierung nach Alter**

In diesem Jahr erreichen die jüngsten Bewerber (geboren 1983-1987) einen mittleren Testwert von 100.8, gefolgt von den ältesten mit einer frühen Maturitätsprüfung (geboren 1959-1979) und Durchschnittswert 99.6. Die Gruppe der 1980-1982 Geborenen ist statistisch davon nicht signifikant zu unterscheiden (Testwert 98.4). Erneut weisen die 1959-1979 Geborenen mit spät erworbener Matur einen geringeren mittleren Testwert auf (96.1) – dieser liegt allerdings 3 Punkte höher als im Vorjahr.

### **Differenzierung nach Geschlecht**

Der Anteil der Bewerberinnen für Human- und Veterinärmedizin bleibt auf dem hohen Niveau des Vorjahres (1998: 55.2% - 2001: 61.5% - 2002: 68.5% - 2003: 67.2%). Die Erwartung identischer mittlerer Testwerte für beide Geschlechter ist schon deshalb gering.

Wiederum sind die Unterschiede beim Testwert nur tendenziell vorhanden und statistisch nicht signifikant – Frauen erreichen nur geringfügig niedrigere Testwerte (die Differenz ist mit rund 1.4 bei Humanmedizin in der deutschen Sprachgruppe mit den Vorjahren vergleichbar). Dass es sich nicht um eine Benachteiligung handelt, belegen die identischen Mittelwerte des Testwertes für die Gruppen nach dem Prüfungsbestehen der ersten Vorprüfung für beide Geschlechter. Eine geschlechtsspezifische Benachteiligung beim EMS kann erwiesenermassen in den Bereich der Legenden verwiesen werden.

### **Differenzierung nach Maturitätsquote und Kanton**

Auch 2003 hat die kantonale Maturitätsquote isoliert betrachtet keinen Einfluss auf den Testwert. Unterschiede zwischen den Kantonen (betrachtet wurden die Kantone mit ausreichender Personenzahl für statistische Analysen) sind auch in diesem Jahr vorhanden, wobei der Kanton Luzern erstmals die schweizweit höchsten und Basel-Stadt seine seit 1998 besten mittleren Testwerte erreicht hat.

### **Schlussfolgerungen**

Ein deutlicher Anstieg der Nachfrage in beiden Disziplinen führte in diesem Jahr dazu, dass die Zulassungsquoten um etwa 15% niedriger als in den Vorjahren lagen. Auch unter solchen Bedingungen macht die Differenzierungsfähigkeit des EMS diesen als Zulassungskriterium geeignet. Die EMS-Session 2003 konnte mit Erfolg abgeschlossen werden und sie setzt die bewährte Testqualität der Vorjahre fort.

## 2 Résumé

Le présent rapport commente les résultats du test d'aptitudes pour les études de médecine en Suisse (AMS), un test nécessaire pour être admis dans certaines facultés de médecine humaine ou vétérinaire au semestre d'hiver 2003/04. Il intègre des données comparatives des tests effectués de 1998 à 2002.

A noter que les résultats de ce rapport ne sont pas suffisamment représentatifs pour permettre de tirer des conclusions générales sur les établissements de formation, les filières d'études ou les régions.

### Statistiques

Appliqué la première en 1998 pour la médecine humaine (en 1999 pour la médecine vétérinaire), un numerus clausus a de nouveau été utilisé en 2003 pour gérer l'admission en médecine aux Universités de Bâle, Berne, Fribourg et Zurich. Cette mesure n'a pas été nécessaire en médecine dentaire puisque la capacité en première année est suffisante.

En effet, alors que les quatre universités ont une capacité totale de 768 places d'études en première année (médecine humaine: 598, médecine vétérinaire: 170), la CRUS enregistrait **1448** pré-inscriptions pour les deux disciplines (1177 en médecine humaine et 271 en médecine vétérinaire) en février 2003.

Au total, 1143 personnes se sont inscrites dans les délais (mai 2003) auprès de la CRUS pour passer le test: elles étaient 1089 le jour J, 54 s'étant retirées ou ne se sont pas présentées. Finalement, **1088 d'entre elles (871 médecine humaine, 217 médecine vétérinaire) ont passé l'AMS en obtenant un résultat valide**, une personne ayant interrompu son test prématurément.

58 personnes (47 médecine humaine, 11 médecine vétérinaire) reprennent leur résultat de l'année dernière, un chiffre en hausse par rapport aux années précédentes.

### Intérêt en hausse pour la médecine

En raison de leur nombre élevé, **seuls 66% environ** des candidats (contre 80% les années précédentes) en médecine humaine et **75%** (contre 90% environ) en médecine vétérinaire ont obtenu une place d'études cette année.

Cette situation s'explique notamment par le fait que certains cantons ont connu une double volée de bacheliers en une seule année (la durée de la maturité ayant été réduite), dont les répercussions se font manifestement sentir sur plusieurs années. A noter toutefois que le nombre de candidats au test augmente aussi dans d'autres cantons, on constate ainsi un intérêt accru pour la médecine au même titre que pour toute formation universitaire: le test ou les exigences élevées en médecine semblent avoir perdu de leur effet dissuasif. Le marché du travail reste attractif, car on s'attend à manquer de médecins dans les hôpitaux, même si les chances de pouvoir ouvrir son propre cabinet diminuent en raison de la saturation de ce secteur.

Cet *intérêt* a un *impact positif* pour les universités. Plusieurs études ont en effet prouvé qu'il existe un lien linéaire entre résultat au test et probabilité de réussir ses études (durée, résultats aux examens). Si le seuil d'admission 2003 est élevé, le taux de réussite des personnes admises le sera aussi.



## Organisation et déroulement

Le test s'est déroulé simultanément dans sept villes et en trois langues le 4 juillet 2003 (en allemand à Bâle, Berne, Coire, St-Gall et Zurich; en français à Fribourg; en italien à Bellinzone). Relevons qu'on peut choisir le lieu du test indépendamment du futur lieu d'études. Aucun incident n'est à signaler: ni avertissement, ni exclusion, et un seul abandon, signe que tous les participants étaient, comme par le passé, disciplinés et motivés.

## Qualité du test

L'avantage de l'AMS, c'est que l'on peut chaque année constituer, grâce à un processus complexe, une version du test de qualité comparable, ce qui justifie la transposition des résultats d'une année sur l'autre. Cette année encore, la fiabilité des sous-tests se situe à un niveau suffisamment élevé et comparable aux années précédentes: 0,92 (stabilité) et 0,81 (consistance). Tous les sous-tests contribuent à la qualité du test, en laissant un seul de côté, diminuerait la fiabilité. La structure factorielle reste elle aussi très stable.

## Egalité de traitement des groupes linguistiques

L'hétérogénéité des trois groupes linguistiques ainsi que la dynamique observée au fil des ans, laissent présager, comme les années précédentes, que les moyennes des résultats au test allaient différer d'un groupe linguistique à l'autre. Mais, comme on le sait désormais de façon certaine, un même résultat au test signifie une même aptitude à étudier la médecine quelque soit le groupe linguistique, car on ne constate aucune différence entre les moyennes des groupes linguistiques répartis en fonction de la réussite au premier propédeutique (cf. Rapport 7 du CTD).

Il n'est possible d'apporter un correctif aux résultats que pour les exercices, où il existe des différences particulières entre groupes linguistiques (DIF: "differential item functioning"), ce correctif équivaut à la différence spécifique. Conformément à la décision que la Conférence universitaire suisse (CUS) a prise en 1999, on utilise une telle procédure (1) uniquement pour les tests à forte composante linguistique, soit cinq sous-tests sur neuf: "Compréhension de questions fondamentales de la médecine et des sciences naturelles", "Problèmes quantitatifs et formels", "Compréhension de textes", "Mémorisation de faits", "Diagrammes et tableaux"; et (2) que lorsqu'on constate, dans un sous-test, des moyennes différentes entre groupes linguistiques.

En 2003, des différences entre les moyennes sont apparues dans cinq sous-tests en français et deux sous-tests en italien. Dans la version italienne toutefois, aucun exercice n'a présenté un écart particulier par rapport aux autres (DIF), si bien qu'il n'a pas été nécessaire d'apporter de correctif. En revanche, il a fallu apporter un correctif à la version française: on a ainsi accordé 2 points au maximum par personne. Il en résulte une fois de plus que les différences sont nettement plus marquées dans les tests à faible composante linguistique que dans les tests à forte composante linguistique.

On peut donc affirmer avec certitude que, grâce au correctif, il n'existe aucune différence au niveau des résultats, due au test lui-même, qui désavantagerait le groupe francophone ou italophone.

## Différences entre classes d'âge

Cette année, la moyenne des plus jeunes candidats (nés entre 1983 et 1987) est de 100,8. Elle n'est guère différente de celle des candidats plus âgés (nés entre 1959 et 1979) ayant passé tôt leur maturité, qui est de 99,6. Le groupe des candidats nés entre 1980 et 1982 ne s'en distingue pas non plus de manière significative d'un point de vue statistique (moyenne: 98,4). Mais le groupe des plus âgés (nés entre 1959 et 1979) ayant passé tardivement leur maturité obtient une fois encore une moyenne inférieure (96,1), même si elle est de 3 points supérieure à l'année passée.

## Différences entre femmes et hommes

La proportion des candidates à la médecine humaine ou vétérinaire reste élevée cette année encore (1998: 55,2% - 2001: 61,5% - 2002: 68,5% - 2003: 67,2%). On ne peut dès lors guère s'attendre à ce que les deux sexes aient des moyennes identiques.

On observe certes que les femmes ont tendance à obtenir des résultats au test sensiblement inférieurs, mais cette tendance n'est statistiquement pas significative (1,4 en médecine humaine dans le groupe germanophone, ce qui est comparable à 2002): le fait que les moyennes au test pour les groupes répartis en fonction de la réussite au 1<sup>er</sup> propédeutique sont les mêmes pour les deux sexes prouvent que les femmes ne sont pas désavantagées. On peut donc écarter à juste titre toute affirmation contraire.

## Différences entre cantons et taux de maturité

Le taux cantonal de maturité, à lui seul, n'a aucune influence sur les résultats. On constate cette année également des écarts entre les cantons (ayant un nombre suffisant de personnes pour une analyse statistique): le canton de Lucerne a obtenu pour la première fois la meilleure moyenne au test du pays, et Bâle-Ville sa meilleure moyenne depuis 1998.

## Conclusions

Le taux d'admission a baissé de 15% cette année par rapport aux années précédentes en raison d'un intérêt nettement accru pour les deux disciplines. Dans ces conditions, l'AMS, grâce à sa capacité de différenciation, reste un critère d'admission adéquate. Le succès de la session 2003 confirme la qualité déjà éprouvée de ce test.

### 3 Aktuelle Diskussionen zum EMS

Die Nachfrage nach Studienplätzen in der Medizin hat in diesem Jahr einen neuen Höchststand erreicht. Dies kann nicht mehr allein durch „Doppelmaturitäten“ erklärt werden (zwei Jahrgänge legen die Maturität wegen der Verkürzung der Maturitätsdauer im gleichen Jahr ab), wenn man die kantonalen Zuwächse der Bewerbungen analysiert (vgl. Tabelle 5).

Es legen mehr Personen eine Wartezeit nach der Maturität ein, um entweder den stark nachgefragten Jahrgängen an der Universität zu entgehen oder sich mit den laufenden Curricula-Reformen im Medizinstudium abzustimmen und dann unter den neuen Studienbedingungen zu beginnen. Ein Zeichen dafür ist, dass immer mehr Personen entweder ihr Testergebnis ins Folgejahr übernehmen oder erneut zum EMS antreten. Mit insgesamt 90 Personen für beide Gruppen im Jahr 2003 werden die „Rückzüge“ von Personen zwischen Eignungstest und Studienaufnahme immer mehr zu „Aufschüben“. Dennoch betrifft die verstärkte Nachfrage vor allem Personen, die sofort nach der Mittelschule ein Medizinstudium aufnehmen möchten (vgl. Abbildung 12) – bei den Älteren bleibt die Nachfrage bisher konstant.

Der Medizinberuf scheint wieder an Attraktivität zu gewinnen – die Mitte der 90er Jahre versuchte „Dissuasion“ durch Bekanntmachen der real hohen Belastungen in Studium und Beruf wirkt offenbar nicht mehr. In der öffentlichen Diskussion wird gegenwärtig ein hoher Bedarf an Krankenhausmediziner für die Zukunft postuliert. Dieser soll u.a. entstehen, weil der überdurchschnittliche Arbeitsaufwand in diesem Beruf für die einzelnen Personen abnimmt – vor allem aus arbeitsrechtlichen Gründen werden die Anforderungen reduziert. Dadurch müssen mehr Personen für die Bewältigung der gleichen Anforderungen eingestellt werden. Der Beruf des Krankenhausarztes mit gut bewältigbaren Anforderungen ist relativ zukunftssicher, wogegen die Berufsaussichten in anderen Berufen und Branchen durch die lang anhaltende Rezession relativ an Attraktivität einbüßen.

Die Kostenexplosion im Gesundheitswesen verbunden mit notwendigen einschränkenden Massnahmen, bleibt ein Dissuasionsfaktor – scheint aber mehr als Beschränkung für niedergelassene Mediziner wahrgenommen zu werden. Für die möglicherweise zunehmende Attraktivität der Krankenhausmediziner könnten in der heutigen „Mediengesellschaft“ nicht zuletzt auch die zahlreichen Filme und Serien auf fast allen Fernsehstationen beitragen, welche die Arbeit im Krankenhaus zumindest keinesfalls dissuasiv darstellen.

Das strukturelle Problem im Gesundheitswesen, die Kostensteigerung durch mehr Behandlungsangebote, die sich eine eigene Nachfrage schaffen, bleibt allerdings ungelöst. Eine vermehrte Mediziner Ausbildung würde diesen Kreislauf vermutlich eher anheizen. Die Schweiz folgt hier auch einem internationalen Trend. Nachdem beispielsweise in Deutschland das Verhältnis Bewerbungen zu Studienplätzen auf 3:1 abgesunken war, bewegt es sich wieder auf ein Verhältnis von 5:1 zu. Ein interessanter Fakt ist hier, dass für eine verbesserte Selbstauswahl der Universitäten, über die so genannte Länderquote, wieder Wünsche nach der Anwendung von Zulassungstests entstanden sind.

Eine erwünschte Folge der vermehrten Nachfrage und des sich damit verschärfenden Numerus Clausus ist, dass die Zugelassenen bessere Testwerte erreicht haben als in früheren Jahren. Dies dürfte sich an den Universitäten darin zeigen, dass weniger Studienverlängerungen und Prüfungswiederholungen auftreten.

Der Eignungstest beruht weiter auf dem bewährten Konzept, welches in den bisherigen Prognosestudien eine sehr gute Vorhersagefähigkeit für Studiendauer und Studienerfolg nachgewiesen hat. Die Analysen basierten auf den Prüfungsergebnissen der EMS-Teilnehmer von 1998 und 1999 (vgl. Bericht 7 des ZTD). Zum einen werden gegenwärtig Analysen für die Kohorten 1998 bis 2001 vorgenommen, um nun auch die langfristige Prognosefähigkeit genauer zu analysieren. Zum anderen hat ein Nachdenken darüber begonnen, ob eine Überarbeitung des EMS aufgrund der erwarteten neuen Studienanforderungen durch die Reformanstrengungen und der neuen Gesetzgebung notwendig ist.

Anlässlich eines Workshops am 9.4.2003 gemeinsam mit Vertretern der medizinischen und veterinärmedizinischen Fakultäten wurde daher eine Standortbestimmung des EMS vorgenommen. Es ging darum, neue Anforderungen abzuleiten, welche in Zukunft die notwendige Prognosekraft des Tests für Studieneignung erhalten.

Das bisherige Konzept des EMS wurde dabei für weiterhin wichtig erachtet. Die erfassten Anforderungen bleiben auch für neue Studienbedingungen relevant. Dennoch ist es sinnvoll, weitere Bereiche stärker zu gewichten, die unter den neuen Studienbedingungen ebenfalls mehr an Gewicht gewinnen. Als mögliche zusätzliche Indikatoren sollten in der Weiterentwicklung des EMS besonders berücksichtigt werden:

- Selbstorganisation und Planung
- Beobachten (Unterschiede, bewegte Situationen, Details)
- Problemlösen (komplex)
- Wesentliches erkennen
- Wiedererkennen (visuell, akustisch, Bewegung)

Zum bereits früher mehrfach diskutierten Bereich der sozialen Kompetenz wurde auch über Erfahrungen aus Belgien referiert, in welchen, im Rahmen des dort entwickelten Eignungstests, filmische Sequenzen nach bestimmten kommunikativen Merkmalen einzuschätzen waren. Diese Indikatoren wiesen allerdings keine hinreichende Prognosegüte für Studienerfolg auf. Der Aufwand bei der Erhebung (Videodarbietung, gleiche Wahrnehmungsbedingungen für alle Teilnehmer in einem Testlokal) steht daher in keinem Verhältnis zum Nutzen.

Es wurde festgestellt, dass eine Revision des EMS nur schrittweise erfolgen soll – die Aussagekraft des Tests muss jedes Jahr gleich und wegen der hohen Zahl der Übernahmen von Testergebnissen auch zwischen den Jahren vergleichbar sein.

Es wurde eine Vereinbarung zwischen der ITB Consulting GmbH und der CRUS bzw. dem ZTD getroffen, um in den Bereichen „Selbstorganisation und Planung“ sowie „Gedächtnis“ neue Testkonzepte zu entwickeln und vorzuerproben, die bei Bewährung in den kommenden Jahren in den EMS eingefügt werden sollen.

## 4 Numerus Clausus (NC) und Medizinstudium

Die Abbildung 1 verdeutlicht die Anmeldeverläufe 1994 bis 1997 (ohne NC) sowie 1998 bis 2003 (mit NC).

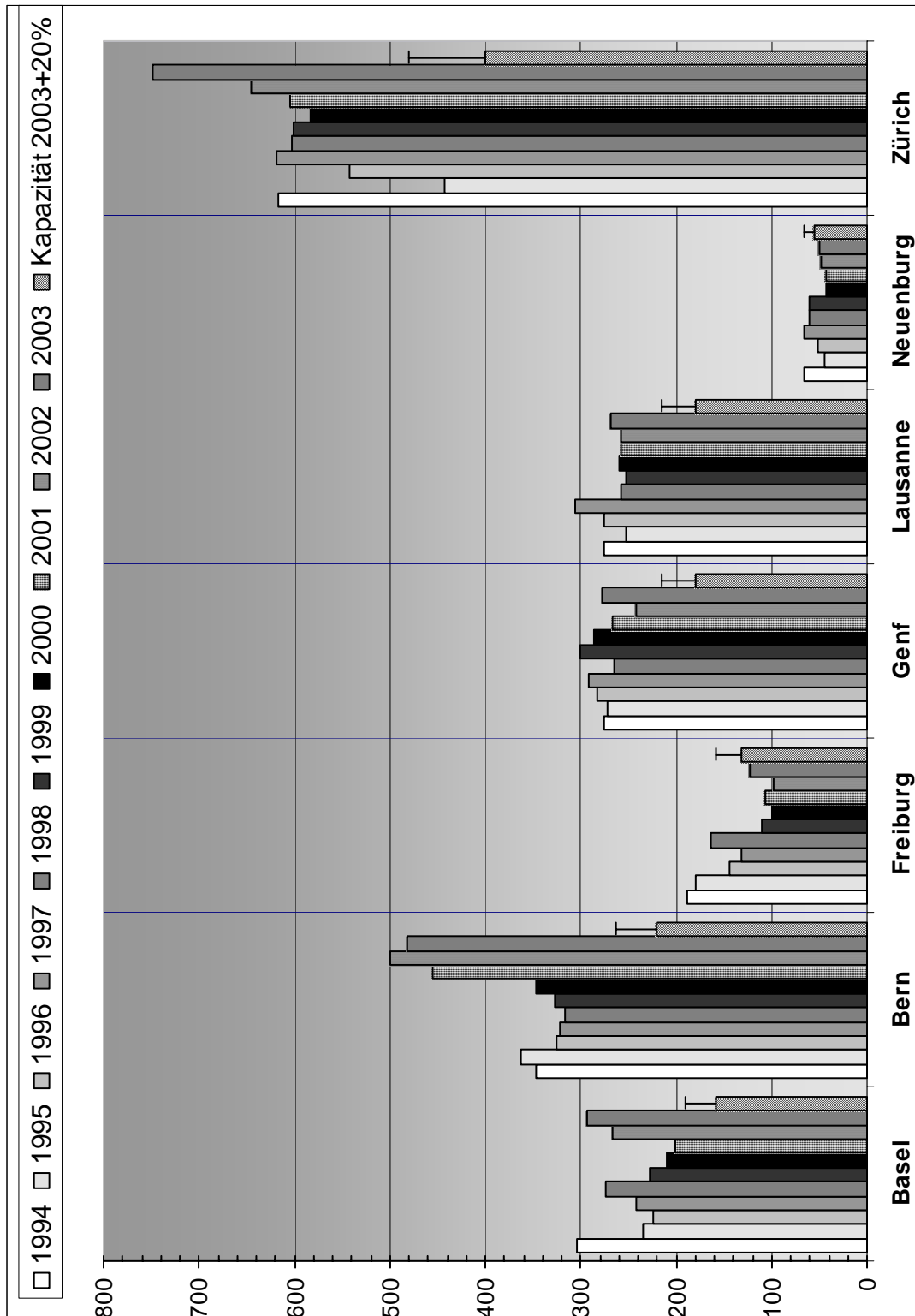


Abbildung 1: Anmeldungen zum Medizinstudium 1994 bis 2003, Anfängerprognose und Kapazitäten 2003 (letzte Säule – mit Markierung des 20%-Überschreitungskriteriums für 2003 als Grenze für die Notwendigkeit einer Zulassungsbeschränkung)

Die letzte Säule zeigt die Kapazitäten und das 20%-Überschreitungskriterium, ab welchem ein NC zur Anwendung kommt. Zürich zeigt in diesem Jahr wie erwartet eine deutliche Zunahme (Jahr nach der Doppelmaturität). Die Bewerberzahl in Bern bleibt auch 2 Jahre nach der Doppelmaturität noch hoch – offensichtlich verteilt sich die zusätzliche Nachfrage auf mehrere Folgejahre. In Basel ist die Nachfrage seit 1994 am zweithöchsten und auch in Freiburg nehmen die Anmeldezahlen nun wieder zu.

Wir nehmen an, dass neben den Doppelmaturitäten auch die allgemeine Nachfrage nach Medizinstudienplätzen wieder zunimmt. Dies wird später auch anhand der Aufschlüsselung der Bewerbungen nach Kantonen deutlich.

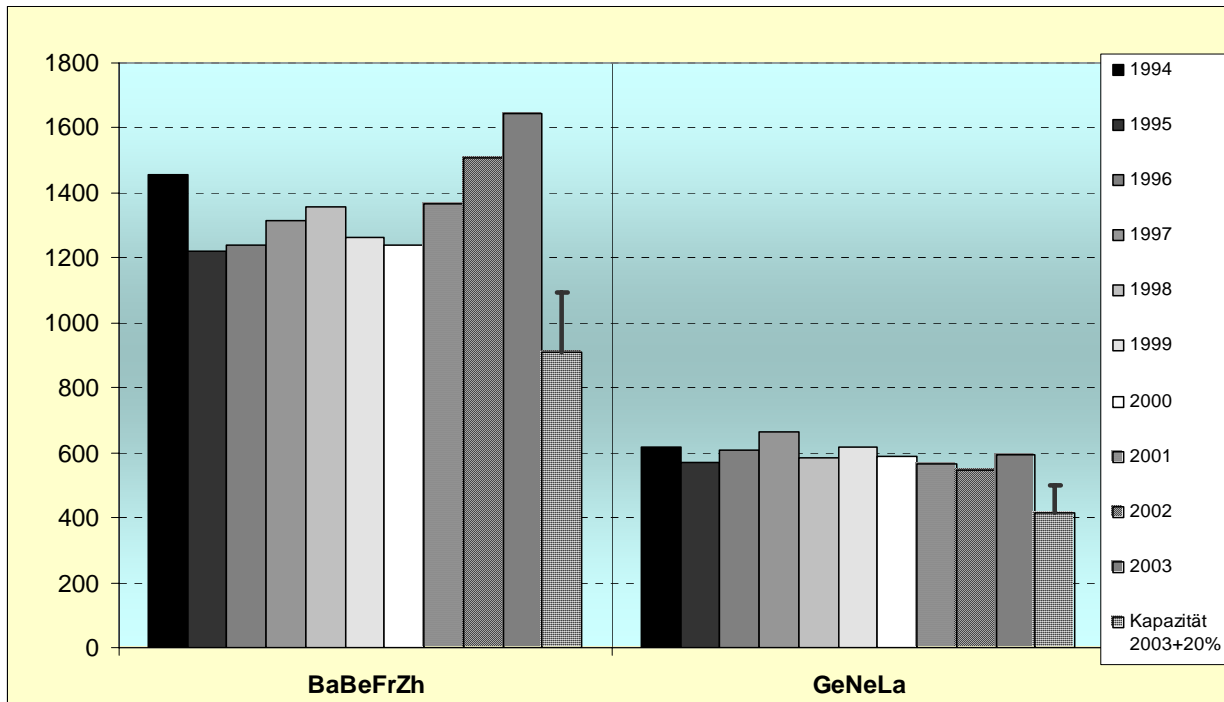


Abbildung 2: Anmeldungen zum Medizinstudium 1994 bis 2003 nach Gruppen (**B**asel, **B**ern, **F**reiburg, **Z**ürich vs. **G**enf, **N**euenburg, **L**ausanne)

Seit 3 Jahren steigt die Zahl der Anmeldungen zum Medizinstudium an – im Jahr 2003 wird für die am Numerus Clausus beteiligten Universitäten ein Maximum erreicht. Auch in der Westschweiz steigt die Zahl der Bewerbungen erstmals seit 3 Jahren wieder an.

Es ist wie bereits beschrieben davon auszugehen, dass die Attraktivität des Medizinstudiums wieder zunimmt, was besonders in der Deutschschweiz zu vermehrten Studienbewerbungen beiträgt.

## 5 Anmeldung zu Medizinstudium und Test 2003

### 5.1 Anmeldungen, Testantritte und Zulassungen

	Humanmedizin						Veterinärmedizin				
	1998	1999	2000	2001	2002	2003	1999	2000	2001	2002	2003
<b>Kapazitäten (am NC beteiligt)</b>	<b>620</b>	<b>546</b>	<b>546</b>	<b>583</b>	<b>648</b>	<b>598</b>	<b>150</b>	<b>150</b>	<b>165</b>	<b>185</b>	<b>170</b>
Anmeldungen bei CRUS Februar d.J.	<b>1020</b>	<b>845</b>	<b>844</b>	<b>974</b>	<b>1080</b>	<b>1177</b>	<b>231</b>	<b>213</b>	<b>217</b>	<b>254</b>	<b>271</b>
... in % zu Kapazität	186	155	154	167	185	<b>197</b>	153	142	132	137	<b>159</b>
<b>Anmeldungen zum EMS im Mai d.J.</b>	<b>797</b>	<b>717</b>	<b>678</b>	<b>752</b>	<b>827</b>	<b>917</b>	<b>194</b>	<b>176</b>	<b>179</b>	<b>206</b>	<b>225</b>
<b>+ Übernahme Test- ergebnis aus Vorjahr</b>	0	5	33	27	40	<b>47</b>	0	5	5	3	<b>11</b>
<b>= Total der Bewerbungen im Mai</b>	<b>797</b>	<b>722</b>	<b>711</b>	<b>779</b>	<b>867</b>	<b>964</b>	<b>194</b>	<b>181</b>	<b>184</b>	<b>209</b>	<b>236</b>
... in % der Kapazität	129	132	130	134	134	<b>161</b>	129	121	112	113	<b>139</b>
<b>Absolvierung EMS</b>	<b>750</b>	<b>673</b>	<b>637</b>	<b>709</b>	<b>770</b>	<b>871</b>	<b>182</b>	<b>164</b>	<b>171</b>	<b>203</b>	<b>217</b>
<b>Gesamtbewerbungen EMS absolviert oder Ergebnis- übertragung</b>	<b>750</b>	<b>678</b>	<b>670</b>	<b>736</b>	<b>805</b>	<b>907*</b>	<b>182</b>	<b>169</b>	<b>176</b>	<b>205</b>	<b>228</b>
Rückzugsquote zwischen Februar und Gesamtbewerbungen	26,5	19,8	20,6	24,4	25,5	<b>22,9</b>	21,2	20,7	18,9	19,3	<b>15,9</b>
Personenanteil mit gültigem Testwert, die einen Platz erhalten	82,7	80,5	81,4	79,2	80,5	<b>65,9</b>	82,4	88,7	93,7	90,2	<b>74,5</b>
<b>Bewerbungs- index</b> Bewerber pro Studienplatz	1,21	1,24	1,23	1,25	1,24	<b>1,52</b>	1,21	1,13	1,06	1,11	<b>1,34</b>

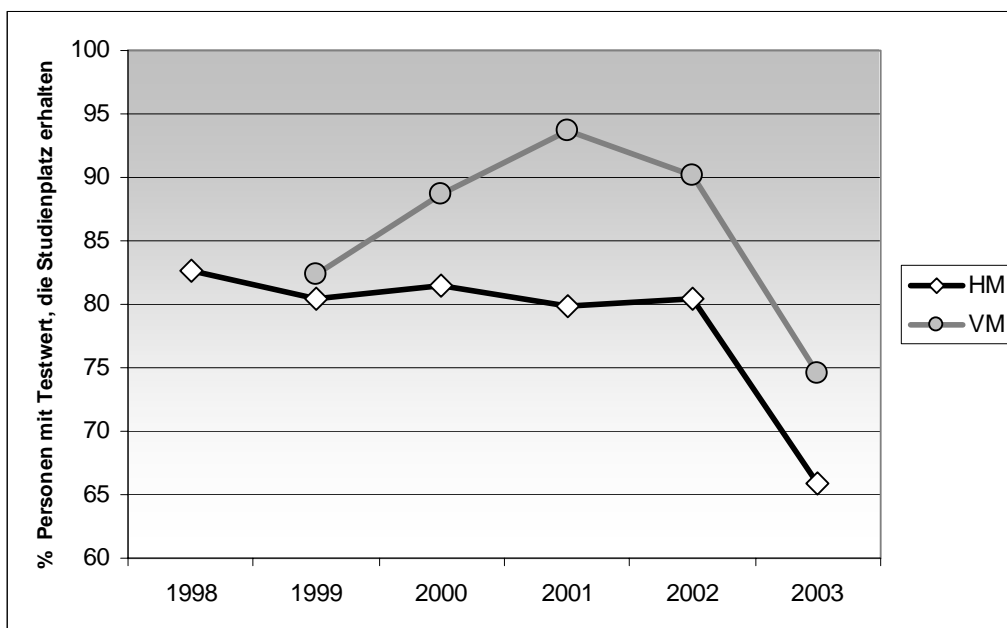
**Tabelle 1:** Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; NC 1998 nur für Humanmedizin – 1999 bis 2003 für Human- und Veterinärmedizin \*11 Personen nahmen am Test teil, um später an eine Universität bzw. Disziplin zu wechseln, für die der Numerus Clausus gilt (zählen hier nicht als Bewerbungen; 9 x von Zahn- in Humanmedizin, 2 x Studienbeginn in Neuenburg).

Tabelle 1 fasst den jährlichen Anmeldeverlauf bis zur Testabsolvierung zusammen. In diesem Jahr kommen in Humanmedizin 1,52 Bewerber auf einen Studienplatz, in Veterinärmedizin sind es 1,34. Dadurch ist ein höherer Testwert für die Zulassung

notwendig als in den Vorjahren und es tritt auch die Situation auf, dass Personen, die Testergebnisse aus dem Vorjahr übertragen haben (und im Vorjahr zugelassen worden wären) in diesem Jahr nicht zugelassen werden können. Bei der Übertragung der Testergebnisse sollte damit gerechnet werden, dass sich die Nachfrage weiter verstärken kann und der Testwert ausreichend hoch sein sollte. In Tabelle 2 und Abbildung 7 (Quelle: CRUS) werden Rückzüge zwischen Anmeldung und Testabsolvierung differenziert.

	Studienbeginn 2000			Studienbeginn 2001			Studienbeginn 2002			Studienbeginn 2003		
	Anmeldungen am 15.2.	Pers. mit gültigem Testwert	Rückzüge in %	Anmeldungen am 15.2.	Pers. mit gültigem Testwert	Rückzüge in %	Anmeldungen am 15.2.	Pers. mit gültigem Testwert	Rückzüge in %	Anmeldungen am 15.2.	Pers. mit gültigem Testwert	Rückzüge in %
Basel HM	174	133 (9)	23.6	175	136 (7)	22.3	239	181 (6)	24.3	256	155 (13)	23.8
Bern HM	177	144 (5)	18.6	282	232 (5)	17.7	305	245 (16)	19.7	294	234 (13)	20.4
Freiburg HM	88	66 (2)	25.0	97	69 (1)	28.9	87	66 (5)	24.1	113	84 (4)	25.7
Zürich HM	405	327 (17)	19.3	420	299 (14)	28.8	449	313 (13)	30.3	518	394 (17)	23.9
<b>Total</b>	<b>844</b>	<b>670</b>	<b>20.6</b>	<b>974</b>	<b>736</b>	<b>24.4</b>	<b>1080</b>	<b>805</b>	<b>25.5</b>	<b>1181</b>	<b>907</b>	<b>23.2</b>
Bern VM	117	99 (0)	15.4	118	104 (4)	11.9	132	106 (2)	19.7	133	108 (4)	18.8
Zürich VM	96	70 (0)	21.7	99	72 (1)	27.3	122	99 (1)	18.9	141	120 (7)	14.9
<b>Total</b>	<b>213</b>	<b>169</b>	<b>20.7</b>	<b>217</b>	<b>176</b>	<b>18.9</b>	<b>254</b>	<b>205</b>	<b>19.3</b>	<b>274</b>	<b>228</b>	<b>16.8</b>

**Tabelle 2:** Anmeldeverlauf für Universitäten, Humanmedizin (HM) und Veterinärmedizin (VM). Personen mit gültigem Testwert – in Klammern: davon Übernahme aus dem Vorjahr



**Abbildung 3:** Prozent der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann – HM: Humanmedizin; VM: Veterinärmedizin



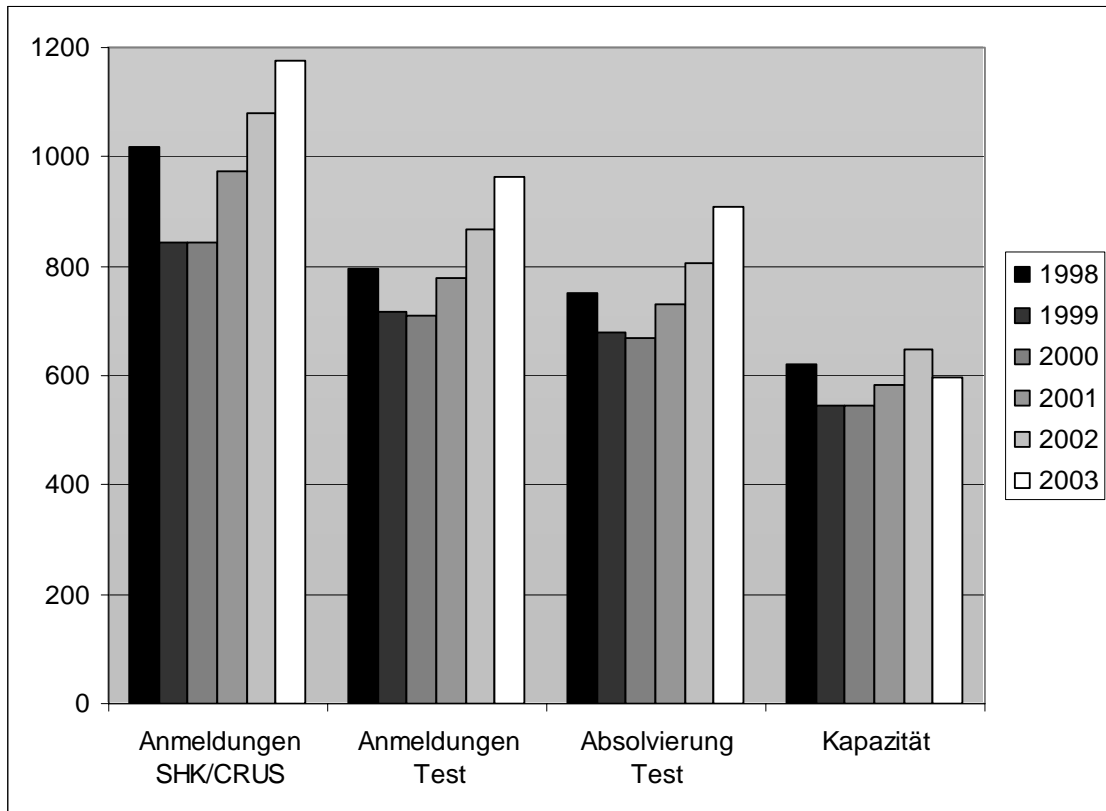


Abbildung 4: Anmeldeverlauf für Humanmedizin 1998 bis 2003 im Vergleich (mit NC)

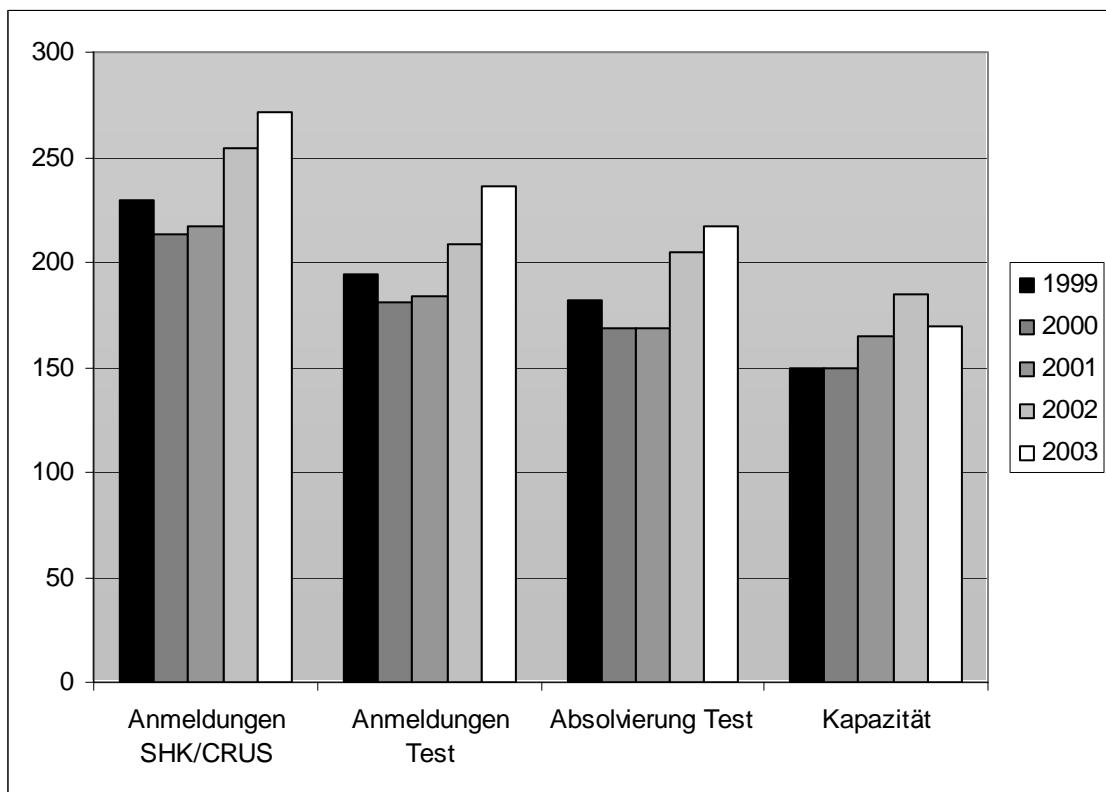


Abbildung 5: Anmeldeverlauf für Veterinärmedizin 1999 bis 2003 im Vergleich (mit NC)

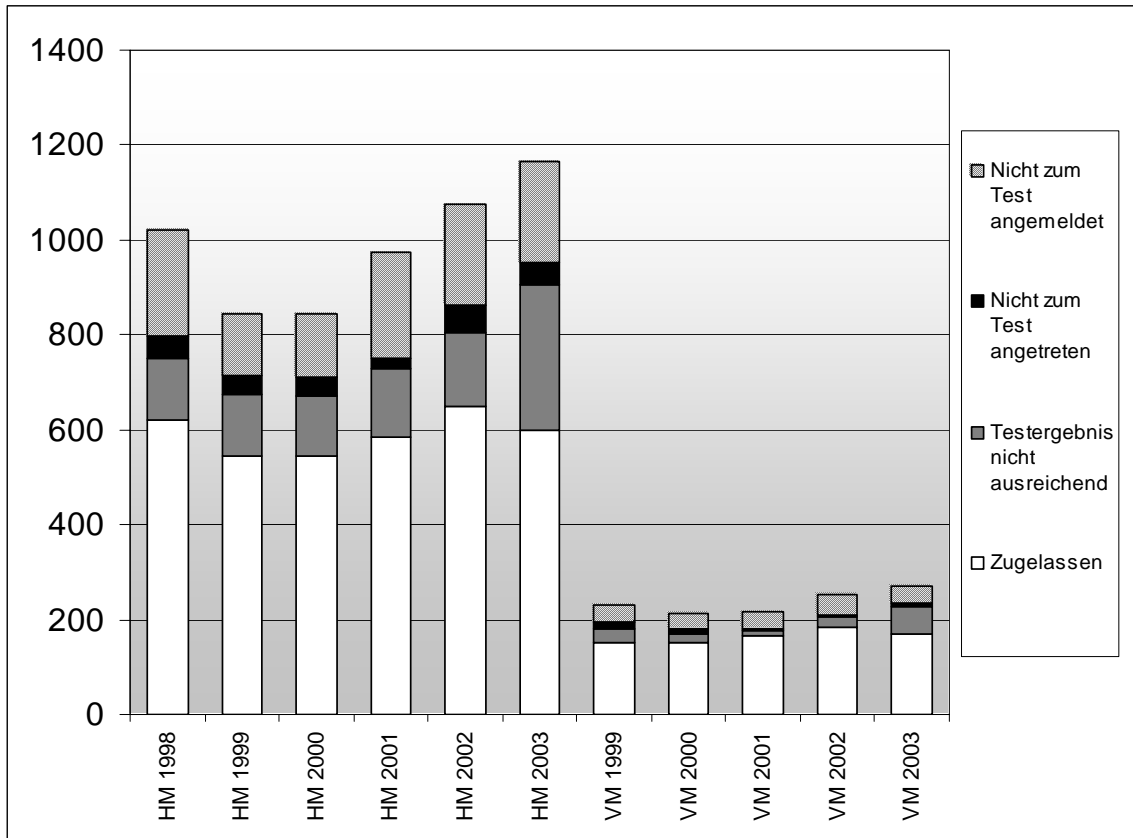


Abbildung 6: Kapazität, Anmeldungen und Testteilnahme für Human- (HM) und Veterinärmedizin (VM). Aufteilung der Voranmeldungen (Februar des jeweiligen Jahres)

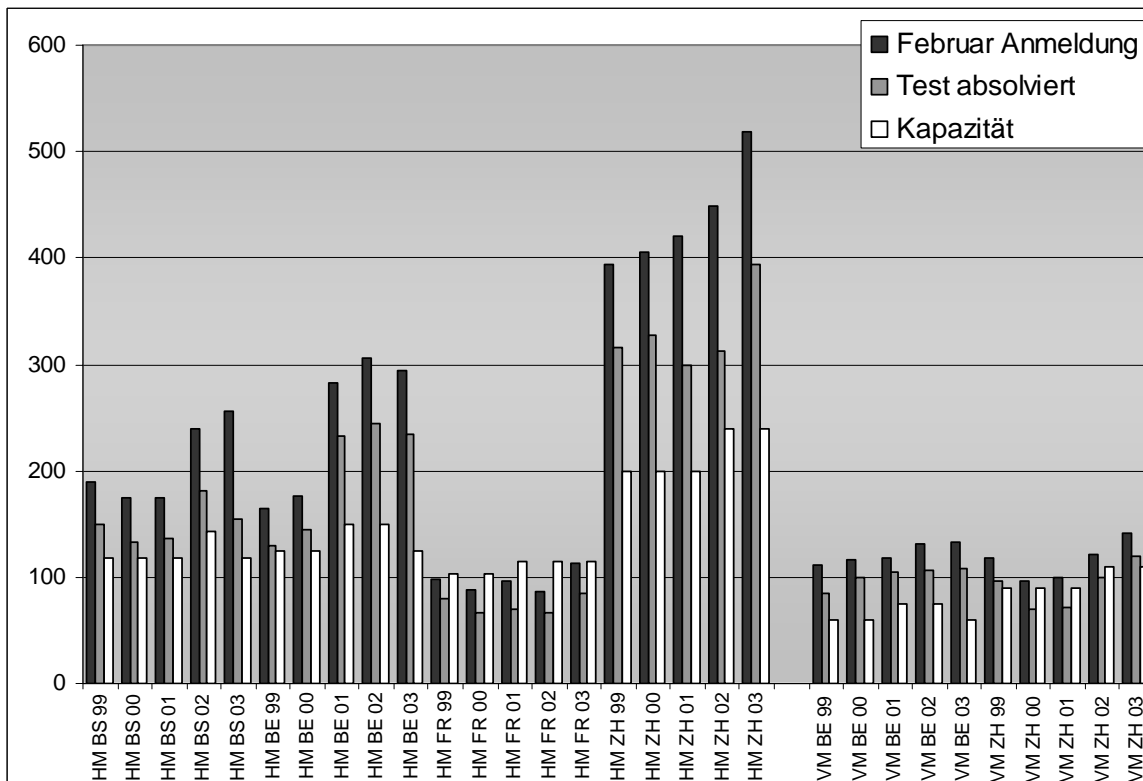


Abbildung 7: Kapazität, Anmeldungen und Testteilnahme für Fächer und Universitäten

## 5.2 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme in 15 Testlokalen in 7 Testorten. Die Grössen variierten im 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).

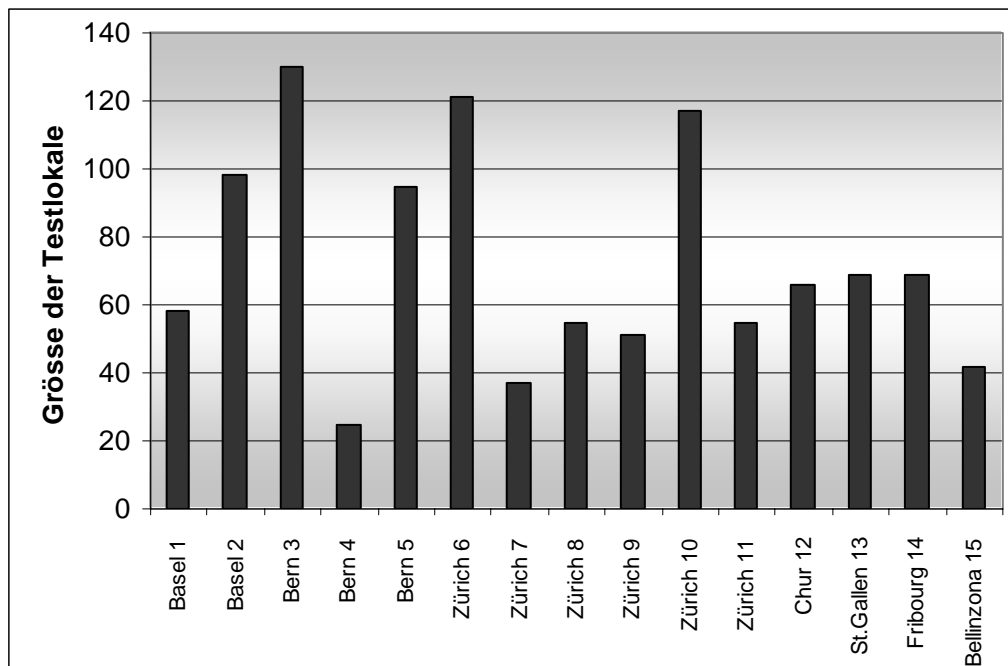


Abbildung 8: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2003

## 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 beiden 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	Vor Anmeldung Universität				Total
	Bern	Basel	Freiburg	Zürich	
Basel	18 11.5%	129 82.7%	4 2.6%	5 3.2%	156 100.0%
Bellinzona	5 11.9%	9 21.4%	6 14.3%	22 52.4%	42 100.0%
Bern	215 86.0%	5 2.0%	18 7.2%	12 4.8%	250 100.0%
Chur	14 21.2%	18 27.3%	3 4.5%	31 47.0%	66 100.0%
Freiburg	35 50.7%	0	32 46.4%	2 2.9%	69 100.0%
St. Gallen	13 18.8%	7 10.1%	5 7.2%	44 63.8%	69 100.0%
Zürich	27 6.2%	22 5.0%	12 2.8%	375 86.0%	436 100.0%
<b>Total</b>	<b>327</b> 30.1%	<b>190</b> 17.5%	<b>80</b> 7.4%	<b>491</b> 45.1%	<b>1088</b> 100.0%

Tabelle 3: Testorte und Wunschuniversität (erste Wahl) 2003

Abbildung 9 macht durch die „Streckung“ aller Universitäten auf 100% das Verhältnis der „lokalen“ Bewerber zu den übrigen vergleichbar.

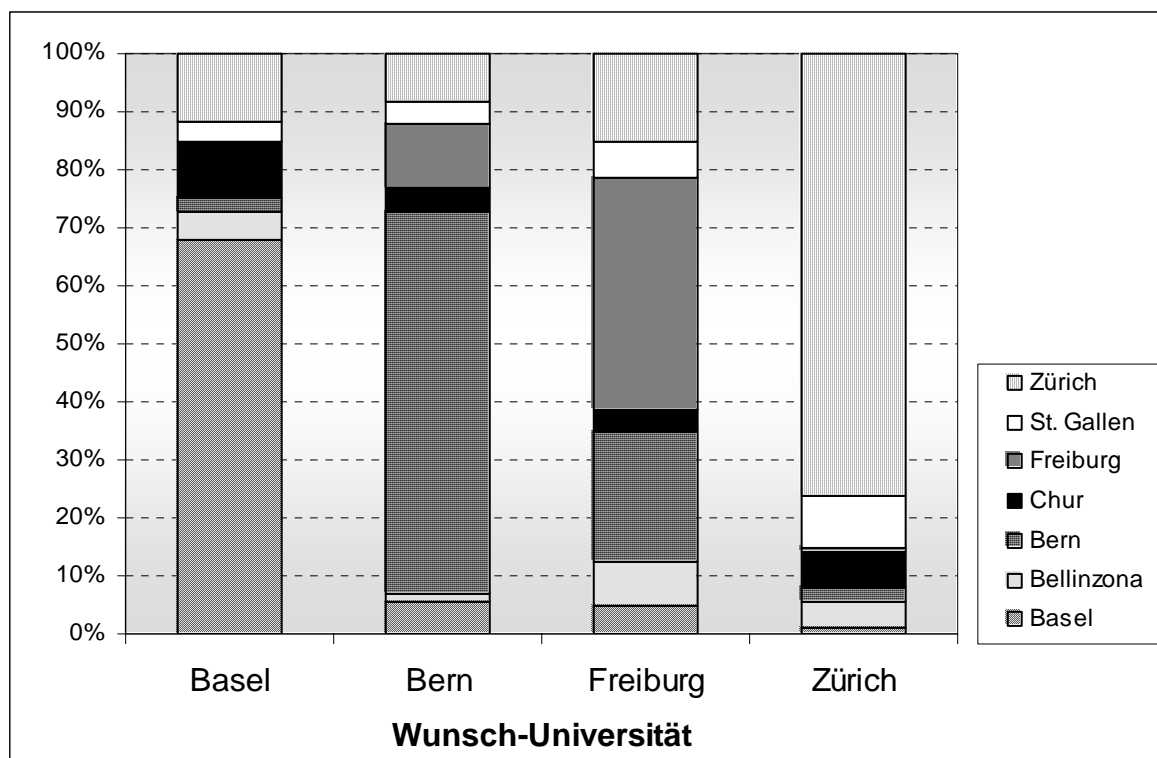


Abbildung 9: Aufteilung der Kandidaten nach Wunschuniversität auf die Testorte 2003

## 5.4 Testorte und Wohnkantone

Nachfolgend sind die Testorte im Verhältnis zum Wohnkanton tabelliert. Angegeben sind auch die Vergleichswerte seit 1998. Doppeljahrgänge der Maturitäten sind in den einzelnen Kantonen und Jahren unterstrichen. In diesem Jahr nimmt auch in mehreren Kantonen ohne Doppelmaturität die Zahl der Bewerbungen zu, was für die generelle Zunahme des Studieninteresses spricht.

Wohnkanton/ ort	Testort 2003							Total					
	BA-SEL	BEL-LIN-ZONA	BERN	CHUR	FREI-BURG	ST. GAL-LEN	ZÜ-RICH	2003	2002	2001	2000	1999	1998 (HM)
AG	23		6				38	<b>67</b>	60	49	71	84	66
AI								<b>0</b>	5	1	<u>4</u>	0	2
AR						7	1	<b>8</b>	6	10	4	6	4
BE	1	1	162		1		2	<b>167</b>	169	<u>186</u>	114	113	100
BL	52							<b>52</b>	44	45	40	59	50
BS	42						1	<b>43</b>	39	39	28	39	40
FR			18		33		1	<b>52</b>	36	27	44	47	34
GE					4		1	<b>5</b>	7	10	11	13	1
GL				1			6	<b>7</b>	4	3	5	4	4
GR	1	1		50			7	<b>59</b>	22	36	19	29	19
JU					5			<b>5</b>	2	1	1	2	0
LU	7		20	1		1	30	<b>59</b>	<u>75</u>	47	41	44	43
NE			2		6			<b>8</b>	<u>7</u>	3	6	5	2
NW	4		5				2	<b>11</b>	<u>6</u>	3	2	3	6
OW	1						1	<b>2</b>	7	3	<u>9</u>	2	3
SG				8		43	10	<b>61</b>	59	48	55	61	<u>64</u>
SH				1			11	<b>12</b>	10	<u>18</u>	8	4	7
SO	19		24			1	1	<b>45</b>	<u>43</u>	25	20	28	26
SZ			1			1	11	<b>13</b>	15	15	18	18	<u>21</u>
TG						14	19	<b>33</b>	34	38	23	20	22
TI		38		1				<b>39</b>	27	38	27	29	30
UR	1						3	<b>4</b>	<u>5</u>	5	2	1	3
VD			1		15		1	<b>17</b>	14	23	17	11	3
VS			6		3			<b>9</b>	20	19	18	15	22
ZG							18	<b>18</b>	19	<u>23</u>	9	10	6
ZH		2	3	2	2		267	<b>276</b>	<u>224</u>	148	194	191	164
FL				2		1		<b>3</b>	4	7	4	2	3
Übriges Aus- land	5		2			1	5	<b>13</b>	10	10	11	17	8
	156	42	250	66	69	69	436	<b>1088</b>	973	880	801	855	750

**Tabelle 4:** Wohnkanton und Testort, Vergleich zu 1998 bis 2003 (1998 nur Humanmedizin). Unterstrichen: Doppeljahrgänge in den Kantonen. Das Fürstentum Liechtenstein (FL) wird als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die Präferenz des gewünschten Studienortes in Abhängigkeit vom Wohnort wird in der Tabelle 5 dargestellt. Diese Präferenzen scheinen vor allem geografisch motiviert zu sein. Zu beachten ist, dass Veterinärmedizin für die gesamte Schweiz seit 1999 nur in Bern und Zürich angeboten wird.

Wohnkanton Wohnort	Vor Anmeldung Universität 2003				Total				
	Bern	Basel	Freiburg	Zürich	2003	2002	2001	2000	1999
AG	9	24		34	<b>67</b>	60	49	71	84
AI					<b>0</b>	5	1	4	0
AR	2	3		3	<b>8</b>	6	10	4	6
BE	156	3	3	5	<b>167</b>	169	186	114	113
BL	6	40	3	3	<b>52</b>	44	45	40	59
BS	5	36		2	<b>43</b>	39	39	28	39
FR	13		38	1	<b>52</b>	36	27	44	47
GE	3			2	<b>5</b>	7	10	11	13
GL			1	6	<b>7</b>	4	3	5	4
GR	12	16		31	<b>59</b>	22	36	19	29
JU	4		1		<b>5</b>	2	1	1	2
LU	27	12	6	14	<b>59</b>	75	47	41	44
NE	8				<b>8</b>	7	3	6	5
NW	6	2	1	2	<b>11</b>	6	3	2	3
OW		2			<b>2</b>	7	3	9	2
SG	10	4	4	43	<b>61</b>	59	48	55	61
SH	2	3	1	6	<b>12</b>	10	18	8	4
SO	19	18	1	7	<b>45</b>	43	25	20	28
SZ	4	1	1	7	<b>13</b>	15	15	18	18
TG	5	1	3	24	<b>33</b>	34	38	23	20
TI	4	10	6	19	<b>39</b>	27	38	27	29
UR	3	1			<b>4</b>	5	5	2	1
VD	15		1	1	<b>17</b>	14	23	17	11
VS	6	1	1	1	<b>9</b>	20	19	18	15
ZG		2	1	15	<b>18</b>	19	23	9	10
ZH	4	4	8	260	<b>276</b>	224	148	194	191
FL	1	2			<b>3</b>	4	7	4	2
Übriges Ausland	3	5		5	<b>13</b>	10	10	11	17
<b>Total</b>	<b>327</b>	<b>190</b>	<b>80</b>	<b>491</b>	<b>1088</b>	<b>973</b>	<b>880</b>	<b>801</b>	<b>855</b>

Tabelle 5: Gewünschte Studienorte nach Wohnkantonen, Vergleich zu 1999 bis 2003. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wohnkanton/ Wohnort	Studienort		Total		
	Bern	Zürich	2003	2002	2001
AG	4	9	13	8	7
AI			0	3	0
AR	1	1	2	0	2
BE	30	2	32	30	39
BL	5	3	8	7	4
BS	1	1	2	4	7
FR	8		8	9	3
GE	3	1	4	6	9
GL		1	1	1	0
GR	4	9	13	4	6
JU	4		4	2	1
LU	2	5	7	13	5
NE	7		7	6	3
NW	2		2	2	0
SG	2	11	13	9	7
SH	1	1	2	0	3
SO	6	1	7	8	4
SZ	1	1	2	4	0
TG		11	11	6	8
TI	2	7	9	9	7
UR	1		1	1	1
VD	15		15	13	21
VS	4		4	7	7
ZG		6	6	3	4
ZH		42	42	46	19
FL			0	1	2
Übriges Ausland	1	1	2	1	2
	<b>104</b>	<b>113</b>	<b>217</b>	<b>203</b>	<b>171</b>

Tabelle 6: Bewerbungen 2003 für Veterinärmedizin nach Wohnkantonen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

## 5.5 Testabsolvierung nach Alter und Geschlecht

Die Zahl der Frauen überwiegt wie bisher auch im Jahr 2003 in den Fächern Human- und Veterinärmedizin. In Veterinärmedizin ist das Verhältnis bereits auf über 4/5 Frauen zu weniger als 1/5 Männer angewachsen.

	Geburtsjahre			Total
	1959 – 1979	1980 – 1982	1983 – 1987	
<b>Humanmedizin</b>				
<b>männlich</b>	27 39.7%	91 42.9%	198 33.5%	<b>316</b> 36.3% von Gesamt 1998 44.8%
<b>weiblich</b>	41 60.3%	121 57.1%	393 66.6%	<b>555</b> 63.2% von Gesamt 1998 55.2%
<b>Gesamt</b>	<b>68</b> 7.8% v. Gesamt 1998 10.4%	<b>212</b> 24.3% v. Gesamt 1998 29.5%	<b>591</b> 67.9% v. Gesamt 1998 60.1%	<b>871</b>
<b>Veterinärmedizin</b>				
<b>männlich</b>	5 26.3%	12 20.0%	23 16.7%	<b>40</b> 18.4% von Gesamt 1999 24.2%
<b>weiblich</b>	14 73.7%	48 80.0%	115 83.3%	<b>177</b> 81.6% von Gesamt 1999 75.8%
<b>Gesamt</b>	<b>19</b> 8.8% v. Gesamt 1999 9.9%	<b>60</b> 27.6% v. Gesamt 1999 26.9%	<b>138</b> 63.6% v. Gesamt 1999 63.2%	<b>217</b>

Tabelle 7: Testabsolventen nach Geburtsjahr und Geschlecht 2003 und im Vergleich zum ersten Jahr des NC pro Disziplin



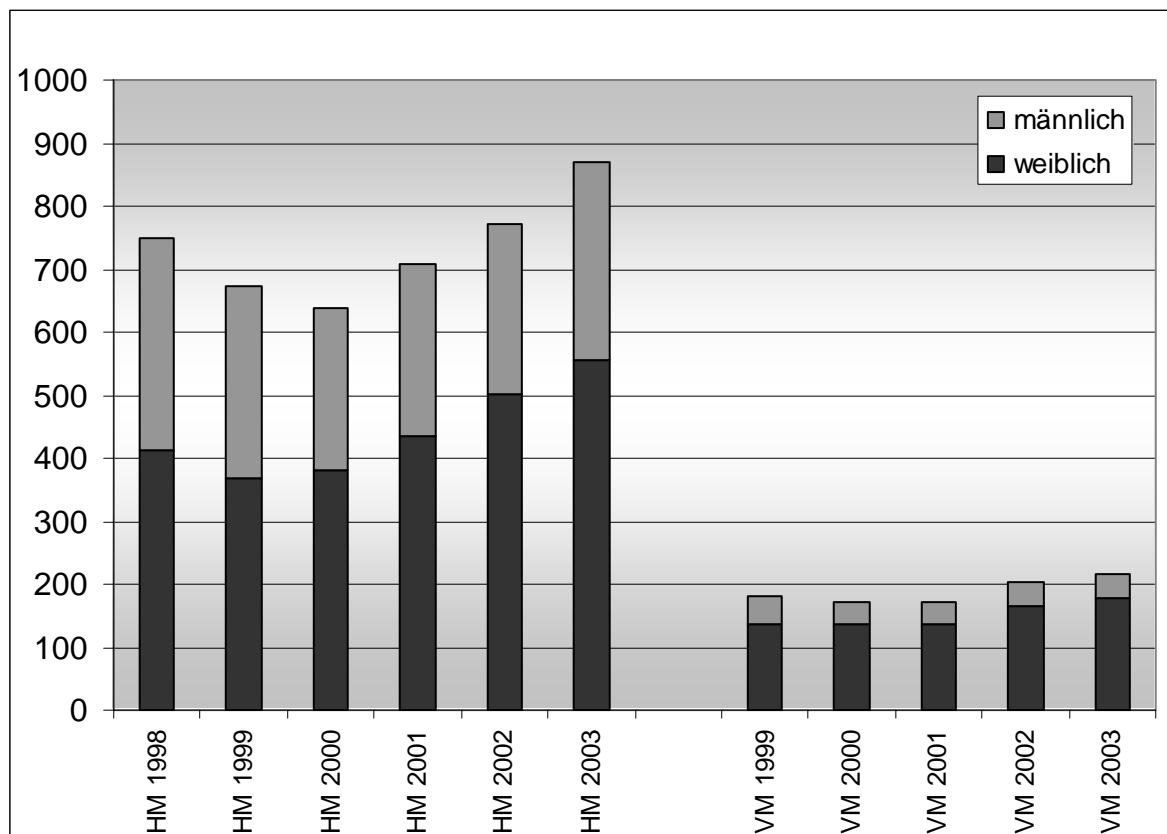


Abbildung 10: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM) 1998 bis 2003 und Veterinärmedizin (VM) 1999 bis 2003

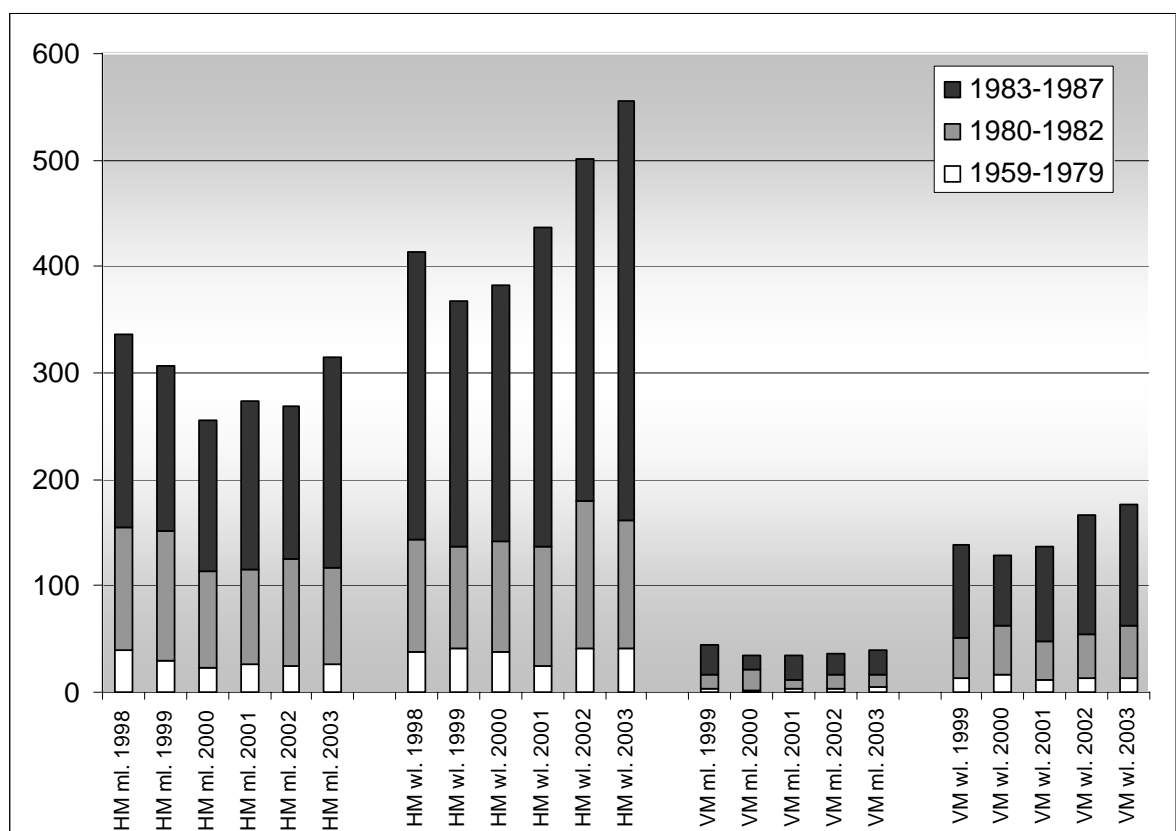


Abbildung 11: Zusammensetzung der Bewerbergruppen nach Geschlecht ml: männlich wl: weiblich und Alter (Geburtsjahre für Kohorte 2003 – 1998 bis 2002 jeweils entsprechend versetzt)

Insgesamt ist aufgrund der unterschiedlichen Zusammensetzung der Bewerbergruppe wiederum nicht damit zu rechnen, dass die Mittelwerte des Testwertes bei globaler Betrachtung für die Geschlechter gleich sind – das Verhältnis beider Teilgruppen ist nicht 50 : 50.

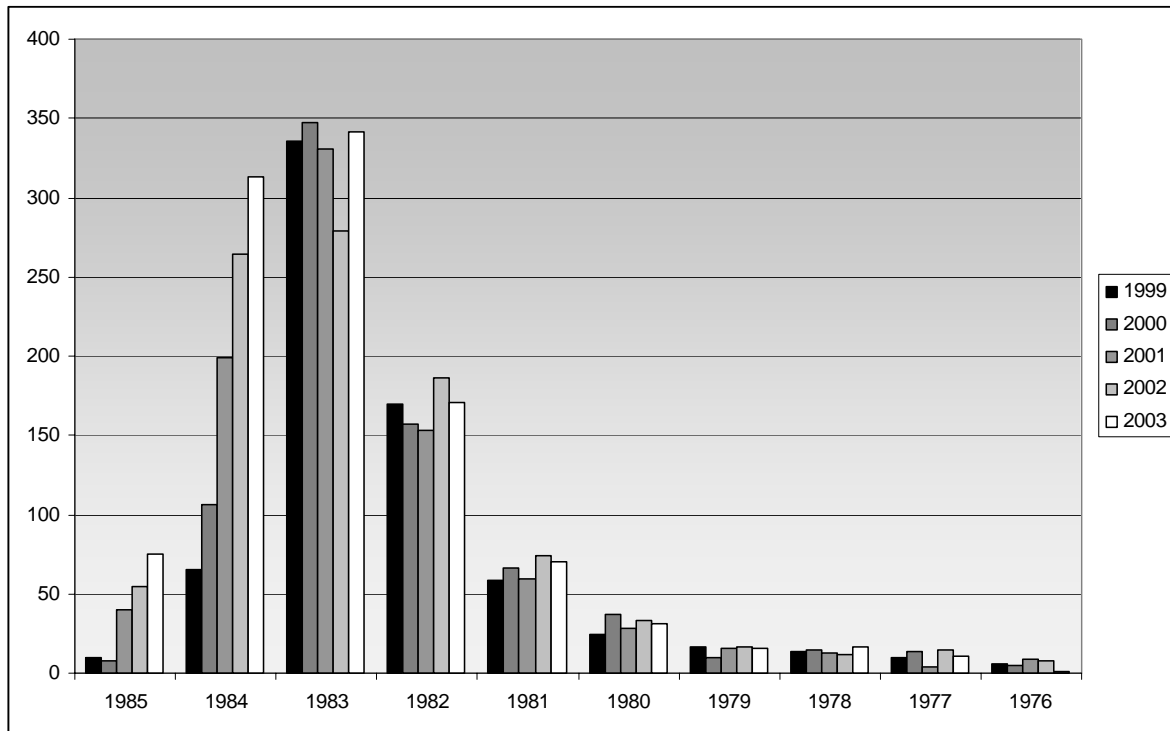


Abbildung 12: Bewerbungen nach Geburtsjahren für 2003 und Entsprechungen für die Jahre 1999 bis 2002

## 5.6 Übernahme des Testergebnisses aus 2002

In Humanmedizin haben 3 Personen (Testwerte 93, 92 und 91), welche ihr Ergebnis aus dem Vorjahr übernommen haben, keinen Studienplatz in der ersten Runde erhalten. In Veterinärmedizin ist dies eine Person (Testwert 86).

Disziplin	Jahr	N	Minimum	Maximum	Mittelwert	Stand.- abw.	Zugelassen 1. Runde
Human- medizin	2000	33	85	119	<b>102.85</b>	7.25	33
	2001	27	94	119	<b>102.96</b>	7.28	27
	2002	40	83	118	<b>104.03</b>	9.16	40
	2003	47	91	121	<b>103.70</b>	8.27	44 (3 nicht)
Vete- rinär- medizin	2000	5	97	107	<b>102.00</b>	3.81	5
	2001	5	89	112	<b>102.80</b>	9.52	5
	2002	3	94	109	<b>101.33</b>	7.51	3
	2003	11	86	118	<b>102.9</b>	9.0	10 (1 nicht)

Tabelle 8: Testwert für Personen mit Übernahme des Testergebnisses aus dem Vorjahr

Erstmals werden 4 Personen aufgrund des Testwertes des Vorjahres nicht zugelassen, weil der Grenzwert im Jahr 2003 höher liegt.

## 5.7 Erneute Testteilnahme 2003

Erste Testteilnahme	Zweite Testteilnahme				
	1999	2000	2001	2002	2003
1998	12	2	1	1	1
1999		19	3	0	0
2000			12	0	3
2001				11	3
2002					<b>32</b>
Gesamt	12	21	16	12	<b>39</b>

Tabelle 9: Wiederholte Testteilnahme: Anzahl der Personen

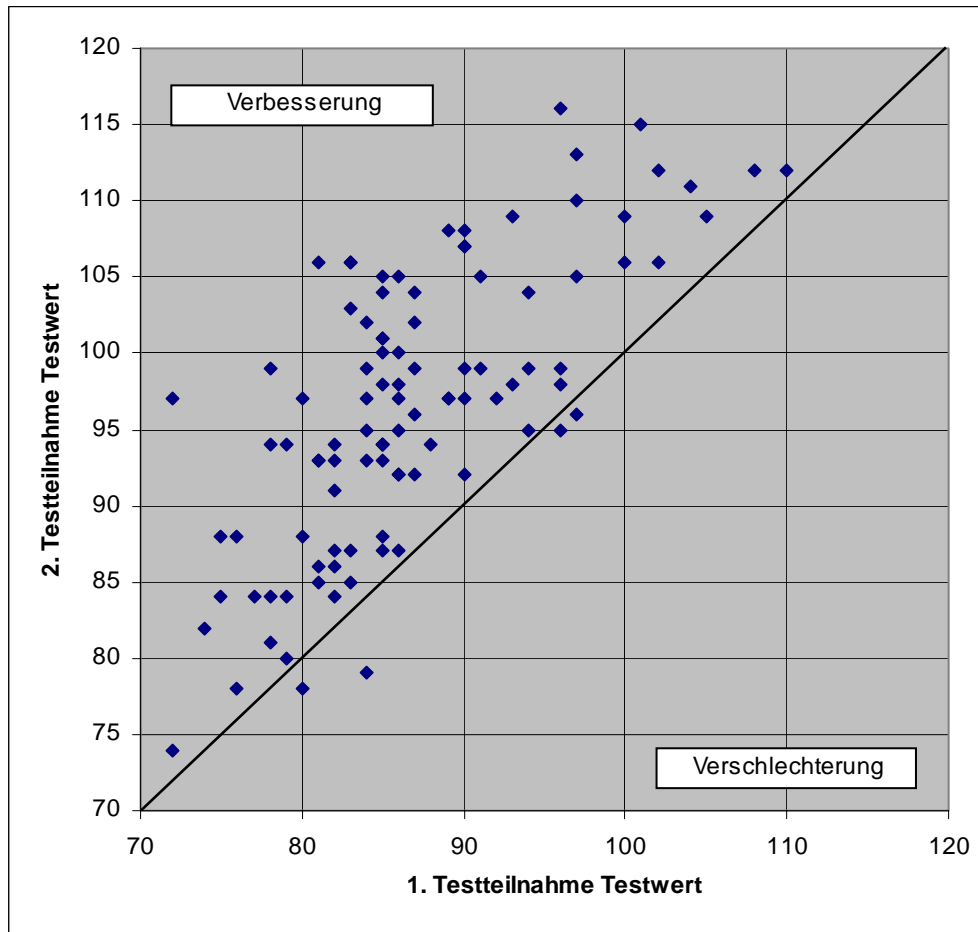
Es ist möglich, erneut zum Test anzutreten und dann das zweite Ergebnis für die Zulassung zu verwenden. Personen, die sich später als ein Jahr nach der ersten Testteilnahme noch einmal bewerben, müssen in jedem Falle noch einmal am Test teilnehmen. Neben der generellen Zulassung hängt vom Testergebnis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

Insgesamt 100 Personen haben seit 1998 zweimal am Test teilgenommen. Die Mehrzahl davon trat bereits im Folgejahr noch einmal an. Dieser Personenkreis könnte das Testergebnis auch übernehmen.

	N	Minimum	Maximum	Mittelwert	Stand.-abweichung
<b>Testwert 1. Teilnahme</b>	100	72	110	<b>86.7</b>	7.8
<b>Testwert 2. Teilnahme</b>	100	74	116	<b>96.6</b>	9.3
<b>Veränderungen</b>	100	-5	25	<b>9.7</b>	6.5

Tabelle 10: Testwertvergleich zwischen erster und zweiter Testteilnahme

Die Korrelation zwischen beiden Messungen beträgt 0.72. 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.



**Abbildung 13:** Testwerte für Personen, die zweimal am Test teilnahmen (1998 bis 2003)

Die insgesamt deutliche Leistungsverbesserung im zweiten Anlauf hängt sicher mit der besseren Vorbereitung auf den EMS zusammen. Eine Befragung zur Vorbereitung in diesem Jahr brachte das Ergebnis, dass sich ein Personenkreis sehr wenig auf den Test vorbereitet. Da als Vorbereitungsempfehlung die Durchführung eines EMS unter realen Zeit-anforderungen gehört, würden diese Personen bei der wiederholten Testdurchführung ziemlich sicher profitieren.

Die insgesamt 90 Personen im Jahr 2003, die bereits in den Vorjahren zum Test angetreten sind, wurden bisher dort als „Rückzüge“ geführt. Es handelt sich vielmehr um „Aufschübe“ des Studienwunsches.

## 5.8 Sprachgruppen

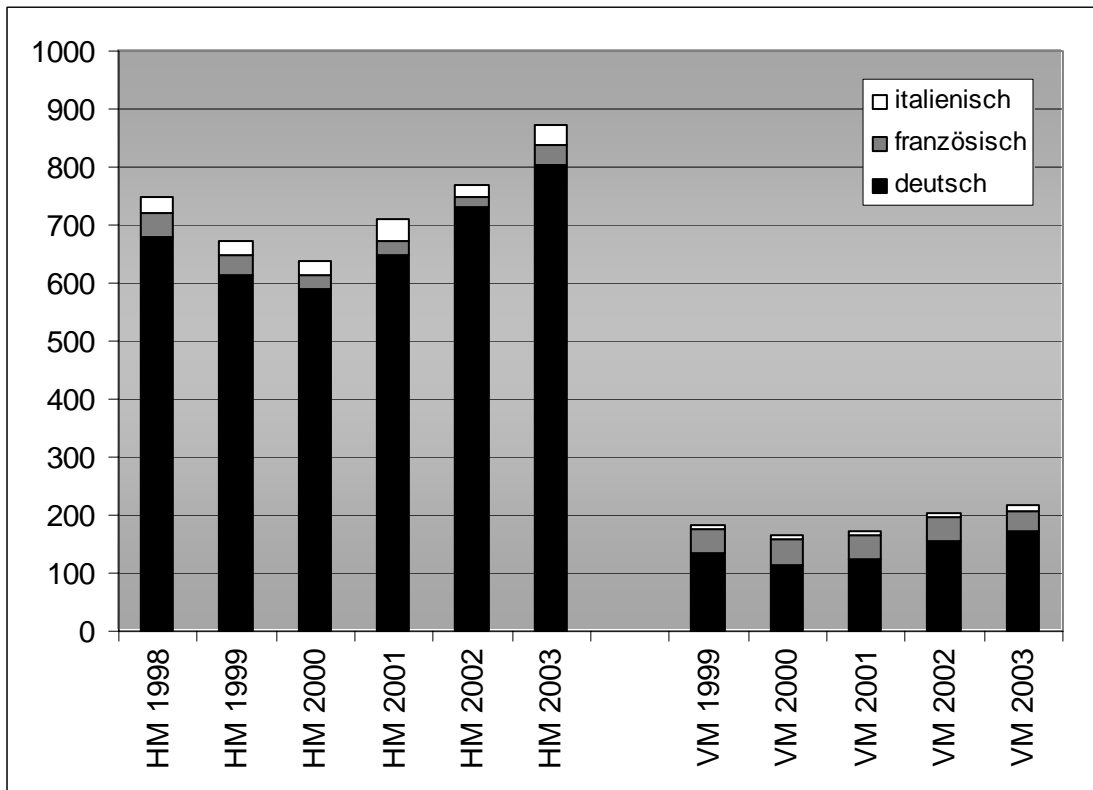


Abbildung 14: Vergleich der Sprachgruppen 1998 bis 2003, Absolutzahlen

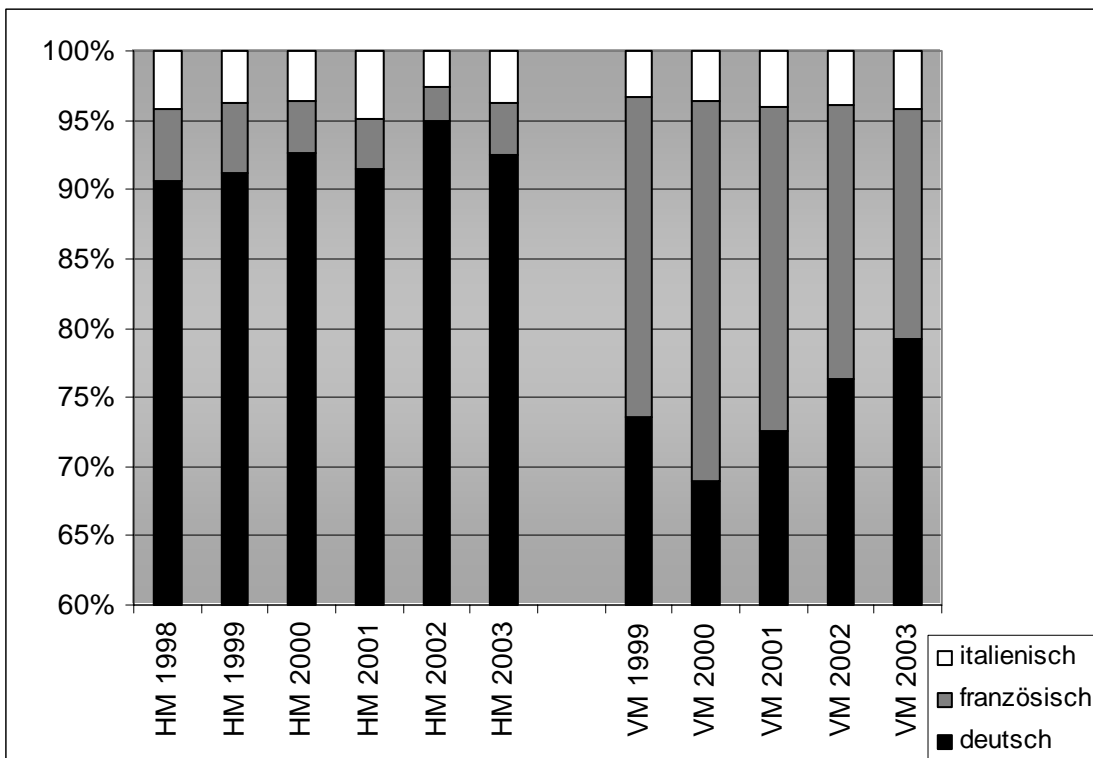


Abbildung 15: Vergleich der Sprachgruppen 1998 bis 2003, Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%)

Disziplin x Sprache x Geschlecht 2003			Klassen Geburtsjahr		
			1959-1979	1980-1982	1983-1987
Humanmedizin	deutsch	männlich	25	85	180
		weiblich	40	111	364
	französisch	männlich	1	5	6
		weiblich	1	5	15
	italienisch	männlich	1	1	12
		weiblich		5	14
Veterinärmedizin	deutsch	männlich	5	8	13
		weiblich	14	41	91
	französisch	männlich		4	9
		weiblich		5	18
	italienisch	männlich			1
		weiblich		2	6

Tabelle 11: Sprachgruppen 2003 nach Alter und Geschlecht aufgeschlüsselt

Die Zusammensetzung der französisch- und italienischsprachigen Gruppe bleibt mit den Vorjahren vergleichbar, der Zuwachs begründet sich vor allem aus der deutschsprachigen Gruppe.

## 5.9 Alter und Maturitätsjahr

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.

Geburts-jahr	Maturitätsjahr																			To-tal		
	78	83	84	85	86	87	89	91	92	93	94	95	96	97	98	99	00	01	02		03	
1959	1																					1
1960																					1	1
1961																			1			1
1964		1	1																			2
1965				1																		1
1966					1																	1
1967																			1	3		4
1968										1												1
1969						1		1													1	3
1970																					2	2
1971							1	1														2
1972								1							1	1						3
1973									1	2	1	1							2			7
1974										1		1									5	7
1975												2	2	1							1	6
1976														1								1
1977													1	2		1	1		2	4		11
1978														2	5	2		1	2	5		17
1979															3	5	2	3		3		16
1980																3	11	6	2	9		31
1981																1	3	20	26	20		70
1982																	1	14	10 2	54		171
1983																		5	17 9	157		341
1984																			44	269		313
1985																			1	72		73
1986																				1		1
1987																				1		1
	1	1	1	1	1	1	1	2	2	4	1	4	3	6	9	13	18	49	36 2	607		1088

Tabelle 12: Geburtsjahr und Maturitätsjahr. Dicke Linien trennen die Altersgruppen (und bei der ältesten Gruppe zusätzlich eine Teilung nach dem Maturitätszeitpunkt), die in 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 13 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Untertests. Aufgabenbeispiele finden Sie im Anhang ab Seite 104.

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	1200 Zeichen 20 Punkte	8
<b>MITTAGSPAUSE 60 Minuten</b>			
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
<b>Gesamttest</b>		<b>204 (178)*</b>	<b>5 Std. 7 Min.</b>

**Tabelle 13:** 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, dass 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 erfolgt<sup>1</sup>);
- 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. Auch an dieser Aufgabenentwicklung 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 13) 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.

<sup>1</sup> Zum Trainingsaspekt siehe Bericht 6 des ZTD, sowie die Argumentationen unter [www.unifr.ch/ztd/ems](http://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 1200 Zeichen der Reihe nach bearbeitet werden – 600 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. Sechshundert 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.

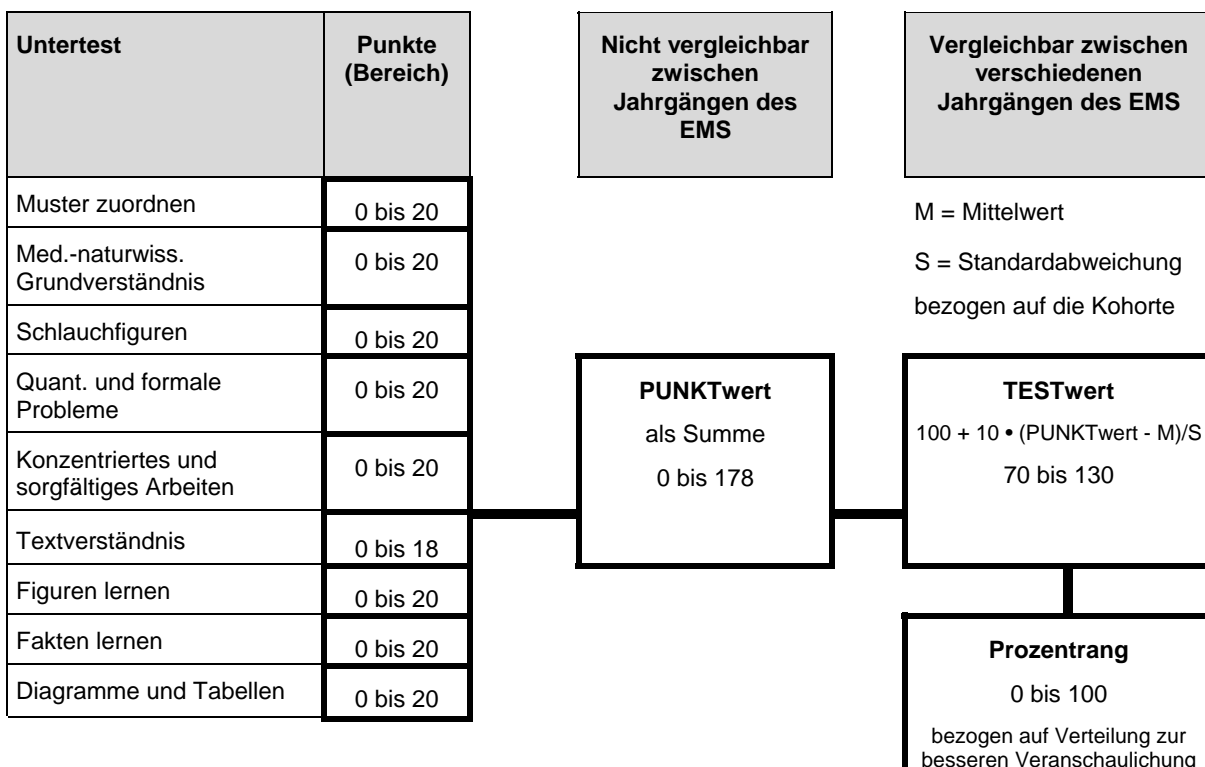


Abbildung 16: Punktwerte der einzelnen Untertests und ihre Zusammenführung über den Punktwert zum Testwert

## 7 Testanwendung in der Schweiz 2003

### 7.1 Organisation der Testabnahme

Bei der sechsten Testabnahme in der Schweiz konnte in allen Testorten auf bewährte und erfahrene Personen und Strukturen zurückgegriffen werden. In jedem Testlokal sind entweder der Testleiter oder deren Stellvertreter mindestens einmal vorher an der Testabnahme beteiligt gewesen. Aus den Protokollen ergeben sich keine Hinweise auf irreguläre Vorkommnisse oder Bedingungen. Die Testabnahme kann daher in allen Testlokalen als uneingeschränkt gültig angesehen werden.

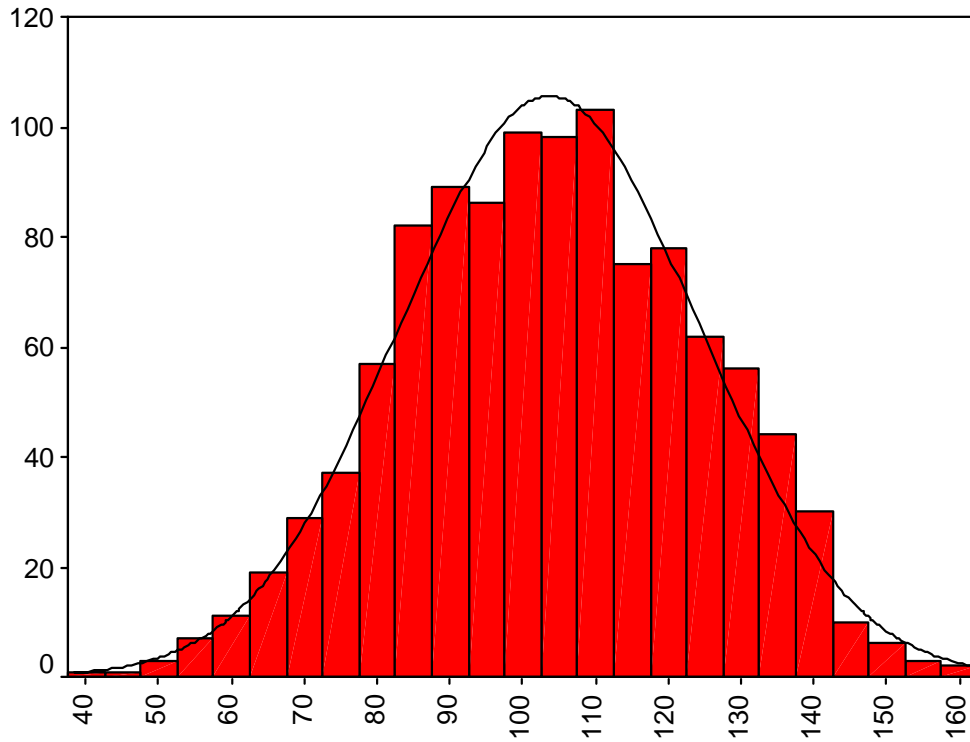
### 7.2 Verteilungsprüfung

Der Punktwert folgt wiederum einer Normalverteilung (Tabelle 14). Metrische Verfahren und die Standardisierung zum Testwert durch lineare z-Transformation sind daher zulässig. Die Punktwerte der Einzeltests sind wie auch in früheren Jahrgängen nicht normalverteilt (Abbildung 17 und Abbildung 18).

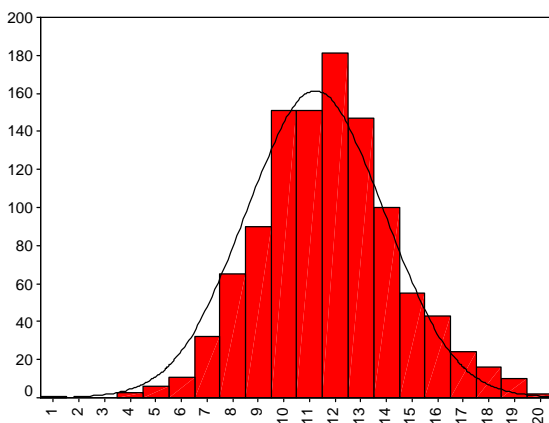
	Punkt- wert	Muster- zuord- nen	Med.- naturw. Grundv.	Schlauch- figuren	Quant.u. formale Probl.	Textver- ständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konz. u. sorgf. Arbeiten
Mittelwert	103.69	10.72	10.37	12.52	10.40	10.09	13.43	11.26	11.49	13.41
Stand.- abweich.	20.52	2.69	3.66	3.80	3.59	3.88	3.92	3.85	3.42	3.63
extremste Differenz	0.03	0.09	0.06	0.06	0.08	0.08	0.08	0.08	0.08	0.09
Kolmog.- Smirnov Z	0.87	3.11	2.08	1.97	2.75	2.52	2.55	2.67	2.51	2.97
Asymp. Sig.	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tabelle 14: Verteilungsprüfung auf Normalverteilung (Kolmogorov-Smirnov Test)

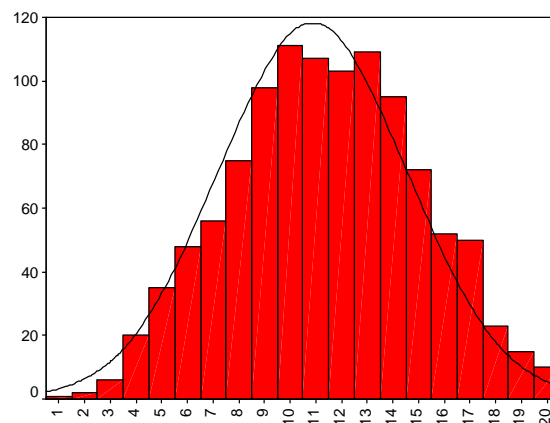
Erneut hat ein grösserer Prozentsatz der Teilnehmer im Untertest „Figuren lernen“ die maximal mögliche Wertung von 20 Punkten erreicht. Gleiches gilt für den Untertest „Konzentriertes und sorgfältiges Arbeiten“. Bei letztgenanntem Untertest kann die Übung ein Faktor sein – es wird bekanntlich empfohlen, diesen Untertest mehrfach vorher zu üben, bis ein stabiles Leistungsniveau erreicht wird – die sensomotorische Koordination als die zugrunde liegende Leistung kann durch Übung verbessert werden. Beim visuellen Gedächtnis sind solche Effekte weniger zu erwarten, da die zu lernenden Muster sinnfrei sind. Eine Kreuztabellierung der Punktwerte in beiden Untertests (vgl. Tabelle 15) zeigt, dass maximale Punktwerte in **einem** der beiden Untertests nicht zufällig mit maximalen Punktwerten im **anderen** Untertest verbunden sind – die Punktwerte dort streuen ebenso wie die Leistungen für andere Punktwerte.



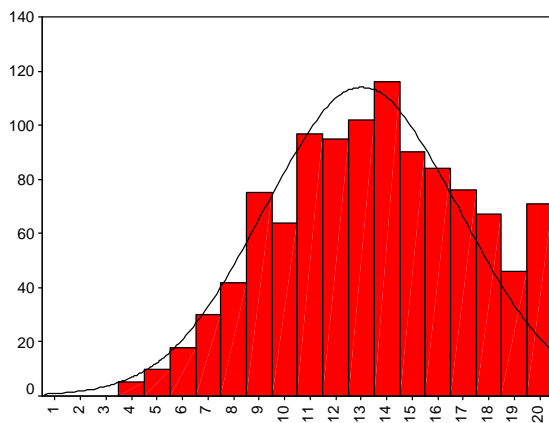
Punktwert



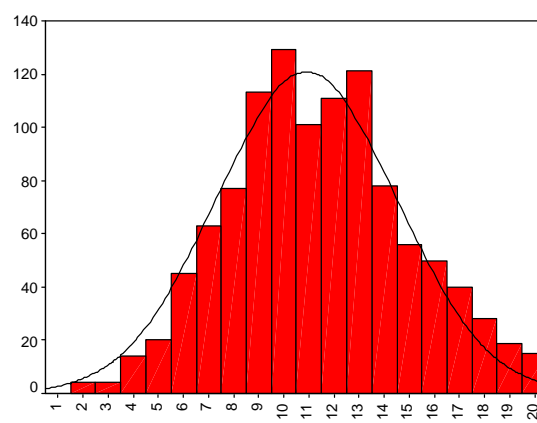
Muster zuordnen



Med.-naturw. Grundverständnis

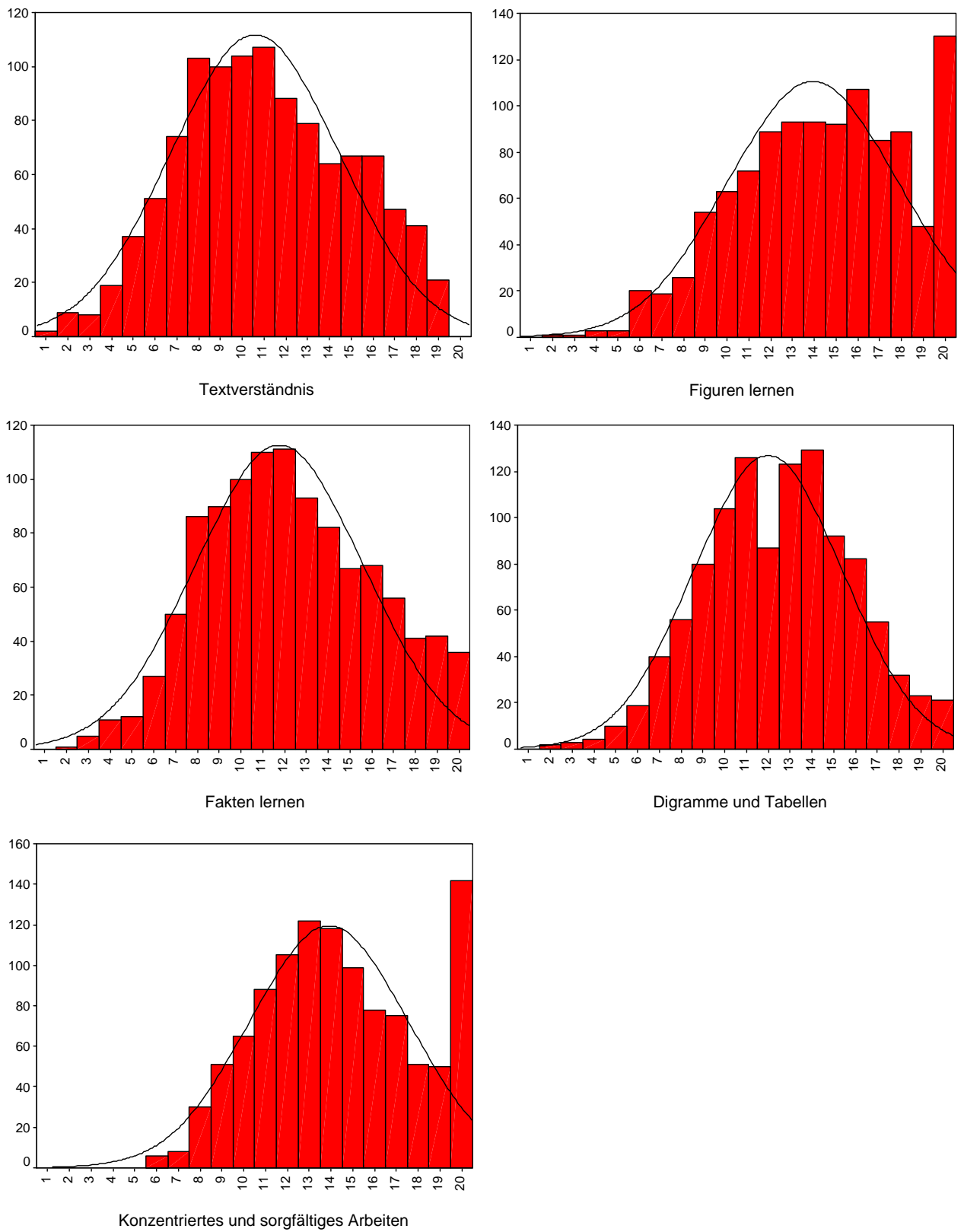


Schlauchfiguren



Quantitative u. Formale Probleme

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



**Abbildung 18:** Häufigkeitsverteilungen für Punkte der Untertests 5-9

Da für die Zulassung bei Anwendung des NC eine Differenzierung immer noch im unteren Leistungsbereich notwendig ist, sind daraus keine negativen Effekte für den Testwert abzuleiten (beide Untertests zeigen eine ausreichende Trennschärfe bezüglich des Gesamtprofils).

		Konzentriertes und sorgfältiges Arbeiten																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Figuren lernen	1					1												1
	2								1									1
	3		1	1					1									3
	4					1			2									3
	5	2	3	2	2	2	2	2	3	1	1							20
	6	1		5	4	3	3	2			1							19
	7				3	6	4	3	3	2	2	1				2		26
	8	1		2	7	6	6	8	6	8	3		2	3		2		54
	9		2	1	8	4	3	4	11	10	6	4	3	2	1	3	1	63
	10			7	8	4	9	8	9	8	4	4	5	4		1	1	72
	11		2	2	3	9	14	12	8	10	8	4	3	6	2	5	1	89
	12			4	6	4	6	12	10	9	8	6	9	8	5	6		93
	13	1		1	6	4	12	15	8	9	5	12	3	5	6	5	1	93
	14			1	2	5	4	7	20	12	12	5	9	1	6	5	3	92
	15			1		5	8	12	18	15	12	11	4	7	7	5	2	107
	16	1		1	1	6	6	6	7	7	11	10	11	4	6	5	3	85
	17			1	1	3	4	6	6	7	9	11	12	4	8	9	8	89
	18					2	2	2	2	13	4	1	4	2	1	8	7	48
	19			1			4	6	4	4	8	3	5	3	1	10	10	59
	20						1		3	3	5	6	5	2	7	22	17	71
		6	8	30	51	65	88	10	12	11	99	78	75	51	50	88	54	108
								5	2	8								8

Tabelle 15: Kreuztabelle für Punktwerte der Untertests „Figuren lernen“ und „Konzentriertes und sorgfältiges Arbeiten“

### 7.3 Vergleich der Testfassungen 1998 bis 2003

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.

Jeder **Testwert** bedeutet in jedem Jahr also das Gleiche und er kann deshalb ohne Benachteiligung ins Folgejahr übernommen werden, wenn ein späterer Studienantritt gewünscht wird.

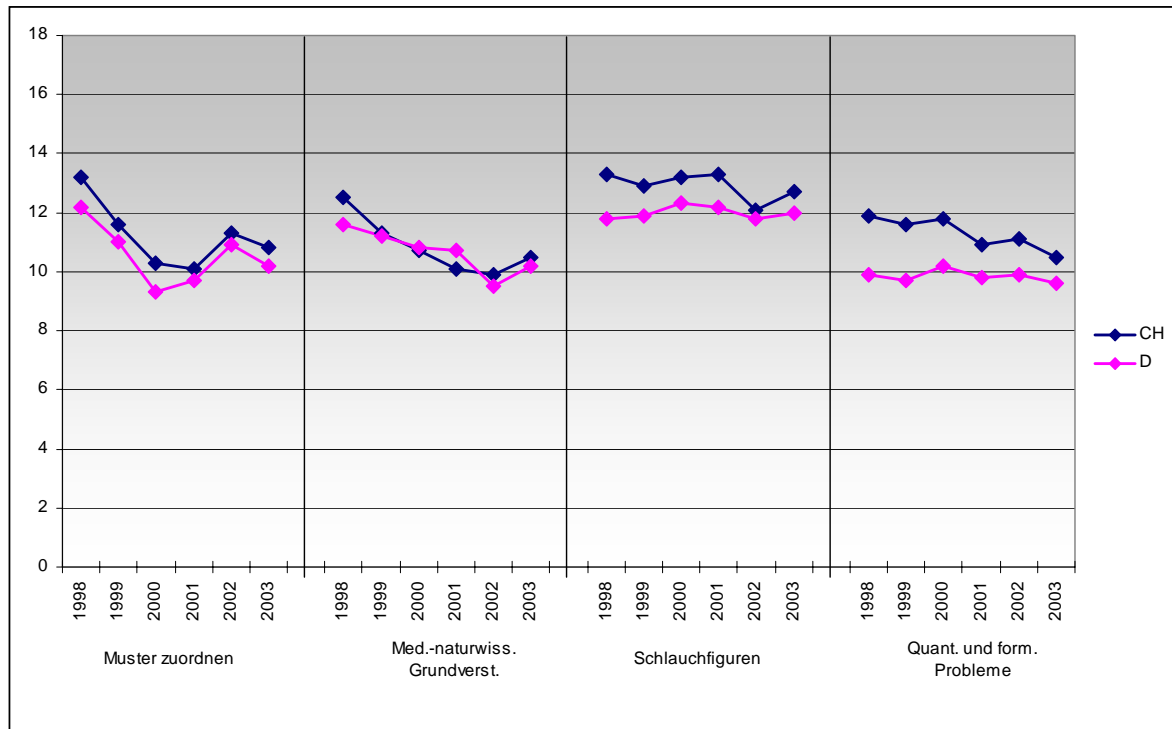
Die „absolute“ Schwierigkeit als **Punktwert** unterliegt andererseits über die Jahre geringfügigen Schwankungen. Tabelle 16 zeigt den Vergleich für die bisherigen Testdurchführungen in der Schweiz mit den entsprechenden geschätzten Punktwerten aus Deutschland. Zur Schätzung wurden die aus Deutschland berichteten Item-Schwierigkeiten herangezogen. Für die Berechnung der Schweizer Punktwerte wurden nur die Ergebnisse der deutschsprachigen Gruppe berücksichtigt.

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

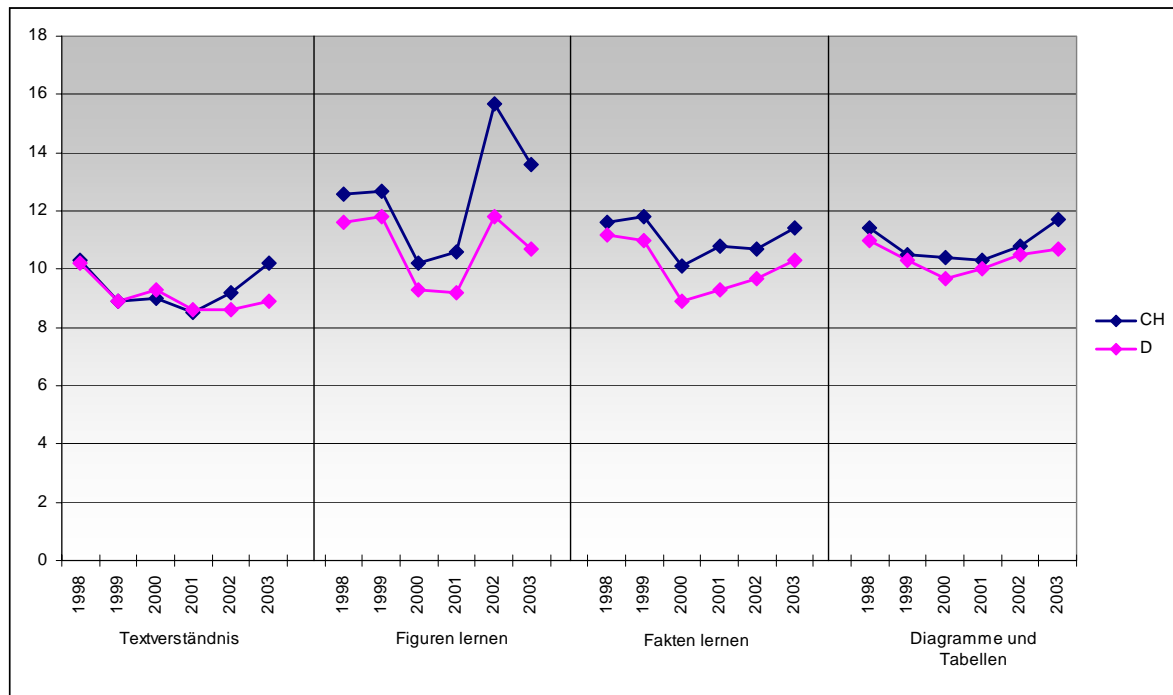
**Tabelle 16:** Mittelwerte (Punkte) der Untertests 1998 bis 2003 deutsche Sprachgruppe Schweiz und Schätzung aufgrund der Item-Schwierigkeitswerte aus Deutschland (HM: Humanmedizin)

Die Gesamtpunktzahl „Schweiz“ liegt in der Grössenordnung des letzten Jahres, mit leicht steigender Tendenz. Der Vergleich mit dem „Erwartungswert“ aus Deutschland, der gegenüber der letztjährigen Version praktisch identisch ist, deutet auf eine vergleichbare Schwierigkeit des Tests hin. Insgesamt kann man daraus auch schlussfolgern, dass in diesem Jahr trotz grösserer Anmeldezahlen kein Rückgang des Leistungsniveaus der Bewerber vorliegt.

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 2003 (Untertests 1-4)



**Abbildung 20:** Mittelwertvergleiche der Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2003 (Untertests 5-8)



Erneut liegen die Ergebnisse im Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“ leicht über den deutschen Vergleichswerten. Im Untertest „Textverständnis“ verdeutlicht sich der 2002 erstmals aufgetretene Unterschied.

Die Kurven für den Untertest „Figuren lernen“ verlaufen weiterhin fast parallel. Der grosse Leistungsunterschied aus 2002 tritt dieses Jahr geringer, aber immer noch deutlich auf.

Nach wie vor ist der Untertest „Textverständnis“ der schwierigste Untertest, wobei dort eine deutliche Leistungssteigerung zu beobachten war, so dass die Ergebnisse jetzt in der Grössenordnung der Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“ und „Quantitative und formale Probleme“ liegen.

## 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.

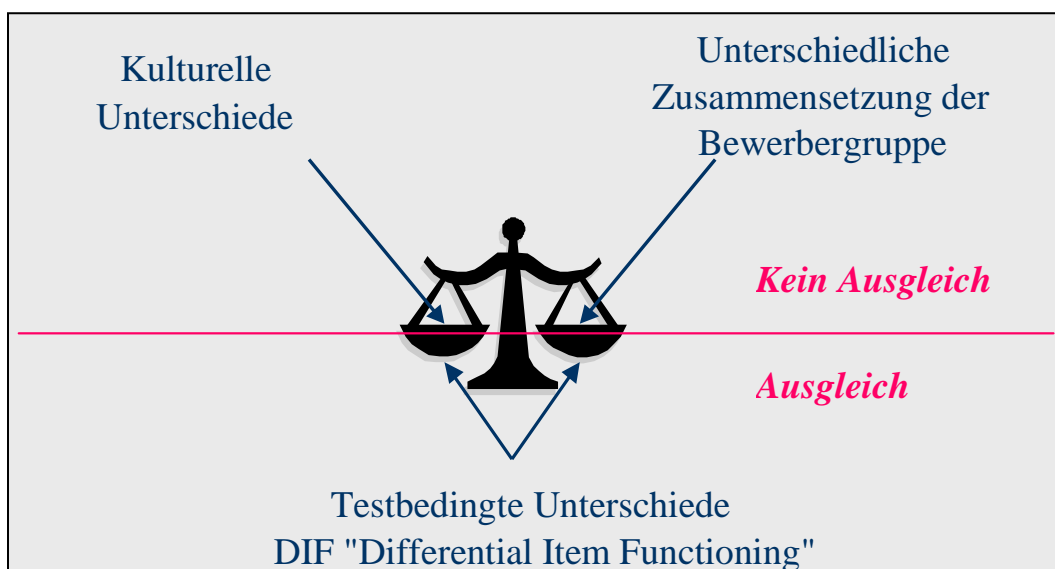


Abbildung 21: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen

Bei den „oberen“ beiden Ursachen handelt es sich um „echte“ Personen-Unterschiede, 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-Verfahren<sup>1</sup> angewendet – eine ausführliche Beschreibung kann beispielsweise in Bericht 5 (1999) nachgelesen werden.

<sup>1</sup> 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 Lösungs-

schwierigkeit 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. Die Ergebnisse der italienisch- und französischsprachigen Gruppen liegen praktisch gleichauf (Abbildung 22).

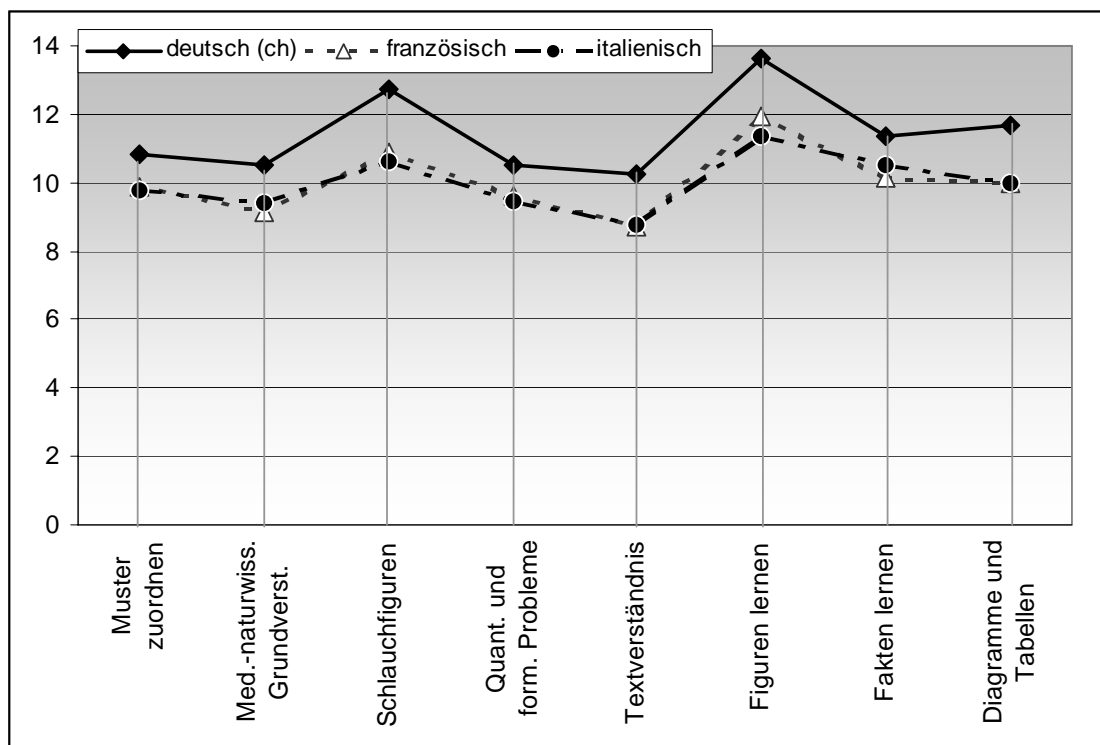
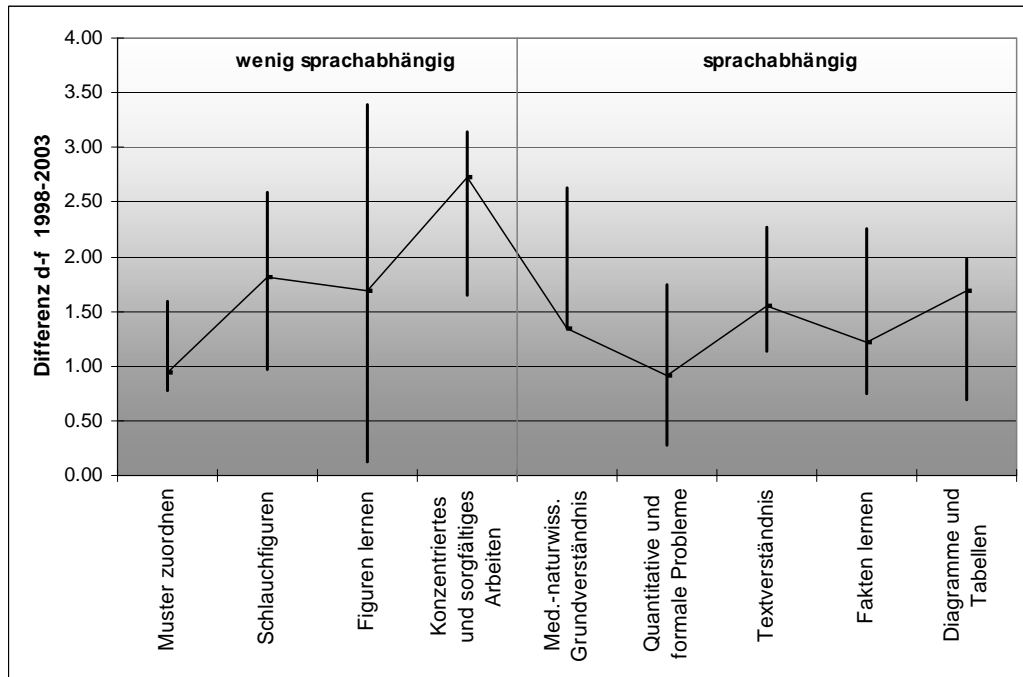


Abbildung 22: Mittelwerte der Untertests für die Sprachgruppen 2003 (unkorrigiert)

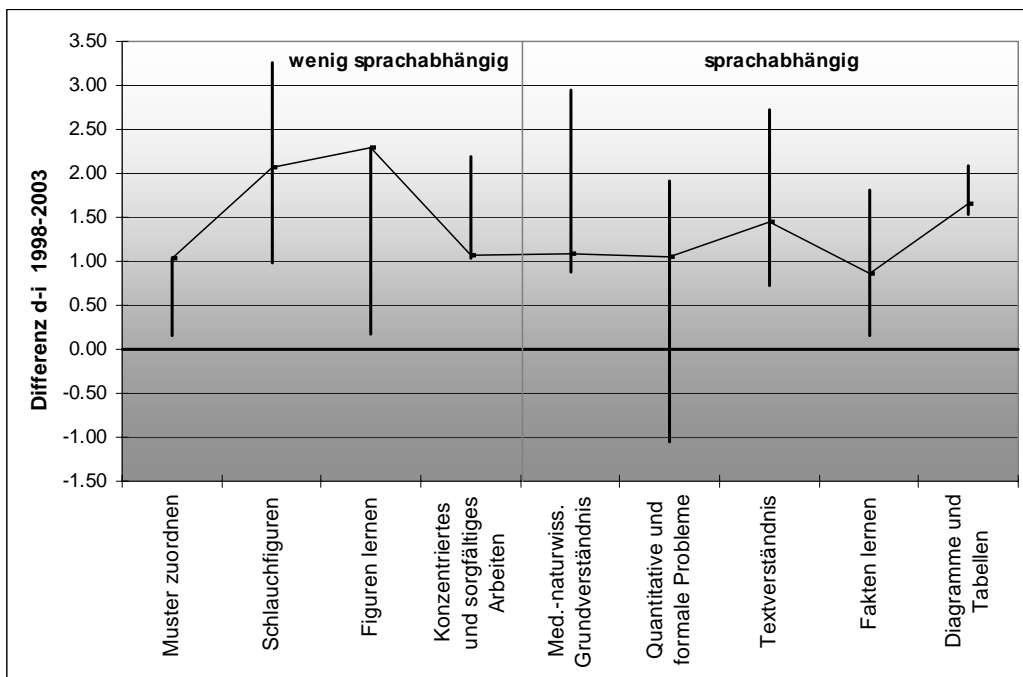
Die Abbildung 23 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern für alle Testanwendungen seit 1998. Die senkrechten Linien bezeichnen die Spannweite (Minimum bis Maximum) der Differenzen über sämtliche Jahrgänge, die Verbindungslinie steht für die Differenzen aus dem aktuellen Jahrgang. Die Differenz fällt im Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“ geringer aus als in den bisherigen Jahrgängen. In den sprachabhängigen Untertests liegen die unkorrigierten Differenzen insgesamt auf tieferem Niveau als in den wenig sprachabhängigen Untertests.



**Abbildung 23:** Differenz Testwert deutsch-französisch 2003 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2003

Im Vergleich zwischen deutsch- und italienischsprachigen Kandidaten werden in den Untertests „Figuren lernen“ und „Muster zuordnen“ neue Höchstwerte erreicht (Abbildung 24). Vergleichsweise gute Ergebnisse erzielten die italienischsprachigen Kandidaten andererseits in den Untertests „Konzentriertes und sorgfältiges Arbeiten“, „Medizinisch-naturwissenschaftliches Grundverständnis“ und „Diagramme und Tabellen“.

**Abbildung 24:** Differenz Testwert deutsch-italienisch 2003 (unkorrigiert) und Spannweite



dieser Differenz (min-max) der Jahre 1998 bis 2003

Insgesamt bleiben sowohl in wenig sprachabhängigen wie auch in sprachabhängigen Untertests Differenzen nachweisbar – tendenziell fallen diese in den wenig sprachabhängigen Untertests sogar deutlicher aus als in den sprachabhängigen Untertests.

Differenzen treten zwischen den einzelnen Jahren in allen Untertests auf, ebenso sind die Schwankungsbreiten zwischen sprachabhängigen und wenig sprachabhängigen Untertests vergleichbar. Adaptationsbedingte Textunterschiede sind deshalb als massgebliche Ursache dafür sehr unwahrscheinlich. Durch das Korrekturverfahren werden vorhandene testbedingte Differenzen zwischen den Sprachversionen ausgeglichen. Dennoch erreichen die Sprachgruppen auch nach dem Ausgleich nicht in jedem Fall vergleichbare Mittelwerte. Aufgrund des hierarchischen Faktormodells mit hohem Generalfaktoranteil an der Varianzaufklärung bleiben die Unterschiede mindestens so hoch wie in den (nicht ausgeglichenen) wenig sprachabhängigen Untertests.

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Testwert ** $D > (F, I)$	deutsch	977	100.63	10.04	70	128
	französisch	69	94.59	7.22	79	110
	italienisch	42	94.43	8.94	73	118
Muster zuordnen ** $D > (F, I)$	deutsch	977	10.82	2.70	0	19
	französisch	69	9.88	2.71	3	16
	italienisch	42	9.79	1.99	5	13
Med.-naturwiss. Grundverständnis <i>homogen</i>	deutsch	977	10.50	3.66	1	20
	französisch	69	9.63	3.18	1	16
	italienisch	42	9.40	3.95	2	19
Schlauchfiguren ** $D > (F, I)$	deutsch	977	12.71	3.82	3	20
	französisch	69	10.90	3.31	3	18
	italienisch	42	10.64	2.83	5	16
Quantitative und formale Probleme <i>homogen</i>	deutsch	977	10.50	3.58	1	20
	französisch	69	9.68	3.33	2	19
	italienisch	42	9.45	3.83	3	19
Textverständnis ** $D > F$	deutsch	977	10.24	3.93	0	18
	französisch	69	8.93	2.82	2	17
	italienisch	42	8.97	3.57	1	18
Figuren lernen ** $D > (F, I)$	deutsch	977	13.63	3.94	1	20
	französisch	69	11.94	3.28	6	20
	italienisch	42	11.33	3.22	2	16
Fakten lernen <i>homogen</i>	deutsch	977	11.37	3.89	1	20
	französisch	69	10.80	2.93	6	17
	italienisch	42	10.50	3.62	3	19
Diagramme und Tabellen ** $D > (F, I)$	deutsch	977	11.66	3.41	1	20
	französisch	69	10.47	2.69	6	16
	italienisch	42	10.10	3.42	3	20
Konzentriertes und sorgf. Arbeiten ** $D > I > F$	deutsch	977	13.62	3.65	5	20
	französisch	69	10.90	2.55	6	18
	italienisch	42	12.55	2.70	7	19

Tabelle 17: Testwert und Punktwerte der Untertests für die Sprachgruppen.

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

Die Mittelwertsdifferenzen für jeden Untertest **nach** dem Sprachausgleich sind der Tabelle 17 zu entnehmen. Die varianzanalytische Prüfung zeigt nach dem Ausgleichsverfahren für die Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“ und „Fakten lernen“ keinen signifikanten Einfluss des Faktors Sprache. In den anderen sprachabhängigen Untertests („Textverständnis“, „Diagramme und Tabellen“) bestehen auch nach dem Sprachausgleich Differenzen, die nun aber mit grosser Sicherheit nicht auf adaptionsbedingte Ursachen zurückzuführen sind.

Dass nur noch diese beiden sprachabhängigen Untertests (gegenüber allen vier wenig sprachabhängigen Untertests) einen Einfluss des Faktors Sprache zeigen, spricht dafür, dass nach dem Ausgleichverfahren keine der „adaptierten“ Sprachformen eine Benachteiligung in Kauf nehmen muss.

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. Für die französischsprachige Form (Tabelle 18) zeigen sich in allen sprachabhängigen Untertests signifikante Differenzen – alle 5 Untertests werden demnach auf Aufgabenebene geprüft.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	977	10.50	3.655	<b>.003</b>	Ja
	f	69	9.16	3.333		
Quantitative und formale Probleme	d	977	10.50	3.581	<b>.041</b>	Ja
	f	69	9.59	3.358		
Textverständnis	d	977	10.24	3.926	<b>.001</b>	Ja
	f	69	8.70	2.846		
Fakten lernen	d	977	11.37	3.894	<b>.011</b>	Ja
	f	69	10.14	3.064		
Diagramme und Tabellen	d	977	11.66	3.415	<b>.000</b>	Ja
	f	69	9.97	2.828		

Tabelle 18: Signifikanzprüfung deutsche vs. französische Sprachform

Die italienische Form (Tabelle 19) weist in zwei sprachabhängigen Untertests signifikante Unterschiede auf: „Textverständnis“ sowie „Diagramme und Tabellen“. Für die übrigen Untertests erfolgt kein Sprachausgleich.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	977	10.50	3.655	.058	Nein
	i	42	9.40	3.945		
Quantitative und formale Probleme	d	977	10.50	3.581	.064	Nein
	i	42	9.45	3.833		
Textverständnis	d	977	10.24	3.926	<b>.019</b>	Ja
	i	42	8.79	3.619		
Fakten lernen	d	977	11.37	3.894	.156	Nein
	i	42	10.50	3.617		
Diagramme und Tabellen	d	977	11.66	3.415	<b>.002</b>	Ja
	i	42	10.00	3.436		

Tabelle 19: 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.

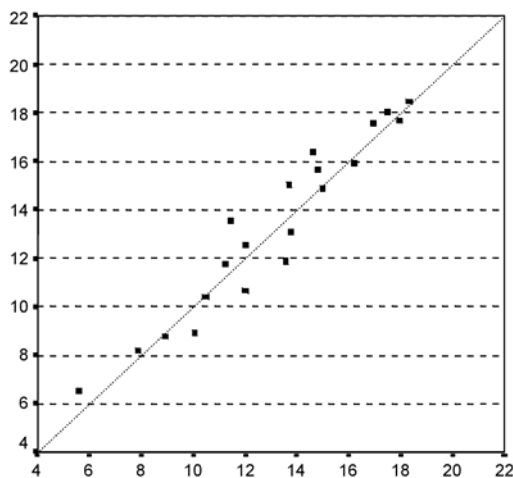
#### **7.4.2.2 Identifikation von DIF-Items und Bestimmung des Korrekturwertes**

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 „ $\Delta$ -Werte“ transformiert. Die Transformation erfolgt über die Formel

$$\Delta = 13 - 4z$$

Dies bedeutet, dass  $\Delta$  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 Medizinisch-naturwissenschaftliches Grundverständnis (MNGV)

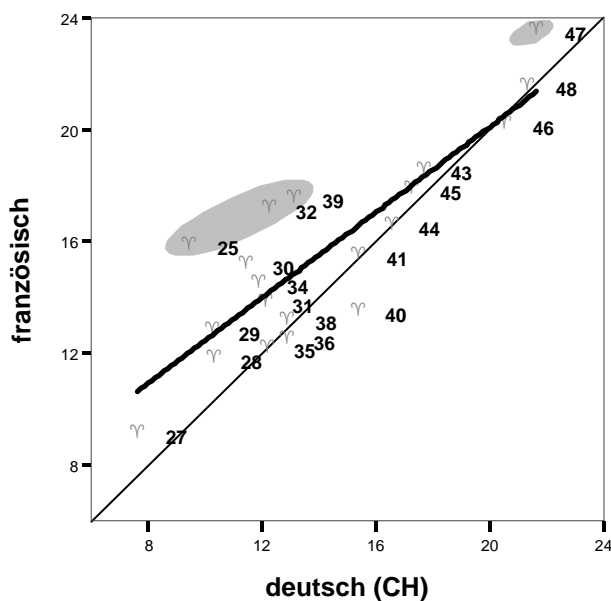


Abbildung 26 zeigt den Delta-Plot für den Untertest „MNGV“ (Vergleich zwischen deutsch- und französischsprachiger Version). Die Items 25, 32, 39 und 47 fallen hier als zu schwer auf.

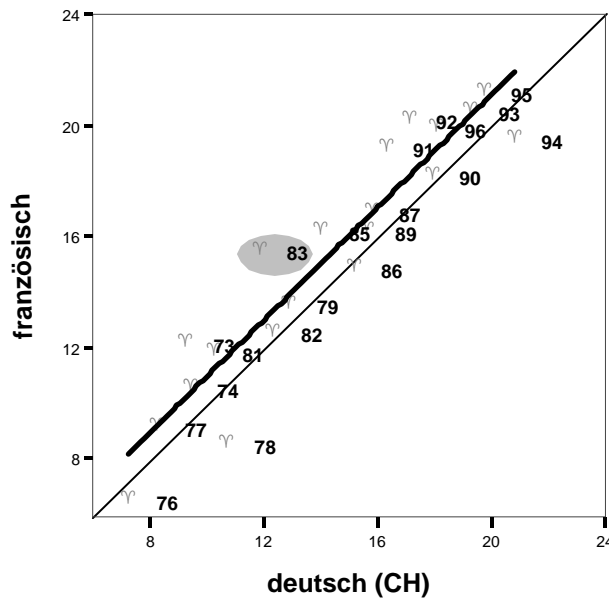
Abbildung 26: Delta-Plot für „MNGV“ (deutsch-französisch)

Maximal können in diesem Untertest den französischsprachigen Teilnehmern, welche alle vier Aufgaben nicht gelöst haben, 0.76 Punkte gutgeschrieben werden.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
25	0.73	0.45	0.64	9.43	15.83	11.35	-2.98	Kein Ausgleich	0.28	Kein Ausgleich
27	0.80	0.74	0.64	7.61	9.12	11.35	1.24			
28	0.69	0.62	0.48	10.28	11.80	15.20	0.73			
29	0.69	0.58	0.71	10.19	12.81	9.69	-0.12			
30	0.64	0.48	0.67	11.40	15.15	10.80	-1.25			
31	0.61	0.54	0.64	12.08	13.81	11.35	0.23			
32	0.61	0.39	0.52	12.20	17.17	14.10	-2.36			
34	0.62	0.51	0.55	11.85	14.48	13.55	-0.45			
35	0.61	0.61	0.64	12.13	12.14	11.35	1.59			
36	0.58	0.59	0.57	12.84	12.47	13.00	1.75			
38	0.58	0.57	0.62	12.86	13.14	11.90	1.24			
39	0.57	0.38	0.40	13.12	17.50	16.85	-2.07			
40	0.47	0.55	0.50	15.33	13.48	14.65	2.47			
41	0.47	0.46	0.50	15.37	15.49	14.65	0.90			
43	0.37	0.33	0.19	17.65	18.51	21.81	-0.12			
44	0.42	0.42	0.36	16.51	16.50	17.96	0.79			
45	0.39	0.36	0.26	17.24	17.84	20.16	0.17			
46	0.25	0.26	0.26	20.49	20.18	20.16	0.27			
47	0.20	0.12	0.10	21.60	23.54	24.02	1.41			
48	0.21	0.20	0.14	21.29	21.52	22.91	-0.73			

Tabelle 20: DIF-Analyse „Medizinisch-naturwissenschaftliches Grundverständnis“

7.4.2.4 Quantitative und formale Probleme



Die folgende Abbildung zeigt den Delta-Plot-Vergleich für den Untertest „Quantitative und formale Probleme“.

Die Regressionsgerade stellt praktisch eine Parallelverschiebung der Diagonalen dar. Einzig Item 83 liegt über dem kritischen Abstand und muss ausgeglichen werden.

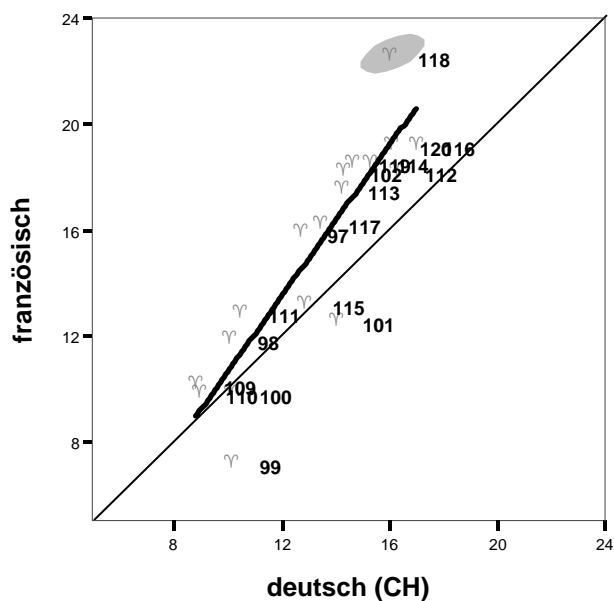
Abbildung 27: Delta-Plot für „Quantitative und formale Probleme“ (deutsch-französisch)

Der maximal erreichbare Sprachbonus beträgt hier für die französischsprachigen Kandidaten 0.16 Punkte.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
73	0.73	0.61	0.67	9.22	12.14	10.80	-1.37			
74	0.73	0.68	0.64	9.38	10.46	11.35	-0.08			
76	0.82	0.86	0.71	7.23	6.44	9.69	1.20			
77	0.78	0.74	0.76	8.22	9.12	8.59	0.03			
78	0.67	0.77	0.81	10.66	8.45	7.49	2.24			
79	0.58	0.55	0.50	12.86	13.48	14.65	0.29			
81	0.69	0.62	0.67	10.24	11.80	10.80	-0.41			
82	0.60	0.59	0.50	12.30	12.47	14.65	0.59			
83	0.62	0.46	0.50	11.82	15.49	14.65	<b>-1.86</b>		<b>0.16</b>	
85	0.53	0.43	0.38	13.98	16.16	17.41	-0.80			
86	0.48	0.49	0.62	15.14	14.82	11.90	0.97			
87	0.45	0.41	0.29	15.78	16.83	19.61	0.02			
89	0.46	0.43	0.45	15.61	16.16	15.75	0.37			
90	0.36	0.35	0.43	17.88	18.17	16.30	0.58			
91	0.43	0.30	0.31	16.27	19.18	19.06	-1.27			
92	0.40	0.26	0.29	17.08	20.18	19.61	-1.40			
93	0.30	0.25	0.29	19.23	20.52	19.61	-0.10			
94	0.24	0.29	0.17	20.77	19.51	22.36	1.70			
95	0.28	0.22	0.21	19.71	21.19	21.26	-0.23			
96	0.35	0.28	0.26	18.05	19.85	20.16	-0.47			

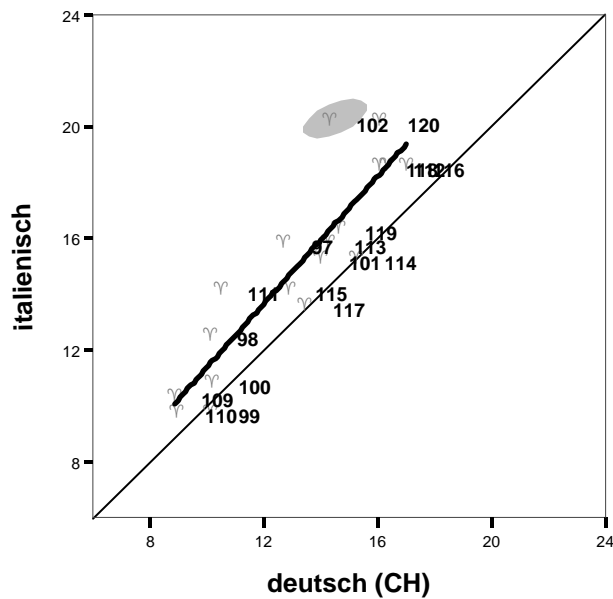
Tabelle 21: DIF-Analyse „Quantitative und formale Probleme“

#### 7.4.2.5 Textverständnis



Für den Untertest „Textverständnis“ fällt im Vergleich der deutschsprachigen mit der französischen Version (Abbildung 28) ein Item aus dem „schwierigen“ Segment bezüglich DIF auf. Hier sind 0,28 Punkte gutzuschreiben falls das Item nicht korrekt beantwortet wurde.

Abbildung 28: Delta-Plot für „Textverständnis“ (deutsch-französisch)



Das Ergebnis für die italienischsprachige Gruppe fällt ähnlich wie der vorhergehende Vergleich aus. Auch hier sind die Differenzen im schwierigeren Bereich etwas ausgeprägter, und auch hier muss nur ein Item mit 0.25 Punkten ausgeglichen werden.

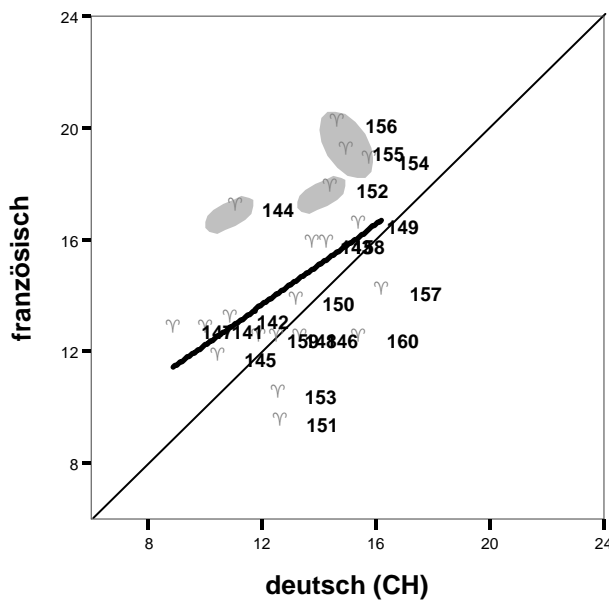
Abbildung 29: Delta-Plot für „Textverständnis“ (deutsch-italienisch)

Im Untertest „Textverständnis“ weist nur jeweils ein Item in der französisch- und italienischsprachigen Version eine auffällige Abweichung auf.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
97	0.59	0.45	0.45	12.67	15.83	15.75	-0.76	-0.83		
98	0.70	0.62	0.60	10.07	11.80	12.45	-0.59	-0.62		
99	0.70	0.83	0.71	10.12	7.11	9.69	2.14	1.23		
100	0.69	0.71	0.67	10.14	9.79	10.80	0.62	0.53		
101	0.53	0.59	0.48	14.00	12.47	15.20	2.25	0.53		
102	0.52	0.35	0.26	14.26	18.17	20.16	-0.81	<b>-2.54</b>		<b>0.25</b>
109	0.75	0.70	0.69	8.82	10.13	10.25	-0.66	-0.11		
110	0.75	0.71	0.71	8.91	9.79	9.69	-0.39	0.33		
111	0.68	0.58	0.52	10.45	12.81	14.10	-0.86	-1.42		
112	0.43	0.35	0.33	16.27	18.17	18.51	0.84	0.06		
113	0.52	0.38	0.45	14.19	17.50	15.75	-0.48	0.31		
114	0.47	0.33	0.48	15.23	18.51	15.20	-0.20	1.45		
115	0.58	0.57	0.52	12.84	13.14	14.10	0.91	0.38		
116	0.40	0.30	0.33	16.96	19.18	18.51	0.83	0.58		
117	0.55	0.43	0.55	13.43	16.16	13.55	-0.33	1.19		
118	0.44	0.16	0.33	16.01	22.53	18.51	<b>-1.87</b>	-0.13	<b>0.28</b>	
119	0.50	0.33	0.43	14.59	18.51	16.30	-0.73	0.25		
120	0.44	0.30	0.26	16.06	19.18	20.16	0.09	-1.19		

Tabelle 22: DIF-Analyse „Textverständnis“

7.4.2.6 Fakten lernen



Im Untertest „Fakten lernen“ kann erneut eine signifikante Differenz zwischen den Leistungen der deutschsprachigen und jener der französischsprachigen Kandidaten beobachtet werden. Gerade dieser eher „leicht“ zu übersetzende Text weist grössere Unterschiede zwischen den Sprachversionen auf, als etwa die komplizierteren Texte des Untertests „Textverständnis“. Besondere Abweichungen weisen die Items 144, 152, 154, 155 und 156 auf (Abbildung 31).

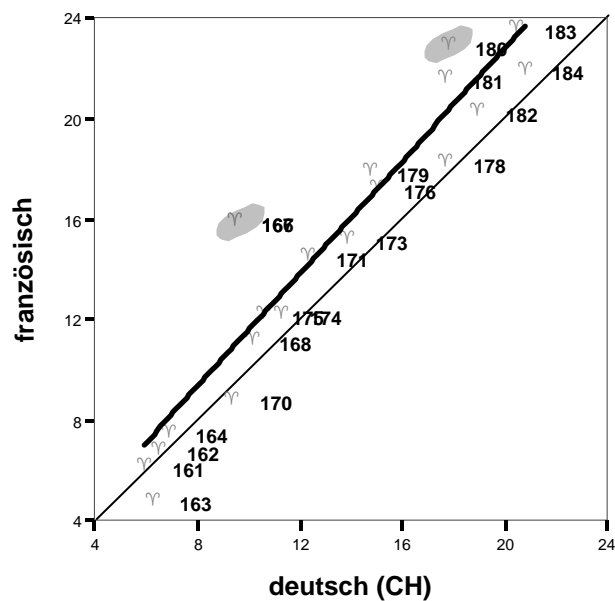
Abbildung 30: Delta-Plot für „Fakten lernen“ (deutsch-französisch)

Der maximal gutzuschreibende Punktwert beträgt 0.98 Punkte.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
141	0.70	0.58	0.50	9.95	12.81	14.65	-0.46			
142	0.67	0.57	0.64	10.83	13.14	11.35	-0.22			
143	0.54	0.45	0.38	13.72	15.83	17.41	-0.71			
144	0.66	0.39	0.57	11.02	17.17	13.00	<b>-3.37</b>		<b>0.27</b>	
145	0.68	0.62	0.64	10.40	11.80	11.35	0.62			
146	0.56	0.59	0.55	13.27	12.47	13.55	1.75			
147	0.75	0.58	0.79	8.86	12.81	8.04	-1.09			
148	0.59	0.59	0.40	12.49	12.47	16.85	1.29			
149	0.47	0.42	0.48	15.35	16.50	15.20	-0.30			
150	0.56	0.54	0.45	13.15	13.81	15.75	0.59			
151	0.59	0.72	0.62	12.58	9.46	11.90	3.80			
152	0.51	0.36	0.55	14.33	17.84	13.55	<b>-1.98</b>		<b>0.15</b>	
153	0.59	0.68	0.64	12.53	10.46	11.35	2.95			
154	0.45	0.32	0.50	15.73	18.84	14.65	<b>-1.99</b>		<b>0.13</b>	
155	0.49	0.30	0.40	14.90	19.18	16.85	<b>-2.74</b>		<b>0.18</b>	
156	0.50	0.26	0.50	14.62	20.18	14.65	<b>-3.72</b>		<b>0.24</b>	
157	0.43	0.52	0.40	16.18	14.15	16.85	2.09			
158	0.52	0.45	0.57	14.21	15.83	13.00	-0.42			
159	0.62	0.59	0.52	11.87	12.47	14.10	0.93			
160	0.47	0.59	0.38	15.37	12.47	17.41	2.98			

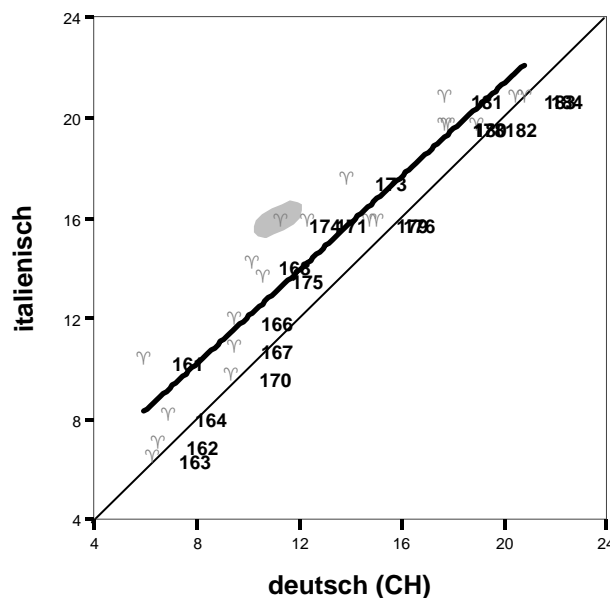
Tabelle 23: DIF-Analyse „Fakten lernen“

### 7.4.2.7 Diagramme und Tabellen



In der französischsprachigen Fassung von „Diagramme und Tabellen“ liegt nahezu eine Parallelverschiebung der Itemschwierigkeiten vor. Auffällig sind nur die Items 166, 167 (identische Werte) und 180.

**Abbildung 31:** Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch)



Auch in der italienischsprachigen Version liegt kein Hinweis auf adaptationsbedingte Differenzen vor. Einziges „problematisches“ Item ist hier Item 174.

**Abbildung 32:** Delta-Plot für „Diagramme und Tabellen“ (deutsch-italienisch)

Insgesamt können den französischsprachigen Kandidaten hier maximal 0.77, den italienischsprachigen Teilnehmern maximal 0.19 Punkte gutgeschrieben werden.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
161	0.88	0.87	0.69	5.93	6.10	10.25	0.63	-1.42		
162	0.85	0.84	0.83	6.49	6.77	6.94	0.61	1.38		
163	0.86	0.93	0.86	6.23	4.76	6.39	1.75	1.61		
164	0.84	0.81	0.79	6.87	7.44	8.04	0.44	0.84		
166	0.73	0.45	0.62	9.43	15.83	11.90	<b>-3.23</b>	-0.24	<b>0.28</b>	
167	0.73	0.45	0.67	9.43	15.83	10.80	<b>-3.23</b>	0.57	<b>0.28</b>	
168	0.69	0.65	0.52	10.14	11.13	14.10	0.43	-1.37		
170	0.73	0.75	0.71	9.34	8.78	9.69	1.39	1.31		
171	0.60	0.51	0.45	12.32	14.48	15.75	-0.18	-1.09		
173	0.53	0.48	0.38	13.86	15.15	17.41	0.53	-1.25		
174	0.65	0.61	0.45	11.28	12.14	15.75	0.61	<b>-1.80</b>		<b>0.19</b>
175	0.68	0.61	0.55	10.57	12.14	13.55	0.08	-0.67		
176	0.49	0.39	0.45	15.00	17.17	15.75	0.04	0.73		
178	0.37	0.35	0.29	17.65	18.17	19.61	1.35	-0.28		
179	0.50	0.36	0.45	14.71	17.84	15.75	-0.62	0.54		
180	0.37	0.14	0.29	17.77	22.87	19.61	<b>-1.69</b>	-0.20	<b>0.22</b>	
181	0.37	0.20	0.24	17.62	21.52	20.71	-0.90	-1.10		
182	0.32	0.26	0.29	18.90	20.18	19.61	0.94	0.58		
183	0.25	0.12	0.24	20.44	23.54	20.71	-0.14	0.82		
184	0.24	0.19	0.24	20.75	21.86	20.71	1.21	1.03		

Tabelle 24: DIF-Analyse „Diagramme und Tabellen“

#### 7.4.2.8 Effekte der Korrektur

Die durch die Ausgleichsprozedur resultierenden Bonuspunkte können Tabelle 25 entnommen werden. Durchschnittlich werden den französischsprachigen Teilnehmern 1.44 Punkte, den italienischsprachigen Teilnehmern 0.29 Punkte gutgeschrieben. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der Bonus in der französischsprachigen Gruppe beträgt also 1 bis 2 Punkte. In der italienischsprachigen Gruppe ist die maximale Gutschrift von 0.44 Punkten zu verzeichnen. Dies bedeutet, dass hier aufgrund des geringen Wertes keine Bonuspunkte vergeben werden.

		n	Minimum	Maximum	m	s
französisch	Med.-naturwiss. Grundverständnis	69	0.0	0.76	0.47	0.24
	Quant. und form. Probleme	69	0.0	0.16	0.09	0.08
	Textverständnis	69	0.0	0.28	0.24	0.10
	Fakten lernen	69	0.0	0.97	0.65	0.25
	Diagramme und Tabellen	69	0.0	0.78	0.50	0.23
	<b>GESAMT</b>	<b>69</b>	<b>0.7</b>	<b>2.17</b>	<b>1.44</b>	<b>0.34</b>
	Ital.	Textverständnis	42	0.0	0.25	0.18
Diagramme und Tabellen		42	0.0	0.19	0.10	0.10
<b>GESAMT</b>		<b>42</b>	<b>0.0</b>	<b>0.44</b>	<b>0.29</b>	<b>0.15</b>

Tabelle 25: Mittelwerte und Standardabweichungen der Korrekturwerte für die Personen der französisch- und italienischsprachigen Gruppen

In Abbildung 33 und Abbildung 34 sind die Verteilungen der (ungerundeten) Bonuspunkte für die vergangenen fünf 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.

Insgesamt darf nicht vergessen werden, dass es sich, gemessen an der Gesamtpunktzahl des Tests, nur um geringe Korrekturen handelt. Nie wurden mehr als (gerundet) 3 Punkte vergeben. Den französischsprachigen Teilnehmern werden durchschnittlich 2 Punkte gutgeschrieben, bei den italienischsprachigen Teilnehmern sind es üblicherweise rund 1 bis 2 Punkte.

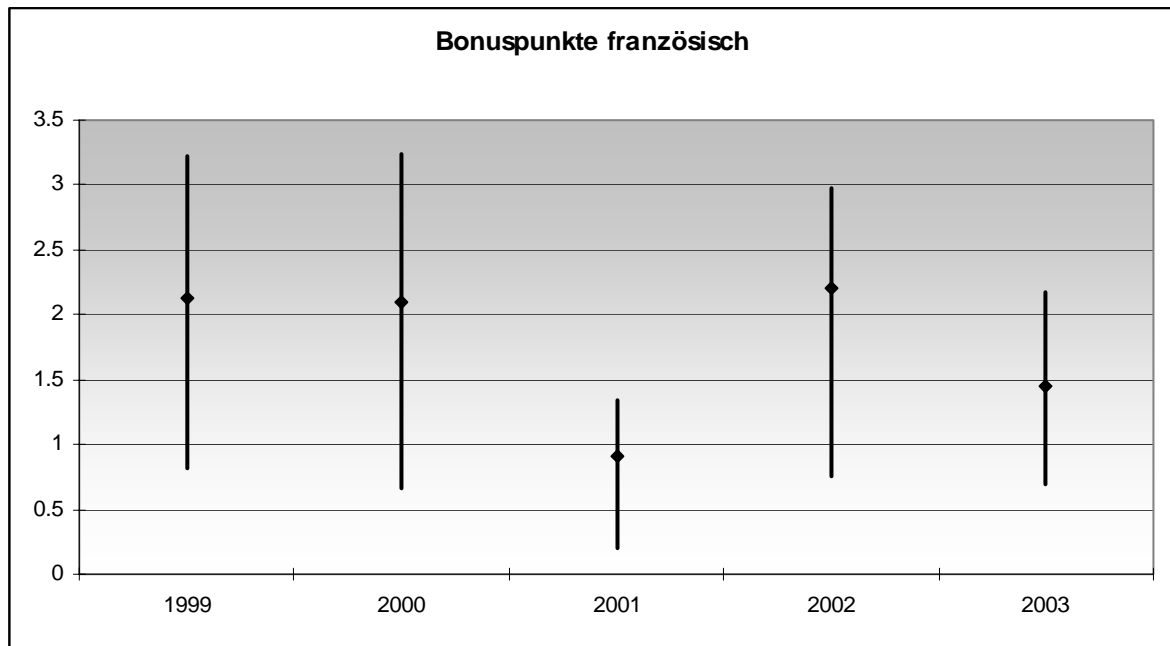


Abbildung 33: Auswirkungen der Korrektur über die Jahre für die französischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

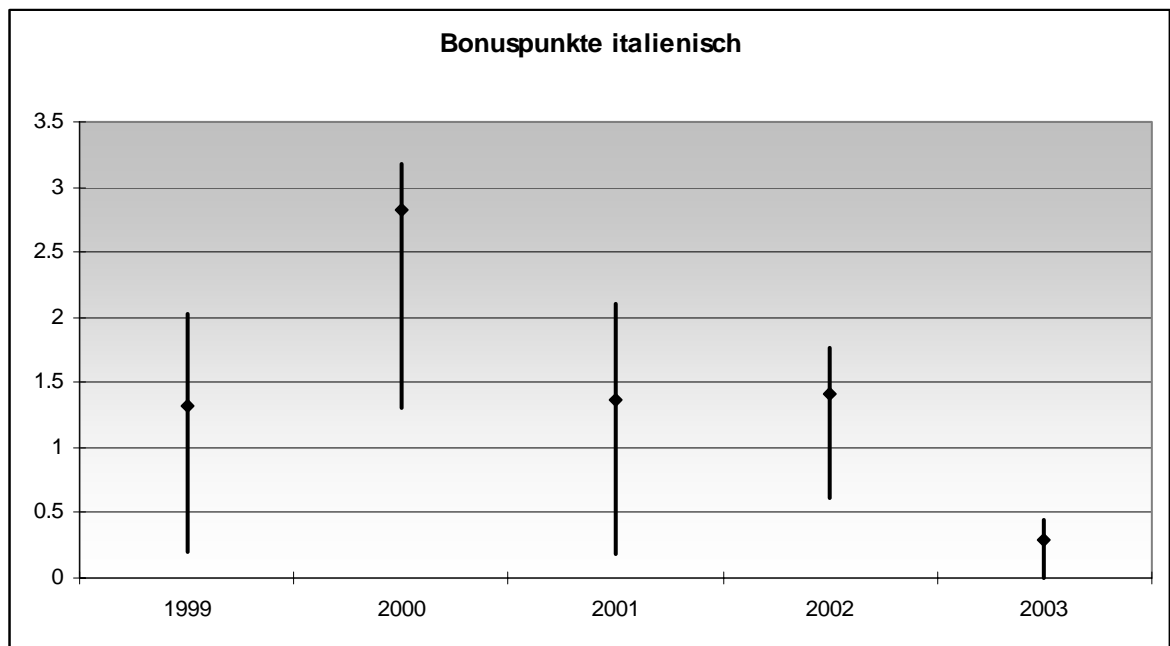


Abbildung 34: Auswirkungen der Korrektur über die Jahre für die italienischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

### 7.4.2.9 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. Diese treten nun erstmals nicht auf. Vergleichsweise gute Ergebnisse der französisch- und italienischsprachigen Teilnehmer zusammen mit den hohen Anmeldezahlen (und damit insgesamt höhere Wartelisten-Quoten) führten zu einer Angleichung der Zulassungsanteile der Sprachgruppen.

Disziplin	Sprache	Warteliste	zugelassen	Total
Humanmedizin	deutsch	195 24.2%	610 75.8%	805 100.0%
	französisch	11 33.3%	22 66.7%	33 100.0%
	italienisch	12 36.4%	21 63.6%	33 100.0%
	<b>Total</b>	<b>218</b> <b>25.0%</b>	<b>653</b> <b>75.0%</b>	<b>871</b> <b>100.0%</b>
Veterinärmedizin	deutsch	36 20.9%	136 79.1%	172 100.0%
	französisch	12 33.3%	24 66.7%	36 100.0%
	italienisch	3 33.3%	6 66.7%	9 100.0%
	<b>Total</b>	<b>51</b> <b>23.5%</b>	<b>166</b> <b>76.5%</b>	<b>217</b> <b>100.0%</b>

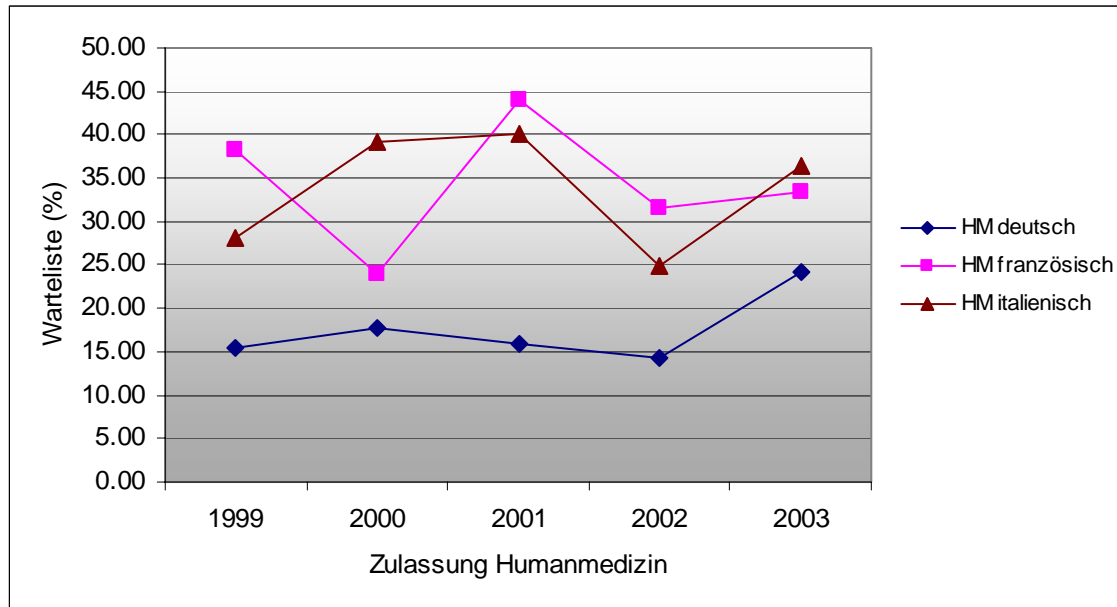
Tabelle 26: Zulassung nach Sprache des Tests

Die Signifikanzprüfung für die Zulassungsunterschiede kann Tabelle 27 entnommen werden. Es sind erstmals keine Zulassungsunterschiede nachzuweisen.

Disziplin		Wert	df	Asymp. Sig. (2seitig)
Humanmedizin	Pearson Chiquadrat	3.750	2	.153
	Likelihood Ratio	3.499	2	.174
Veterinärmedizin	Pearson Chiquadrat	3.052	2	.217
	Likelihood Ratio	2.880	2	.237

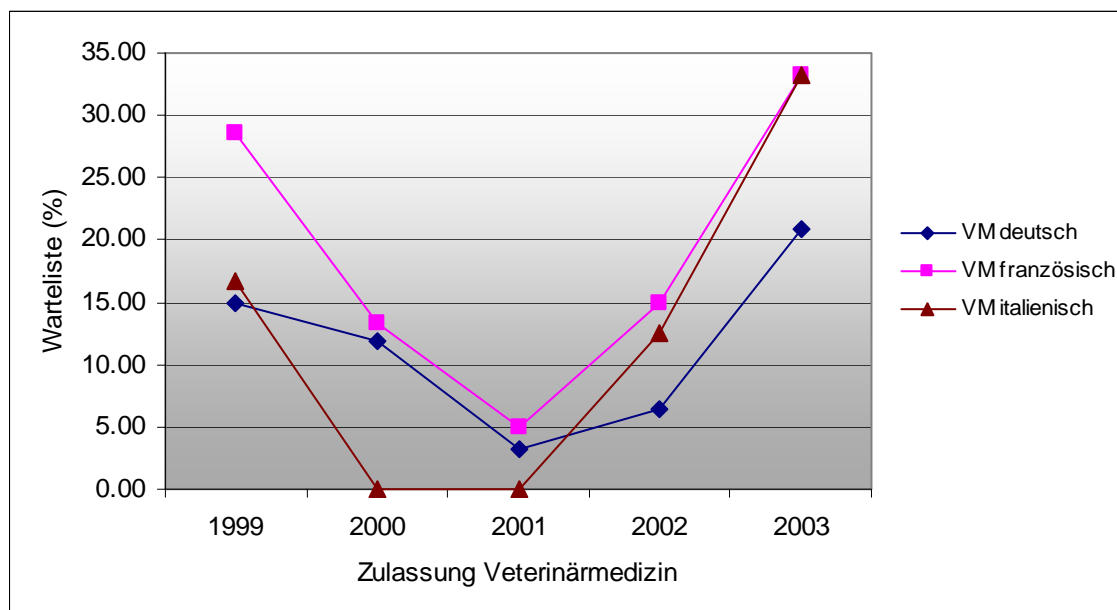
Tabelle 27: Prüfung auf Verteilungsunterschiede für Sprachgruppen getrennt nach Disziplinen: kein signifikanter Unterschied für Humanmedizin und Veterinärmedizin

In den folgenden Abbildungen sind die Wartelisten-Quoten für Humanmedizin und Veterinärmedizin dargestellt. Hohe Werte bedeuten eine geringere Zulassungsquote.



**Abbildung 35:** Zulassungsquoten Humanmedizin nach Sprache 1999 bis 2003

In der Veterinärmedizin wurde die Zulassungsquote aufgrund der hohen Anmeldezahlen weiter verringert – in allen Sprachversionen stieg dort der Anteil der Personen auf der Warteliste (Abbildung 36).



**Abbildung 36:** Zulassungsquoten Veterinärmedizin nach Sprache 1999 bis 2003

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.

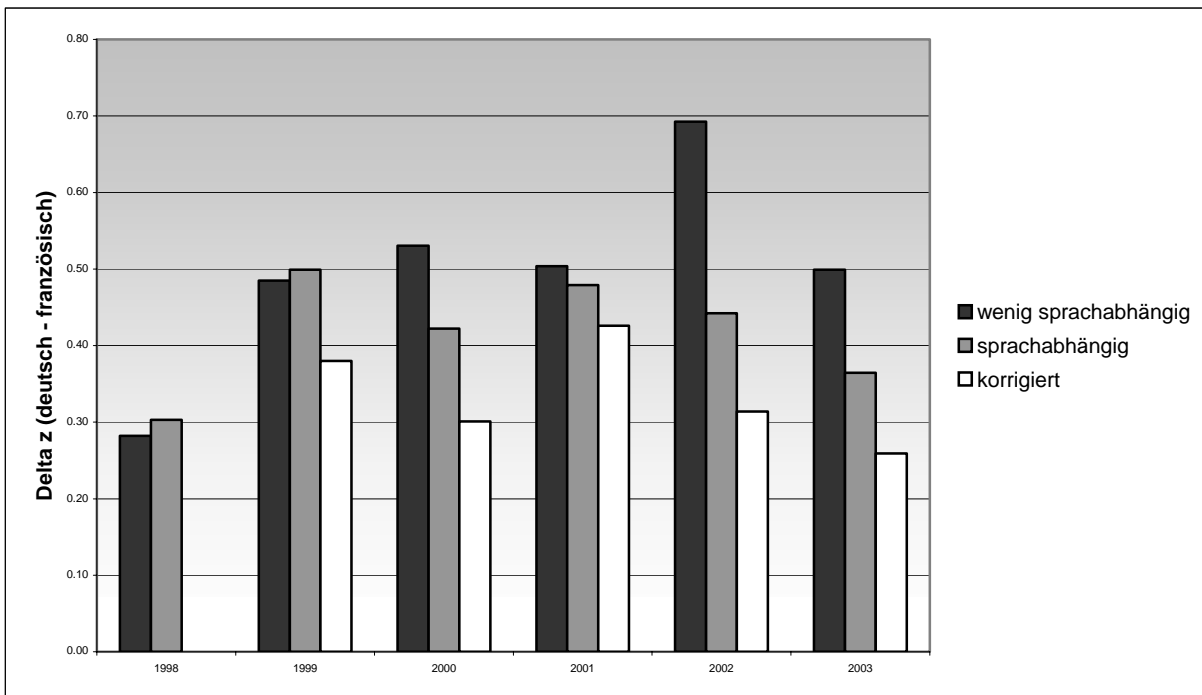
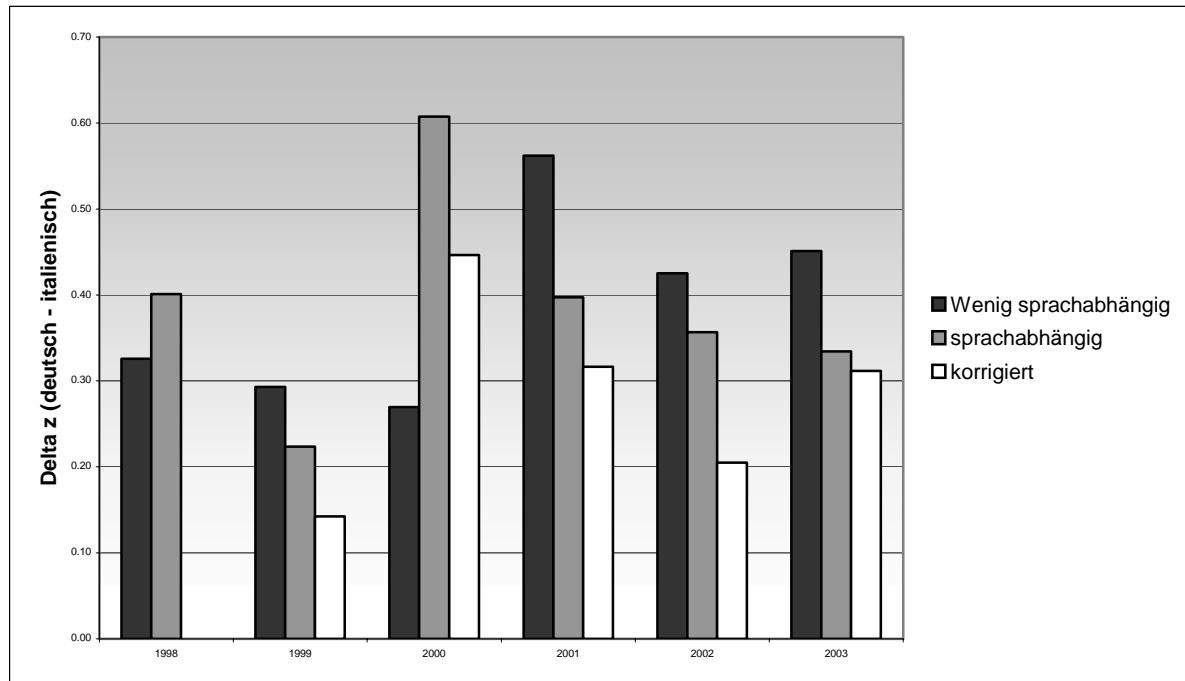


Abbildung 37: Abweichung (Delta  $z$ ) für den Testwert zwischen deutscher und französischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.



**Abbildung 38:** Abweichung (Delta z) für den Testwert zwischen deutscher und italienischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

1998 wurde für beide Sprachversionen kein aufgabenspezifisches Korrekturverfahren verwendet. An dessen Stelle wurde eine separate Normierung des Testwerts vorgenommen, wodurch ebenfalls Benachteiligung der Sprachminderheiten vermieden wird, allerdings ohne Differenzierung zwischen testbedingten und „realen“ Differenzen in den Testleistungen.

### 7.4.3 Leistungen in identischen Testaufgaben 1998 – 2003

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	2002	886	16.21	3.44	.00
	2003	977	16.95	4.11	
französisch	2002	59	14.64	3.71	.10
	2003	69	15.72	3.58	
italienisch	2002	28	15.36	3.77	.98
	2003	42	15.33	4.49	

Tabelle 28: Verteilungskennwerte identischer Einstreuaufgaben 2002 und 2003

Abbildung 39 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 (Effektgrösse = 0.20). In der französischen Version ist der Unterschied aufgrund der geringeren Fallzahl statistisch nicht signifikant, eine Verbesserung ist aber auch hier tendenziell zu beobachten (Effektgrösse = 0.30). Die Leistungen der italienischsprachigen Gruppe unterscheiden sich nicht von denjenigen des Vorjahres (Effektgrösse = 0.01). Die Effektgrößen sind durchwegs als „gering“ einzustufen (Cohen, 1988).

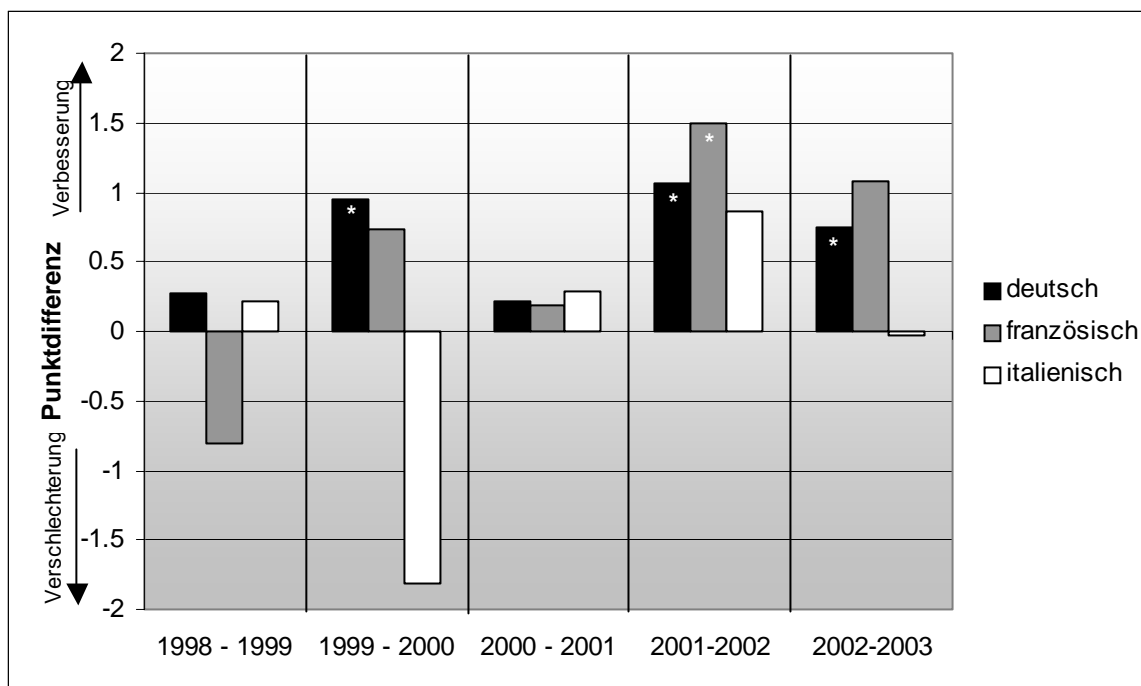


Abbildung 39: Verlauf der Leistungen in identischen Testaufgaben 1998 bis 2003 (Differenzverlauf); \*= signifikante Differenz

## 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, insbesondere auch hinsichtlich der hohen Punktzahlen im Untertest „Figuren lernen“.

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testwert	<i>Zwischen Gruppen</i>	4152.11	12	346.01	3.54	<b>0.00</b>
	<i>Innerhalb Gruppen</i>	94135.50	964	97.65		
	<i>Total</i>	98287.61	976			
Muster zuordnen	<i>Zwischen Gruppen</i>	178.94	12	14.91	2.08	<b>0.02</b>
	<i>Innerhalb Gruppen</i>	6912.78	964	7.17		
	<i>Total</i>	7091.72	976			
Med.-naturwiss. Grundverständnis	<i>Zwischen Gruppen</i>	213.23	12	17.77	1.34	0.19
	<i>Innerhalb Gruppen</i>	12827.02	964	13.31		
	<i>Total</i>	13040.25	976			
Schlauchfiguren	<i>Zwischen Gruppen</i>	177.81	12	14.82	1.02	0.43
	<i>Innerhalb Gruppen</i>	14067.51	964	14.59		
	<i>Total</i>	14245.33	976			
Quant. und formale Probleme	<i>Zwischen Gruppen</i>	243.68	12	20.31	1.60	0.09
	<i>Innerhalb Gruppen</i>	12272.57	964	12.73		
	<i>Total</i>	12516.25	976			
Textverständnis	<i>Zwischen Gruppen</i>	407.39	12	33.95	2.24	<b>0.01</b>
	<i>Innerhalb Gruppen</i>	14635.08	964	15.18		
	<i>Total</i>	15042.47	976			
Figuren lernen	<i>Zwischen Gruppen</i>	485.31	12	40.44	2.65	<b>0.00</b>
	<i>Innerhalb Gruppen</i>	14701.08	964	15.25		
	<i>Total</i>	15186.38	976			
Fakten lernen	<i>Zwischen Gruppen</i>	362.50	12	30.21	2.02	<b>0.02</b>
	<i>Innerhalb Gruppen</i>	14435.11	964	14.97		
	<i>Total</i>	14797.61	976			
Diagramme und Tabellen	<i>Zwischen Gruppen</i>	269.56	12	22.46	1.95	<b>0.03</b>
	<i>Innerhalb Gruppen</i>	11110.89	964	11.53		
	<i>Total</i>	11380.45	976			
Konzentr. sorgf. Arbeiten	<i>Zwischen Gruppen</i>	533.01	12	44.42	3.42	<b>0.00</b>
	<i>Innerhalb Gruppen</i>	12504.38	964	12.97		
	<i>Total</i>	13037.39	976			

Tabelle 29: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den Testlokalen

Die 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 40). Erstmals ist in diesem Jahr ein wieder geringer Anstieg des Anteils an Kandidaten zu verzeichnen. Er stieg von 34,8% (2002) auf 36,3% (2003). In der Veterinärmedizin bleibt der Frauenanteil weiterhin praktisch stabil auf 81,6%.

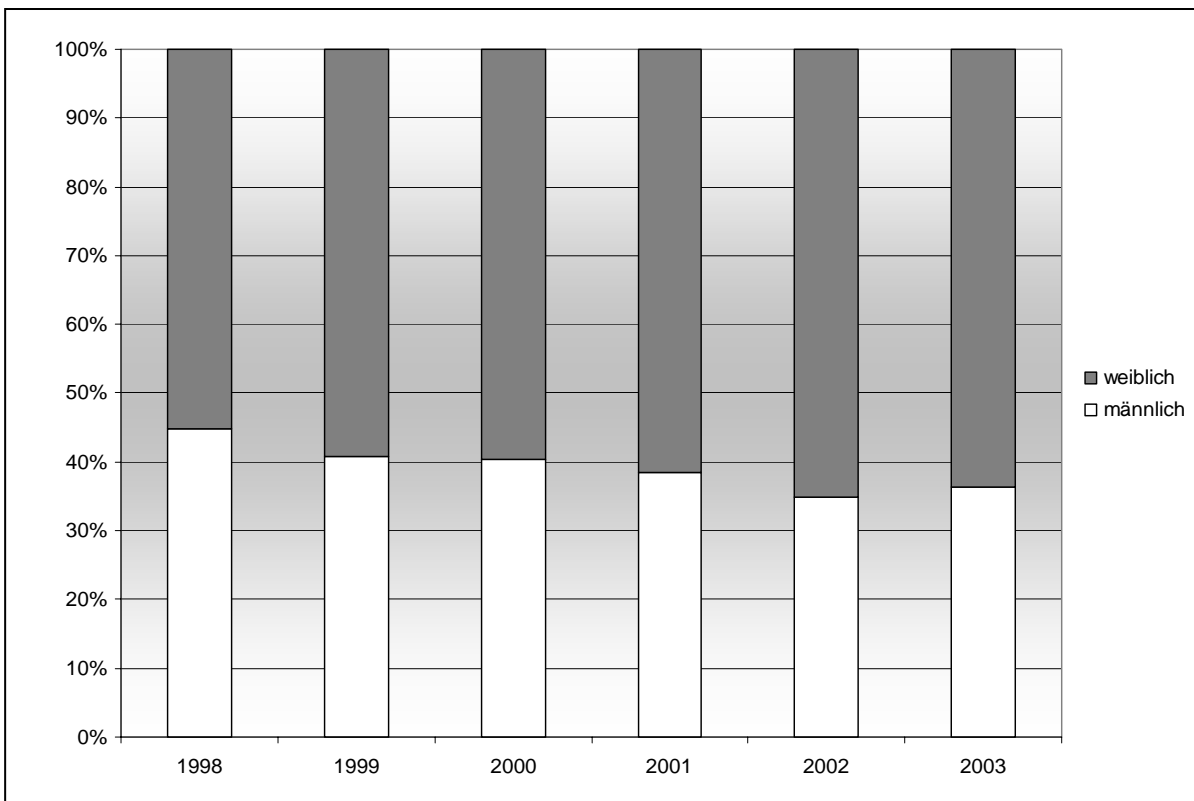


Abbildung 40: Bewerberkohorten (Humanmedizin) 1998 bis 2003 nach Geschlecht

Was heisst nun Chancengleichheit? Es kann über 2 Jahre nachgewiesen werden, dass sich die mittleren Testwerte für Personen, welche die erste Vorprüfung bestehen, nicht zwischen Männern und Frauen unterscheiden. Wenn der Testwert Eignung erfasst, so haben gleich Geeignete tatsächlich die gleichen Chancen.

Die Erfolgswahrscheinlichkeit, die erste Vorprüfung zu bestehen, unterscheidet sich allerdings um 7% zugunsten der Männer. Wegen der Mittelwertgleichheit der Erfolgreichen ist es sehr wahrscheinlich, dass dieser Unterschied in einer unterschiedlichen Repräsentativität der Bewerbergruppen für die beiden Geschlechter begründet liegt.



Chancengleichheit bedeutet also nicht die Nivellierung von **tatsächlich vorhandenen** Unterschieden. Der Test muss die vorhandenen Realitäten unverzerrt abbilden.

Auch in diesem Jahr liegen die Testwerte der deutschsprachigen männlichen Teilnehmer leicht über denjenigen der Kandidatinnen (Tabelle 30). Die Differenz bleibt wie auch schon in den Vorjahren gering (Tabelle 31). Die Tendenz zum Unterschied bleibt also bestehen, allerdings bewegt sich die Differenz in unbedeutenden Grössenordnungen (Effektgrösse  $\ll 0.2$ ). Unter Berücksichtigung der Erfolgswahrscheinlichkeiten für die Vorprüfungen kann eine Benachteiligung der Kandidatinnen durch den Eignungstest nach wie vor ausgeschlossen werden.

		Humanmedizin			Veterinärmedizin		
		deutsch	französisch	italienisch	deutsch	französisch	italienisch
männlich	Mittelwert	<b>101.92</b>	<b>97.00</b>	<b>99.29</b>	<b>99.15</b>	<b>96.15</b>	<b>96.00</b>
	Stand.-abw.	9.81	6.88	10.13	9.21	7.77	
	N	290	12	14	26	13	1
weiblich	Mittelwert	<b>100.49</b>	<b>95.76</b>	<b>91.37</b>	<b>98.83</b>	<b>91.39</b>	<b>93.00</b>
	Stand.-abw.	10.21	6.81	7.79	9.74	6.76	6.76
	N	515	21	19	146	23	8

Tabelle 30: Statistiken für den Testwert nach Geschlechtern und Disziplin: Werte für Gruppen unter 30 Personen (kursiv) nicht ausreichend für statistische Vergleiche

Sprache	Disziplin	Levene's Test Gleichheit der Varianzen			t-Test Gleichheit des Mittelwerts		
			F	Sig.	t	df	Sig.
deutsch	Humanmedizin	Varianzen homogen	1.400	.23	1.937	803	.053
	Veterinärmedizin	Varianzen homogen	.095	.75	.158	170	.875
franzö-sisch	Humanmedizin	Varianzen homogen	.224	.63	.501	31	.620
	Veterinärmedizin	Varianzen homogen	.362	.55	1.925	34	.063
italienisch	Humanmedizin	Varianzen homogen	1.311	.26	2.540	31	<b>.016</b>

Tabelle 31: Prüfung Varianzhomogenität und Mittelwert (t-Test) für Testwert zwischen Geschlechtern. Italienische Sprachgruppe Veterinärmedizin wg. Fallzahl nicht berechnet

Abbildung 41 zeigt den Vergleich zwischen männlichen Bewerbern und Kandidatinnen, aufgeteilt nach Disziplin, Abbildung 42 den Boxplot für die Testwerte nach Sprache und Geschlecht geschichtet.

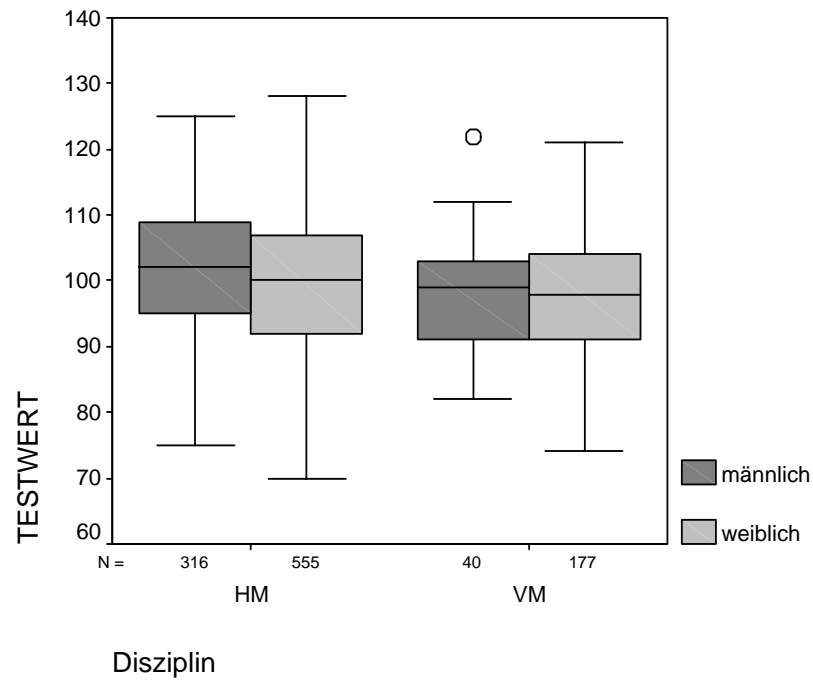


Abbildung 41: Boxplot für den Testwert nach Geschlecht und Disziplin (Humanmedizin und Veterinärmedizin)

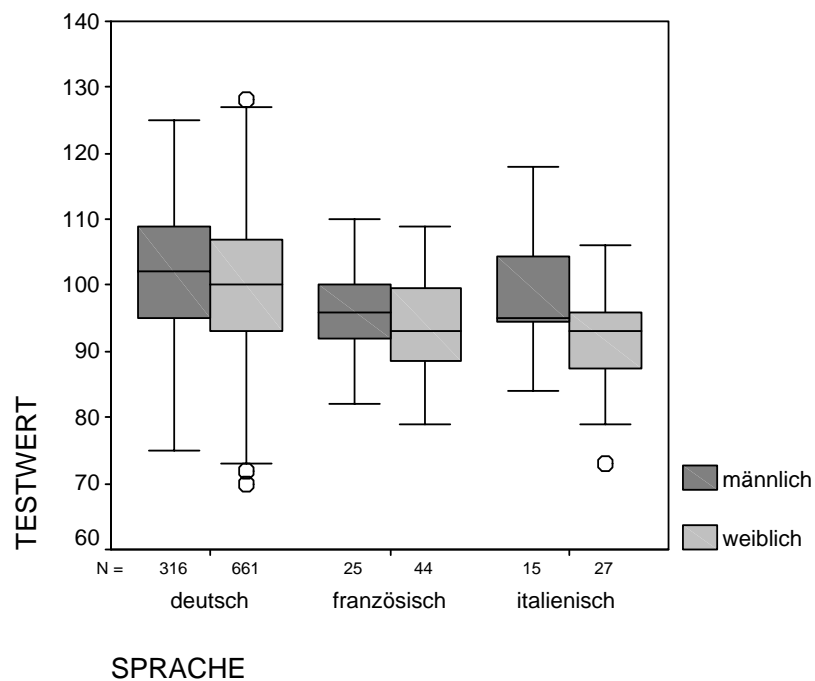


Abbildung 42: Boxplots für die Testwerte nach Geschlecht, getrennt für die drei Sprachgruppen

Wiederum zeigen sich Differenzen für die einzelnen Untertests zwischen den Geschlechtern. Diesmal sind sämtliche Differenzen signifikant – erstmals auch im Untertest „Muster zuordnen“. Gleichzeitig ist dies der einzige Untertest, bei welchem sich die „Richtung“ der Unterschiede geändert hat. In den ersten Durchführungen erzielten die Kandidatinnen höhere Ergebnisse, in den letzten beiden Jahrgängen waren die männlichen Teilnehmer besser.

Frauen erzielen tendenziell bessere Werte in den Untertests „Figuren lernen“, „Fakten lernen“ und „Konzentriertes und sorgfältiges Arbeiten“. Die männlichen Testteilnehmer erreichen höhere Punktzahlen in den Untertests „Schlauchfiguren“, „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“ sowie „Diagramme und Tabellen“. Sämtliche Effektgrößen bewegen sich im unteren (0.2) bis maximal mittleren (0.5) Bereich.

Untertest	Geschlecht	m	s	Levene's Test Varianz			EG	t-test Mittelwert		
					F	Sig.		t	df	Sig.
Muster zuordnen	männlich	11.08	2.67	<i>homogen</i>	0.027	0.87	0.20	3.045	1086.00	<b>0.00</b>
	weiblich	10.55	2.68							
Med.-naturwiss. Grundverst.	männlich	10.95	3.53	<i>homogen</i>	0.406	0.52	0.24	3.644	1086.00	<b>0.00</b>
	weiblich	10.09	3.70							
Schlauchfiguren	männlich	13.41	3.73	<i>homogen</i>	0.414	0.52	0.35	5.465	1086.00	<b>0.00</b>
	weiblich	12.09	3.76							
Quantitative und formale Probleme	männlich	11.55	3.68	<i>Nicht homogen</i>	3.859	0.05	0.48	7.331	657.99	<b>0.00</b>
	weiblich	9.85	3.41							
Textverständnis	männlich	10.57	3.76	<i>homogen</i>	0.058	0.81	0.19	2.862	1086.00	<b>0.00</b>
	weiblich	9.85	3.92							
Figuren lernen	männlich	12.77	3.98	<i>homogen</i>	0.326	0.57	0.25	-3.927	1086.00	<b>0.00</b>
	weiblich	13.76	3.85							
Fakten lernen	männlich	10.30	3.64	<i>Nicht homogen</i>	4.005	0.05	0.38	-5.928	741.23	<b>0.00</b>
	weiblich	11.72	3.86							
Diagramme und Tabellen	männlich	12.58	3.36	<i>homogen</i>	0.022	0.88	0.49	7.549	1086.00	<b>0.00</b>
	weiblich	10.95	3.32							
Konzent. u. sorgfältiges Arbeiten	männlich	12.97	3.53	<i>homogen</i>	2.655	0.10	0.18	-2.764	1086.00	<b>0.01</b>
	weiblich	13.62	3.66							

**Tabelle 32:** Mittelwerte (m) und Standardabweichungen (s) geschlechtsspezifisch und Ergebnisse der Prüfungen auf Varianzhomogenität und Mittelwertunterschiede mit Effektgrösse (EG) (Näherung wenn keine Varianzhomogenität)

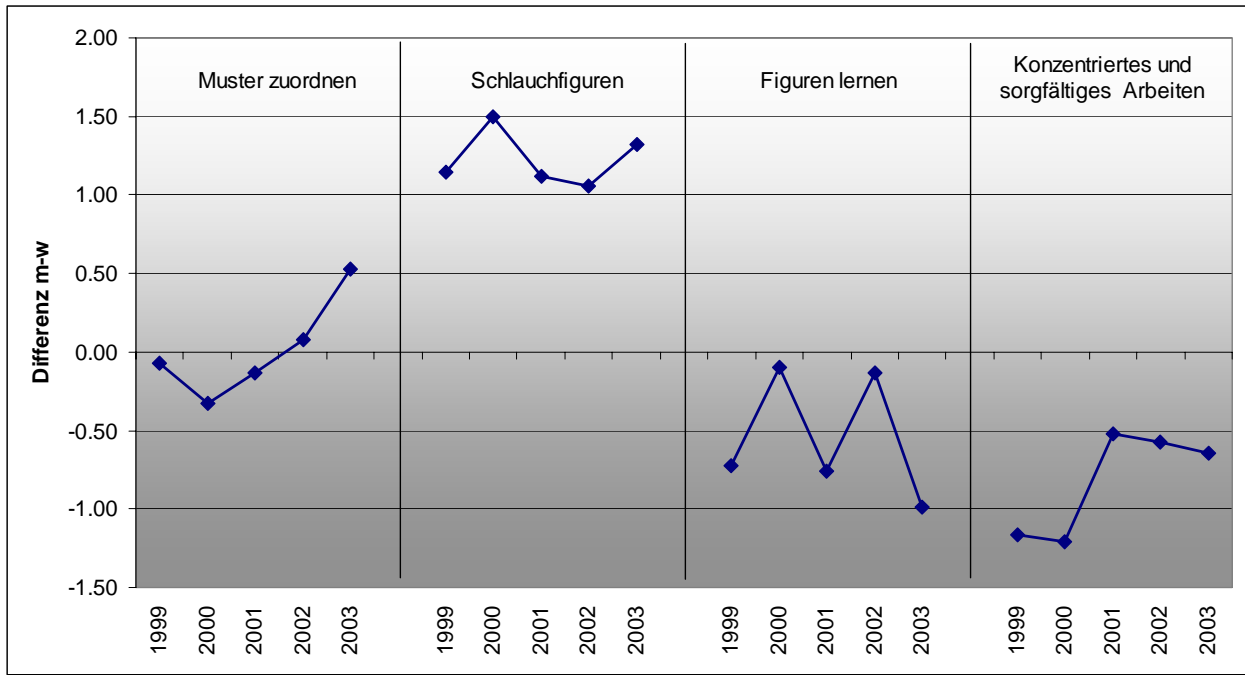


Abbildung 43: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (wenig sprachabhängige Untertests)

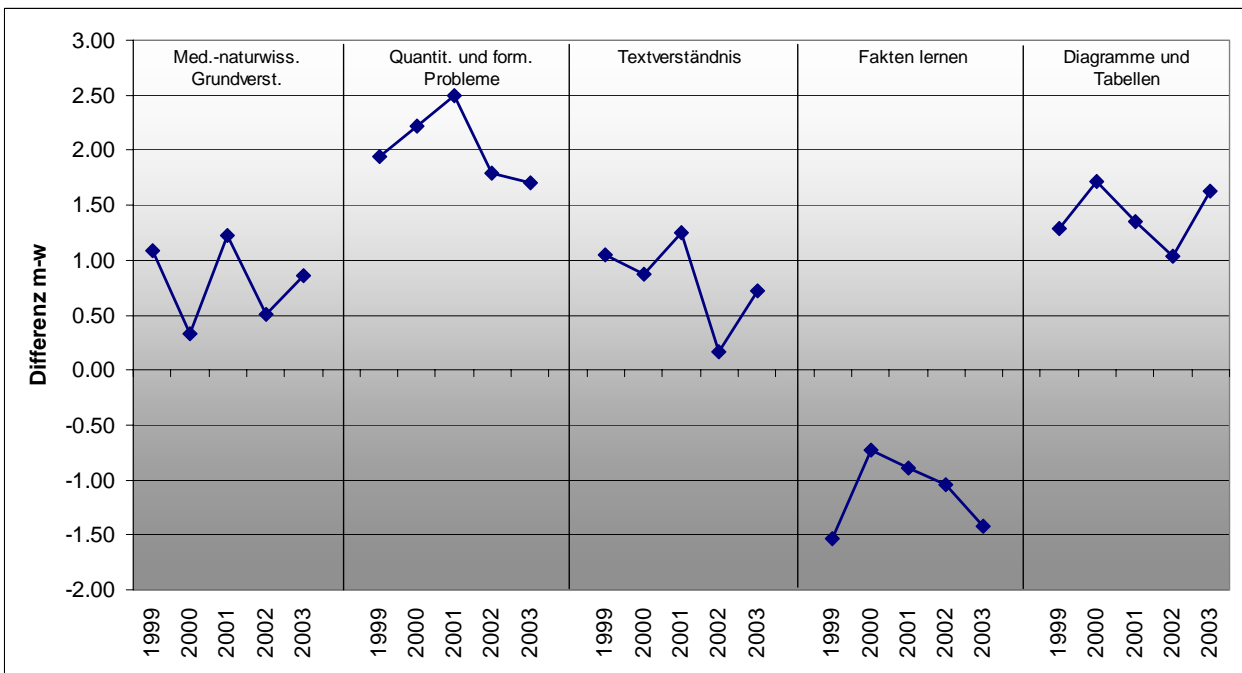


Abbildung 44: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (sprachabhängige Untertests)

Bezüglich der Zulassung sind diesmal in der Disziplin „Humanmedizin“ signifikante Unterschiede zwischen den Geschlechtern festzustellen (Tabelle 34). Bei den Männern beträgt sie 80.4%, bei den Frauen 71.9% (Tabelle 33). Dies kann mit der grösseren Wartelisten-Quote und gleichzeitig höheren Varianz der durch die Teilnehmerinnen erzielten Testwerte erklärt werden. Aufgrund der grösseren Streuung sind die Kandidatinnen sowohl im unteren wie oberen Leistungsbereich stärker vertreten als ihre männlichen Kollegen.

Disziplin	Geschlecht	Warteliste	zugelassen	Total
Humanmedizin	männlich	62 19.6%	<b>254</b> <b>80.4%</b>	316 100.0%
	weiblich	156 28.1%	<b>399</b> <b>71.9%</b>	555 100.0%
	Total	218 25.0%	653 75.0%	871 100.0%
Veterinärmedizin	männlich	8 20.0%	32 80.0%	40 100.0%
	weiblich	43 24.3%	134 75.7%	177 100.0%
	Total	51 23.5%	166 76.5%	217 100.0%

Tabelle 33: Zulassungsquoten geschlechtsspezifisch

Disziplin		Wert	df	Asymp. Sig. (2-seitig)
Humanmedizin	Pearson Chiquadrat	7.731	1	<b>.01</b>
	Likelihood Ratio	7.930	1	<b>.01</b>
Veterinärmedizin	Pearson Chiquadrat	.335	1	.563
	Likelihood Ratio	.344	1	.557

Tabelle 34: Prüfung Unterschiede der Zulassung – geschlechtsspezifisch: signifikante Unterschiede bei Humanmedizin

## 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 2003 innerhalb der Deutschschweiz

Eine der wichtigsten Fragen ist jeweils, ob sich die Leistungen für die Kantone nach unterschiedlichen Maturitätsquoten unterscheiden. Diesbezüglich konnten in den letzten Jahren keine Unterschiede verzeichnet werden. Basierend auf den Quoten für 2001<sup>1</sup> wurden die Kantone wie immer in die drei Gruppen geteilt, die hohe, mittlere und niedrige Maturitätsquoten repräsentieren (Tabelle 35). Gegenüber früheren Einteilungen sind aufgrund der neueren Kennziffern dieses Jahr einige Verschiebungen zu registrieren.

hoch			mittel			niedrig		
Kanton	Maturitäts- quote	Anzahl	Kanton	Maturitäts- quote	Anzahl	Kanton	Maturitäts- quote	Anzahl
BS	25.9	43	AG	15.7	67	SG	13.7	61
SH	19.4	12	SZ	15.4	13	LU	13.3	59
ZH	18.8	276	GR	15.4	59	BE	13.0	167
BL	18.8	52	AR	14.8	8	UR	11.7	4
			NW	14.8	11	TG	11.6	33
			OW	14.7	2	AI	11.1	0
			GL	14.5	7			
			ZG	14.3	18			
			SO	14.2	45			
Summe		383	Summe		230	Summe		324

Tabelle 35: Gruppierung der Kantone nach der Maturitätsquote (Basis 2001)

Zwischen den nach Maturitätsquoten kategorisierten Kantonen bestehen keine Unterschiede bezüglich des Testwerts (Tabelle 36). Wo Unterschiede auftreten (Untertests „Fakten lernen“ und „Konzentriertes und sorgfältiges Arbeiten“) sind die Leistungen in den Kantonen mit hoher Maturitätsquote sogar höher als in den anderen. Es gibt nach wie vor keinen Grund dafür, in der unterschiedlichen Maturitätsquote eine generelle Ursache zu sehen, aus der sich Schlussfolgerungen für einzelne Kantone bezüglich einer Auswahl für die Mittelschule ergeben.

<sup>1</sup> Zahlen aus: Bundesamt für Statistik (2002). Maturitäten 2001. *Maturités 2001*. Neuchâtel 2002

	Maturitäts- quote	Mittel- wert	Stand.- abw.		Quadrat- summe	df	MQ	F	Sig.
Testwert	hoch	101.58	10.17	Zwischen Gruppen	395.213	2	197.606	1.986	.13
	mittel	100.13	9.78						
	niedrig	100.37	9.88						
Muster zuordnen	hoch	11.05	2.73	Zwischen Gruppen	20.926	2	10.463	1.455	.23
	mittel	10.70	2.46						
	niedrig	10.80	2.77						
Medizin.-natur- wiss. Grundver- ständnis	hoch	10.63	3.53	Zwischen Gruppen	5.289	2	2.645	.199	.82
	mittel	10.43	3.70						
	niedrig	10.56	3.74						
Schlauchfigure n	hoch	12.97	3.94	Zwischen Gruppen	29.927	2	14.964	1.031	.35
	mittel	12.58	3.70						
	niedrig	12.62	3.72						
Quantitat. u. formale Probleme	hoch	10.68	3.53	Zwischen Gruppen	35.027	2	17.514	1.384	.25
	mittel	10.63	3.54						
	niedrig	10.26	3.60						
Textverständnis	hoch	10.39	3.87	Zwischen Gruppen	7.412	2	3.706	.241	.78
	mittel	10.17	4.04						
	niedrig	10.24	3.91						
Figuren lernen	hoch	13.75	4.20	Zwischen Gruppen	4.811	2	2.405	.155	.85
	mittel	13.67	3.69						
	niedrig	13.59	3.79						
Fakten lernen	hoch	11.72	3.95	Zwischen Gruppen	119.052	2	59.526	3.892	<b>.02</b>
	mittel	10.81	3.84						
	niedrig	11.38	3.91						
Diagramme und Tabellen	hoch	11.69	3.40	Zwischen Gruppen	3.256	2	1.628	.141	.86
	mittel	11.74	3.29						
	niedrig	11.60	3.46						
Konzentr. u. sorgf. Arb.	hoch	14.15	3.69	Zwischen Gruppen	134.766	2	67.383	5.104	<b>.00</b>
	mittel	13.32	3.82						
	niedrig	13.42	3.43						

**Tabelle 36:** Mittelwerte und Standardabweichungen für drei Gruppen der Kantone nach der Maturitätsquote

Für Vergleiche zwischen den Kantonen werden wiederum wie in den Vorjahren die gleichen, zahlenmässig stärksten Kantone berücksichtigt (Tabelle 37). Die Abbildung 45 zeigt den Verlauf der Testwerte für die Kantone über die fünf bisherigen Testdurchführungen.

Die Rangreihe der Kantone ist über die einzelnen Jahre bis auf einige Ausnahmen relativ stabil (Abbildung 45). Ein besonders gutes Ergebnis wurde dieses Jahr für den Kanton Luzern registriert. Dort wurde dieses Jahr der beste Testwert-Durchschnitt erreicht, und gleichzeitig der zweitbeste Wert seit Einführung des Eignungstests (St. Gallen erzielte den höchsten Wert bisher im Jahr 2000). Basel-Stadt konnte die Leistungen des letzten Jahres ebenfalls wieder deutlich steigern, und erzielte eine neue „Bestleistung“ (erstmalig über 98

Punkte). Der Kanton Aargau fand nach einem eher schwachen Jahr zu den „normalen“ Werten zurück. Alle übrigen Kantone weisen im Vergleich zum Vorjahr Punkterückgänge auf.

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	<b>101.76</b>	10.66	11.34	12.34	11.36	11.15	13.88	10.90	12.03	13.70
	S	<b>9.33</b>	2.35	3.81	3.63	3.60	4.32	3.44	3.89	3.23	3.54
BE	M	<b>98.95</b>	10.23	10.59	12.47	9.86	10.11	12.98	11.12	11.28	12.88
	S	<b>10.19</b>	2.63	3.78	3.66	3.56	3.96	3.84	4.09	3.59	3.28
BL	M	<b>100.15</b>	10.98	10.75	12.67	10.69	9.83	13.27	10.69	11.46	13.83
	S	<b>11.28</b>	3.08	3.30	4.28	3.66	4.05	4.59	4.15	2.95	3.88
BS	M	<b>98.19</b>	10.33	10.33	12.56	9.44	9.30	13.40	10.67	11.28	12.63
	S	<b>10.21</b>	2.22	3.92	3.72	3.25	4.41	3.79	3.74	3.98	3.08
LU	M	<b>103.51</b>	11.92	11.51	12.46	11.44	11.36	14.10	11.54	12.54	14.03
	S	<b>9.49</b>	2.82	3.16	3.79	3.70	3.94	3.27	4.01	3.35	3.03
SG	M	<b>100.54</b>	11.10	9.90	13.11	10.10	9.48	14.02	11.64	11.59	13.90
	S	<b>9.20</b>	3.06	3.72	3.83	3.20	3.43	4.20	3.66	3.21	3.96
ZH	M	<b>102.42</b>	11.18	10.67	13.15	10.86	10.64	13.94	12.06	11.78	14.47
	S	<b>9.89</b>	2.76	3.55	3.93	3.51	3.74	4.22	3.89	3.38	3.66

Tabelle 37: Vergleich der Kantone für Testwert und die einzelnen Untertests

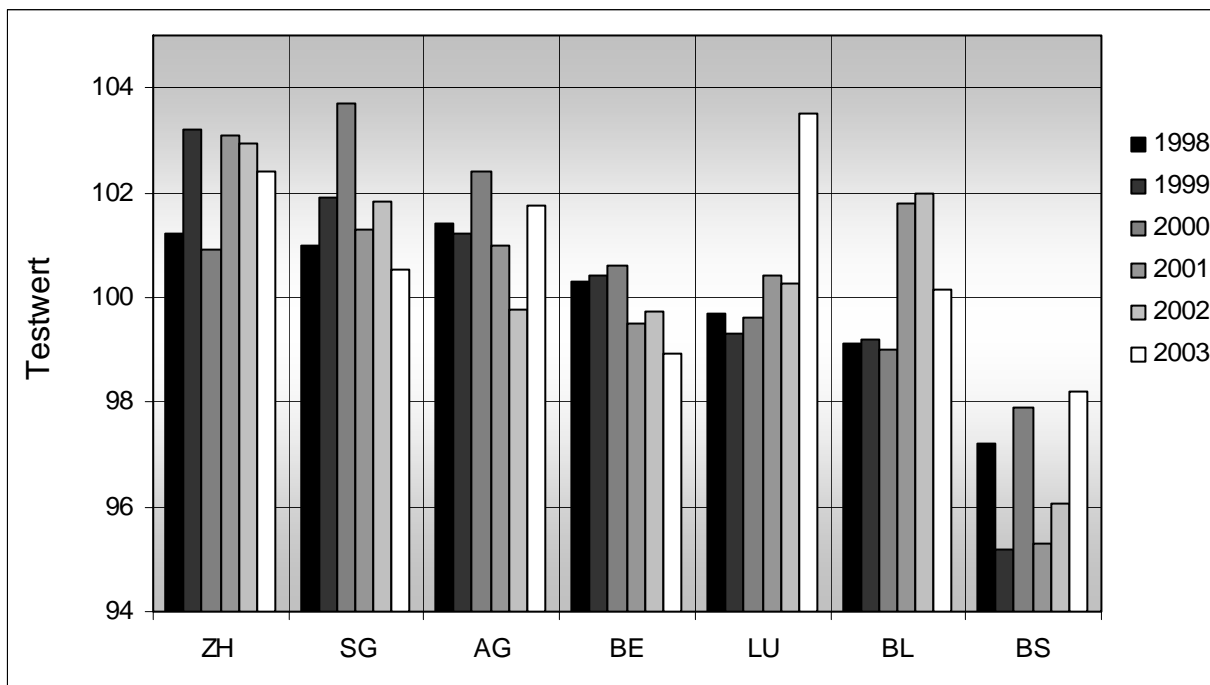


Abbildung 45: Testwert für Kantone im Vergleich



## 7.8 Vergleiche für Altersgruppen

Für die Überprüfung der Alterseffekte im Jahr 2003 wurden die Gruppen auf der Basis der Tabelle 38 gebildet. Die Altersgruppen wurden um 1 Jahr gegenüber 2002 verschoben. Die älteste Gruppe wurde nahe dem Median des Maturitätsalters geteilt.

Maturitätsjahr	Geburtsjahr			Total
	1959-1979	1980-1982	1983-1987	
1978	1			1
1983	1			1
1984	1			1
1985	1			1
1986	1			1
1987	1			1
1989	1			1
1991	2			2
1992	2			2
1993	4			4
1994	1			1
1995	4			4
1996	3			3
1997	6			6
1998	9			9
1999	9	4		13
2000	3	15		18
2001	4	40	5	49
2002	8	130	224	362
2003	25	83	500	608
	<b>87</b>	<b>272</b>	<b>729</b>	<b>1088</b>

Tabelle 38: Gruppenbildung für das Jahr der Maturitätsprüfung bezogen auf die Geburtsjahre

Erneut zeigen die älteren Jahrgänge mit später Maturität niedrigere Testwerte als alle anderen Gruppen (Abbildung 46).

	Geburtsjahr und Maturität	Mittelwert	Standardabw.	Quadratsumme	df	MQ	F	Sig.
Testwert	1959 - 1979, frühe Mat.	99.57	11.53	1791.32	3	597.11	6.05	<b>0.00</b>
	1959 - 1979, späte Mat.	96.10	10.79					
	1980 - 1982	98.44	10.52					
	1983 - 1987	100.84	9.55					
Muster zuordnen	1959 - 1979, frühe Mat.	9.89	2.81	40.99	3	13.66	1.90	0.13
	1959 - 1979, späte Mat.	10.45	2.50					
	1980 - 1982	10.68	2.59					
	1983 - 1987	10.81	2.72					
Med.-naturwiss. Grundverständnis	1959 - 1979, frühe Mat.	10.83	3.68	273.05	3	91.02	6.89	<b>0.00</b>
	1959 - 1979, späte Mat.	8.85	3.48					
	1980 - 1982	9.74	3.55					
	1983 - 1987	10.66	3.67					
Schlauchfiguren	1959 - 1979, frühe Mat.	12.21	4.39	22.77	3	7.59	0.53	0.66
	1959 - 1979, späte Mat.	11.90	4.18					
	1980 - 1982	12.62	3.87					
	1983 - 1987	12.53	3.71					
Quantitative und formale Probleme	1959 - 1979, frühe Mat.	10.87	4.37	177.86	3	59.29	4.66	<b>0.00</b>
	1959 - 1979, späte Mat.	9.13	4.01					
	1980 - 1982	9.91	3.46					
	1983 - 1987	10.63	3.53					
Textverständnis	1959 - 1979, frühe Mat.	10.19	4.28	248.22	3	82.74	5.57	<b>0.00</b>
	1959 - 1979, späte Mat.	8.83	3.82					
	1980 - 1982	9.43	3.93					
	1983 - 1987	10.39	3.80					
Figuren lernen	1959 - 1979, frühe Mat.	12.87	3.85	61.20	3	20.40	1.33	0.26
	1959 - 1979, späte Mat.	12.98	4.85					
	1980 - 1982	13.17	4.10					
	1983 - 1987	13.59	3.80					
Fakten lernen	1959 - 1979, frühe Mat.	10.94	3.86	184.88	3	61.63	4.20	<b>0.01</b>
	1959 - 1979, späte Mat.	10.30	3.94					
	1980 - 1982	10.70	3.84					
	1983 - 1987	11.54	3.82					
Diagramme und Tabellen	1959 - 1979, frühe Mat.	11.74	4.19	251.70	3	83.90	7.31	<b>0.00</b>
	1959 - 1979, späte Mat.	9.83	4.16					
	1980 - 1982	10.94	3.36					
	1983 - 1987	11.76	3.29					
Konzentriertes und sorgfältiges Arbeiten	1959 - 1979, frühe Mat.	13.23	3.56	6.89	3	2.30	0.17	0.91
	1959 - 1979, späte Mat.	13.65	3.85					
	1980 - 1982	13.32	3.90					
	1983 - 1987	13.44	3.52					

**Tabelle 39:** Testwert und Punktwerte für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung; Varianzanalytische Prüfung des Unterschiedes

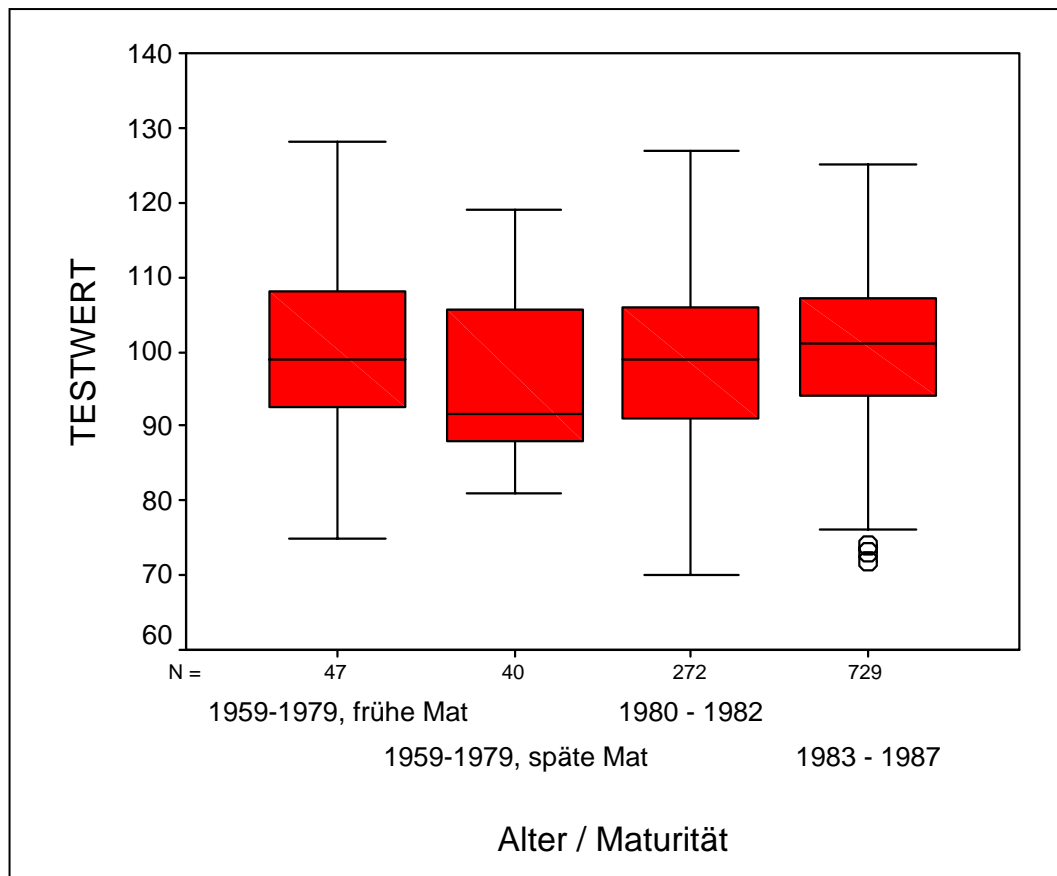


Abbildung 46: Boxplots für Testwert der Altersgruppen (Geburtsjahre) und Maturität (Mat)

Eine Betrachtung der Zulassungsquoten (Tabelle 40) zeigt einen Wert von 89% für die Gruppe der jüngsten Teilnehmer, rund 67% für die zweitjüngste, 70% für die älteste Gruppe mit früher Maturität, und 45% für die älteren Jahrgänge mit später Maturität. Damit schneidet die ältere Gruppe mit später Maturität auch hier wie in den Vorjahren deutlich schlechter ab als die anderen Kandidaten.

Geburtsjahr und Maturität	Warteliste	zugelassen	n
1959-1979, frühe Maturität	14 29.8%	33 70.2%	47
1959-1979, späte Maturität	22 55.0%	18 45.0%	40
1980 - 1982	90 33.1%	182 66.9%	272
1983 - 1987	143 11.0%	586 89.0%	729
Total	269 24.7%	819 75.3%	1088

Tabelle 40: Zulassungsquoten für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung

## 7.9 Vergleiche nach Wunschuniversitäten

Die Zulassung zu den Universitäten erfolgt aufgrund des Testwertes (Tabelle 41). Erstmals unterscheiden sich die Zulassungsquoten in der Humanmedizin signifikant (Tabelle 45), wenn auch die Effektgrösse klein ausfällt (Cramer's  $V = .150$ ). Die höchste Zulassungsquote wird unter jenen Personen erreicht, welche Zürich als ihre Wunschuniversität nennen.

Disziplinen	Wunsch-Universität	Zulassung		Total
		Warteliste	zugelassen	
Humanmedizin	Bern	62	159	221
		28.1%	71.9%	100%
	Basel	52	138	190
		27.4%	72.6%	100%
	Freiburg	29	51	80
36.3%		63.8%	100%	
Zürich	73	305	378	
	19.3%	80.7%	100%	
Total	216*	653	869	
	24.9%	75.1%	100%	
Veterinärmedizin	Bern	30	74	104
		28.8%	71.2%	100%
	Zürich	21	92	113
		18.6%	81.4%	100%
	Total	51	166	217
23.5%		76.5%	100%	

\* 2 Anmeldungen für Neuenburg

**Tabelle 41:** Zulassungsquoten für Disziplin und Wunschuniversität (erste Wahl)

	Value	df	Asymp. Sig.
Pearson Chi-Square	19.57	4	<b>0.00</b>
Likelihood Ratio	19.05	4	<b>0.00</b>
Cramer's V	.150		

**Tabelle 42:** Signifikanz für Zulassungsquoten nach Wunschuniversität (HM)

Die Verteilung der Testwerte aller Testteilnehmer nach Wunschuniversitäten ist in Abbildung 47 und Tabelle 43 dargestellt. Eine signifikante Differenz besteht einzig zwischen Freiburg und Zürich – diese bleibt auch nach dem Wegfall der Personen auf der Warteliste bestehen.

Der Unterschied dieser beiden Universitäten ist in Zusammenhang mit den Ergebnissen bezüglich der Sprachversionen zu betrachten. 40% der Personen, die Freiburg als Wunschuniversität angegeben haben, haben den Test in französischer Sprache absolviert, 52.5% in deutsch. In Zürich und Basel lag der Anteil der deutschsprachigen Teilnehmer über 95%, in Bern bei 88%.

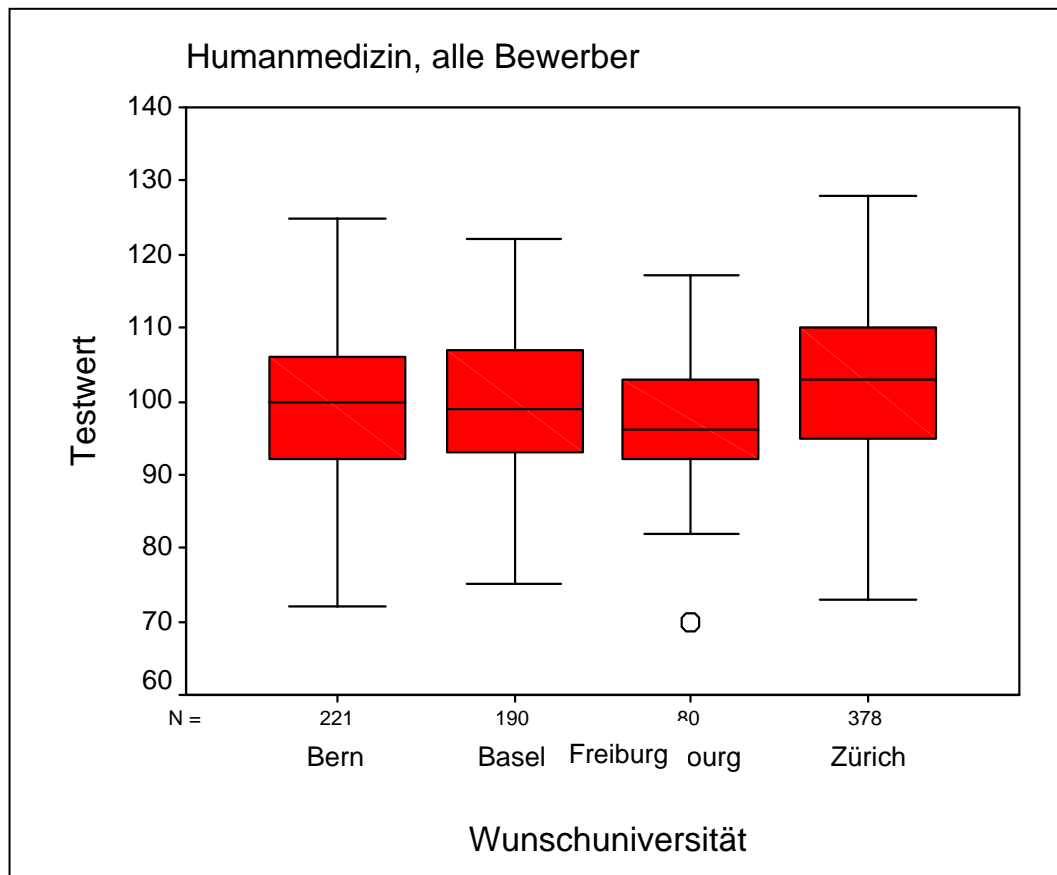


Abbildung 47: Boxplots für Testwerte aller Teilnehmer nach Wunschuniversitäten

Wunsch- universität	Teilnehmer gesamt			zugelassen		
	N	Mittel- wert	Standard- abw.	N	Mittel- wert	Standard- abw.
Bern	221	99.86	10.40	159	104.85	7.27
Basel	190	99.37	9.74	138	103.82	6.93
Freiburg	80	97.25*	8.81	51	102.27*	6.24
Zürich	378	102.43*	9.88	305	105.91*	7.33
Total	869	100.63	10.02	653	104.92	7.22

Tabelle 43: Testwerte nach Wunschuniversitäten, alle Teilnehmer und Zugelassene (HM),

\*signifikante Differenzen

## 8 Ergebnisse zur Testgüte

### 8.1 Zuverlässigkeit

Die Reliabilität (deutsche Ergebnisse vgl. Trost et al., 1995) kann anhand zweier Koeffizienten verglichen werden. Zunächst wurde die Zuverlässigkeit der Untertests abgeschätzt. Beim Untertest „Konzentriertes und sorgfältigen Arbeiten“ ist diese Berechnung aufgrund seiner anderen Struktur nicht möglich. Seine Zuverlässigkeit wird indirekt dadurch nachgewiesen, dass das Weglassen dieses Untertests die Güte des Punktwertes verringern würde.

Die internen Konsistenzen (Cronbachs Alpha) schätzen die Messgenauigkeit anhand der Korrelationen jeder Aufgabe mit allen anderen des entsprechenden Untertests. Wie sich zeigt, liegen 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. Andererseits liegen dieses Jahr auch einige Konsistenzen über den aus Deutschland bekannten Werten.

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

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

Tabelle 44: Zuverlässigkeit 1998 bis 2003; Vergleich der Skalenkonsistenzen für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

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

Tabelle 45: Zuverlässigkeit 1998 bis 2003; Vergleich der Reliabilitäten nach der Testhalbierungsmethode für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

Die **Reliabilität des Punktwertes** insgesamt – geschätzt nach der Testhalbierungsmethode (gerade vs. ungerade Itemnummern) auf der Basis aller neun Untertests liegt mit 0.92 weiterhin in einem sehr hohen Bereich und in der aus Deutschland bekannten Grössenordnung.

Wie in anderen Jahren würde ein Weglassen von einzelnen Untertests zu einer Verringerung der Gesamtkonsistenz des Testprofils führen (Tabelle 46). Diese fällt mit 0.81 erneut hoch aus.

Untertest	Konsistenz ohne den Untertest 1998	Konsistenz ohne den Untertest 1999	Konsistenz ohne den Untertest 2000	Konsistenz ohne den Untertest 2001	Konsistenz ohne den Untertest 2002	Konsistenz ohne den Untertest 2003
Muster zuordnen	.82	.80	.80	.78	.79	.80
Med.-nat. Grundverständnisse	.79	.78	.78	.77	.77	.78
Schlauchfiguren	.80	.79	.79	.78	.78	.79
Quant. u. form. Probleme	.80	.79	.78	.78	.77	.78
Textverständnis	.80	.79	.78	.77	.78	.77
Figuren lernen	.81	.80	.80	.78	.78	.79
Fakten lernen	.82	.81	.80	.79	.79	.80
Diagramme und Tabellen	.80	.78	.78	.77	.77	.78
Konz. u. sorgf. Arbeiten	.82	.81	.80	.79	.79	.80
<b>TESTPROFIL</b>	<b>.83</b>	<b>.81</b>	<b>.81</b>	<b>.80</b>	<b>.80</b>	<b>.81</b>

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.

Die Korrelationen der Untertests mit dem Punktwert sind weiterhin hoch. Ausnahme ist auch hier der Untertest „Muster zuordnen“, wo der Zusammenhang erstmals 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		0.19	0.36	0.24	0.18	0.32	0.15	0.25	0.30	0.49
Med.-naturw. Grundverst.	0.19		0.26	0.56	0.71	0.18	0.30	0.62	0.10	0.69
Schlauchfiguren	0.36	0.26		0.32	0.30	0.38	0.22	0.34	0.43	0.64
Quant. u. formale Probl.	0.24	0.56	0.32		0.53	0.22	0.21	0.62	0.18	0.68
Textverständnis	0.18	0.71	0.30	0.53		0.25	0.32	0.61	0.17	0.73
Figuren lernen	0.32	0.18	0.38	0.22	0.25		0.39	0.20	0.47	0.61
Fakten lernen	0.15	0.30	0.22	0.21	0.32	0.39		0.21	0.30	0.56
Diagramme und Tabellen	0.25	0.62	0.34	0.62	0.61	0.20	0.21		0.16	0.70
Konzent. u. sorgfält. Arbeiten	0.30	0.10	0.43	0.18	0.17	0.47	0.30	0.16		0.55
Punktwert	0.49	0.69	0.64	0.68	0.73	0.61	0.56	0.70	0.55	

Tabelle 47: Korrelationen zwischen Punktwerten und Untertests

Abbildung 48 zeigt, dass die Größenordnung der Korrelationen mit der Gesamtpunktzahl für die einzelnen Untertests über die Jahre stabil bleibt. Der Untertest „Muster zuordnen“ fällt dieses Jahr als einziger etwas ab.



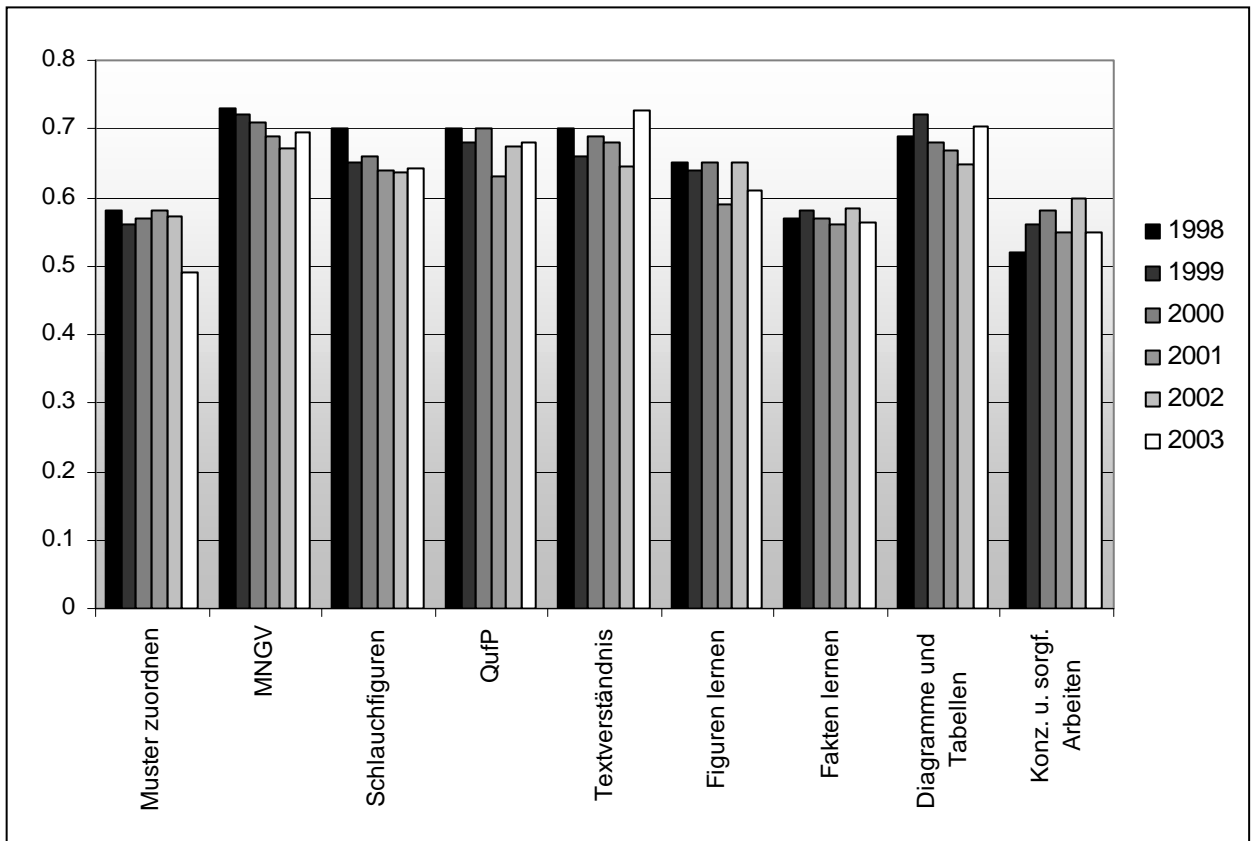


Abbildung 48: Korrelationen zwischen Punktwerten und Untertests nach Jahren

Die geringsten Zusammenhänge sind unter den Untertests „Muster zuordnen“, „Fakten lernen“ sowie „Konzentriertes und sorgfältiges Arbeiten“ zu finden. Die bereits bekannte Drei-Faktoren-Struktur bleibt auch in diesem Jahr stabil. Die Ladungsmuster stimmen weiterhin gut überein. Tabelle 48 zeigt die Vergleichbarkeit der unrotierten Lösungen über die letzten fünf Jahre. Erneut laden alle Untertests signifikant im ersten unrotierten Faktor (Tabelle 50).

Unrotierte Lösung															
Faktor	Eigenwert					% Varianz					Kumuliert %				
	2003	2002	2001	2000	1999	2003	2002	2001	2000	1999	2003	2002	2001	2000	1999
1	3.64	3.61	3.53	3.70	3.67	40.48	40.14	39.21	41.12	40.78	40.48	40.14	39.21	41.12	40.78
2	1.60	1.54	1.63	1.46	1.36	17.78	17.12	18.14	16.18	15.16	58.26	57.26	57.35	57.30	55.94
3	0.92	0.77	.85	.84	.91	10.28	8.53	9.39	9.31	10.16	68.54	65.79	66.74	66.61	66.10
4	0.67	0.68	.62	.66	.64	7.50	7.51	6.93	7.34	7.10	76.03	73.30	73.67	73.95	73.20
5	0.54	0.58	.57	.58	.59	5.95	6.41	6.35	6.40	6.50	81.98	79.71	80.03	80.35	79.70
6	0.50	0.53	.55	.53	.54	5.54	5.94	6.06	5.93	5.95	87.52	85.64	86.09	86.28	85.65
7	0.48	0.49	.50	.47	.53	5.38	5.40	5.53	5.23	5.86	92.90	91.05	91.61	91.51	91.50
8	0.35	0.42	.42	.39	.42	3.93	4.67	4.63	4.31	4.64	96.83	95.72	96.24	95.82	96.15
9	0.29	0.39	.34	.38	.35	3.17	4.28	3.76	4.18	3.85	100	100	100	100	100

Tabelle 48: Verlauf der unrotierten Lösung

In Tabelle 49 sind die rotierten Zwei- und Dreifaktorenlösungen dargestellt. Die Ladungen der einzelnen Untertests können Tabelle 51, beziehungsweise Tabelle 52 entnommen werden. Die Zweifaktorenlösung klärt 58% der Gesamtvarianz auf, die Dreifaktorenlösung 69%.

Varimax-rotierte Lösungen														
Eigenwert					% Varianz					Kumuliert %				
2003	2002	2001	2000	1999	2003	2002	2001	2000	1999	2003	2002	2001	2000	1999
<b>Zwei-Faktorenlösung</b>														
2.87	2.66	2.72	2.74	2.72	31.90	29.60	30.19	30.49	30.25	31.90	29.60	30.19	30.49	30.25
2.37	2.49	2.44	2.41	2.31	26.36	27.66	27.16	26.80	25.69	58.26	57.26	57.35	57.30	55.94
<b>Drei-Faktorenlösung</b>														
2.84	2.61	2.71	2.71	2.68	31.57	29.05	30.06	30.12	29.81	31.57	29.05	30.06	30.12	29.81
1.83	1.67	1.82	1.92	1.81	20.34	18.53	20.21	21.29	20.12	51.91	47.57	50.27	51.41	49.93
1.50	1.64	1.48	1.37	1.46	16.623	18.22	16.47	15.20	16.17	68.54	65.79	66.74	66.61	66.10

Tabelle 49: Varianzanteile der einzelnen Faktorenlösungen (rotierte Lösungen)

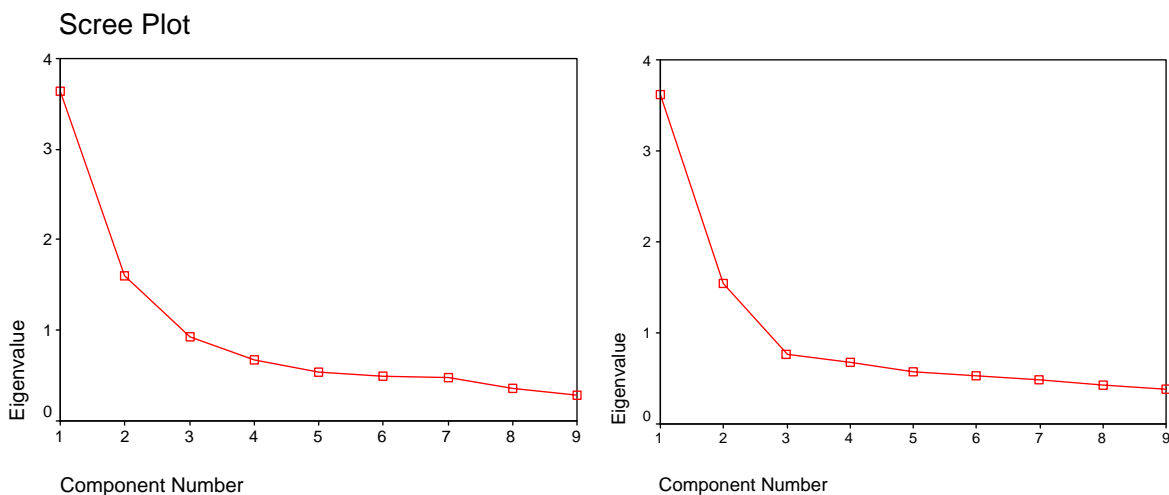


Abbildung 49: Scree-Plot (Eigenwerteverlauf) der Faktoren 1 bis 9 für die Untertests des EMS 2003 (links) und 2002 (rechts)

Im „Scree-Plot“ (Abbildung 49) wird der Eigenwerteverlauf der Faktoren dargestellt. Der Verlauf für den EMS 2003 befindet sich links, rechts davon zum Vergleich der Verlauf aus dem Vorjahr. Die Ein-Faktorenlösung ist erneut deutlich hervorgehoben, die Zweifaktorenlösung und bedingt auch die Dreifaktorenlösung unterscheiden sich weniger von den jeweils nachfolgenden. Ab der Dreifaktorenlösung sind die Eigenwerte kleiner als 1 – der Varianzanteil des Faktors ist kleiner als der eines einzelnen Untertests.

Die Abbildung 50 zur Struktur der Untertests fasst die Ergebnisse zusammen. Die Ladungsmatrix des ersten Faktors der unrotierten Lösung zeigt substantielle Ladungen für alle Untertests. Ein „Generalfaktormodell“ ist indiziert. Der allgemeine Fähigkeitsfaktor ist sehr varianzstark. Die Varianzaufklärung bleibt über alle bisherigen Testdurchführungen konstant.

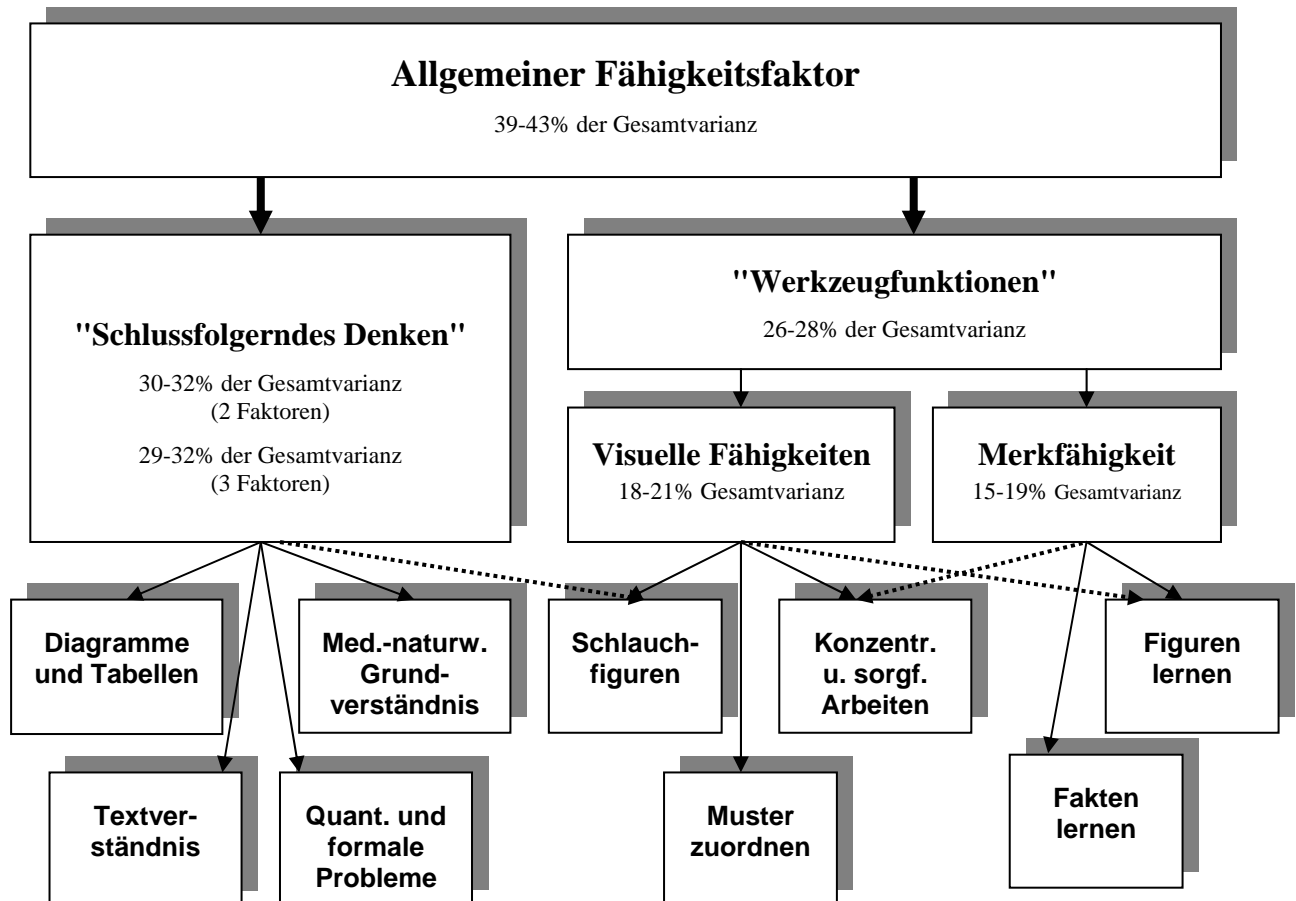


Abbildung 50: Struktur der Untertests des EMS, hierarchische Faktorenstruktur, angegeben auch der Wertebereich der Varianzaufklärung durch den Faktor 1998 bis 2003

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

Tabelle 50: Ladungen im ersten Faktor der unrotierten Lösung 1998 bis 2003 („Generalfaktor“ bzw. allgemeiner Fähigkeitsfaktor)

	Faktor 1						Faktor 2						Kommunalitäten (h <sup>2</sup> )					
	98	99	00	01	02	03	98	99	00	01	02	03	98	99	00	01	02	03
Diagramme und Tabellen	.83	.76	.79	.77	.77	.83				.21	(.17)	(.17)	.71	.65	.66	.63	.63	.73
Med.-naturwiss. Grundverständnis	.79	.81	.78	.82	.77	.86	.22	.21	.22	(.18)	.21	(.11)	.68	.71	.66	.70	.63	.76
Textverständnis	.77	.80	.80	.82	.82	.83				(.16)	(.11)	(.18)	.64	.66	.67	.70	.68	.74
Quant. und formale Probleme	.77	.78	.80	.81	.79	.77					(.18)	.20	.62	.63	.68	.66	.65	.65
Konzentr. u. sorgf. Arbeiten							.78	.79	.81	.79	.82	.79	.62	.62	.66	.62	.68	.62
Figuren lernen	.23	(.16)	(.15)	(.16)	.20		.70	.72	.64	.66	.70	.78	.55	.55	.44	.46	.52	.65
Muster zuordnen				(.15)	.20		.66	.64	.73	.69	.63	.58	.48	.43	.54	.50	.43	.62
Fakten lernen	.20	.25	.27	.15	(.11)	.26	.59	.51	.51	.62	.67	.50	.40	.38	.34	.41	.47	.78
Schlauchfiguren	.45	.33	.30	.25	.29	.26	.54	.59	.65	.66	.60	.66	.50	.46	.51	.50	.45	.60

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

In der Zwei-Faktorenlösung zeigen sich „Schlussfolgerndes Denken“ und „Werkzeugfunktionen“ als Faktoren. Der Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“ ist dieses Jahr der ladungsstärkste Untertest für den ersten Faktor, für den zweiten Faktor ist dies wie bisher „Konzentriertes und sorgfältiges Arbeiten“.

Auch die Drei-Faktorenlösung bleibt stabil: Der Faktor „Werkzeugfunktionen“ teilt sich in die Komponenten „Wahrnehmung“ und „Gedächtnis“ auf – der Faktor „Schlussfolgerndes

Denken“ bleibt bezüglich der Ladungen und des Varianzanteiles gegenüber der Zwei-Faktorenlösung stabil.

Faktor	Schweiz Gesamt								Deutschland		
	1		2		3		h <sup>2</sup>		1	2	3
	2002	2003	2002	2003	2002	2003	2002	2003			
Diagramme und Tabellen	<b>.76</b>	<b>.83</b>	.26	.22			.65	.74	<b>.82</b>		
Medizin.-naturwiss. Grundverständnis	<b>.78</b>	<b>.86</b>			.23	.16	.67	.76	<b>.81</b>		.20
Quantitat. u. formale Probleme	<b>.78</b>	<b>.77</b>	.20	.25			.65	.65	<b>.80</b>	.18	
Textverständnis	<b>.82</b>	<b>.82</b>			.13	.24	.70	.74	<b>.79</b>		.20
Muster zuordnen	.15	.15	<b>.78</b>	<b>.77</b>	.14		.65	.62		<b>.81</b>	
Konzentr. u. sorgfält. Arbeiten			<b>.59</b>	<b>.61</b>	.58	.50	.68	.63		<b>.70</b>	.40
Schlauchfiguren	.25	.26	<b>.70</b>	<b>.71</b>	.18	.17	.59	.60	.35	<b>.71</b>	
Fakten lernen	.13	.24			<b>.83</b>	<b>.85</b>	.71	.78	.21		<b>.87</b>
Figuren lernen	.20		.25	.49	<b>.72</b>	<b>.64</b>	.62	.65	.13	.47	<b>.64</b>

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

Die ladungsstärksten Untertests in den jeweiligen Faktoren sind „Medizinisch-naturwissenschaftliches Grundverständnis“, „Muster zuordnen“ und „Fakten lernen“.

### 8.3 Item-Trennschärfen

Die Item-Trennschärfen (korrigierte Item-Total-Korrelation) folgen den deutschen Kennwerten. Fehlende Itemnummern bezeichnen Einstreuaufgaben.

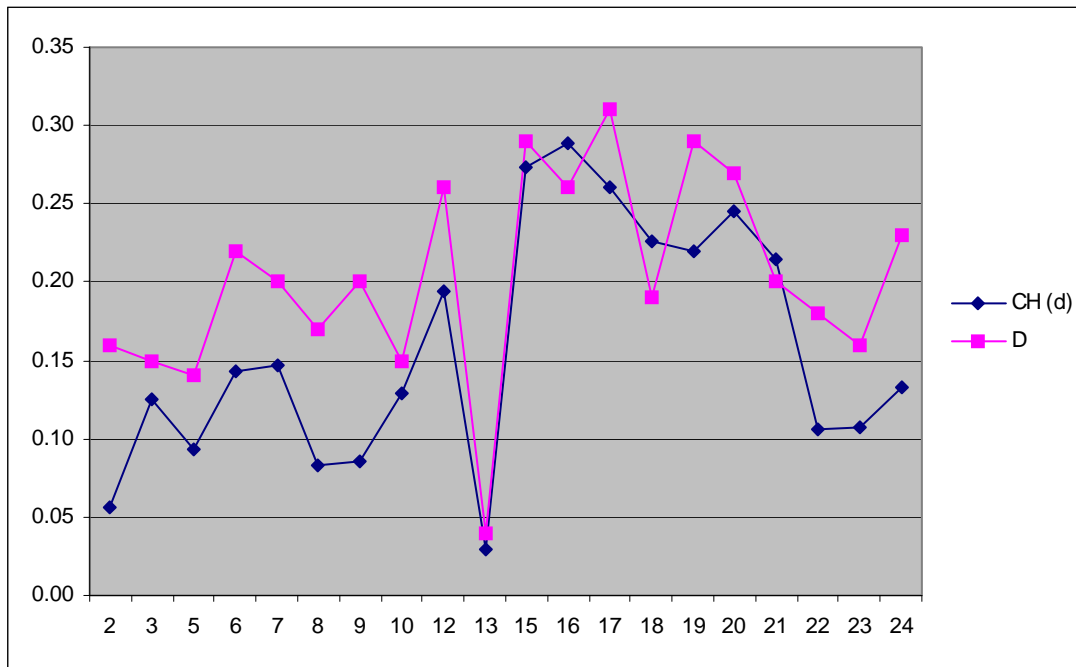


Abbildung 51: Trennschärpen für den Untertest „Muster zuordnen“

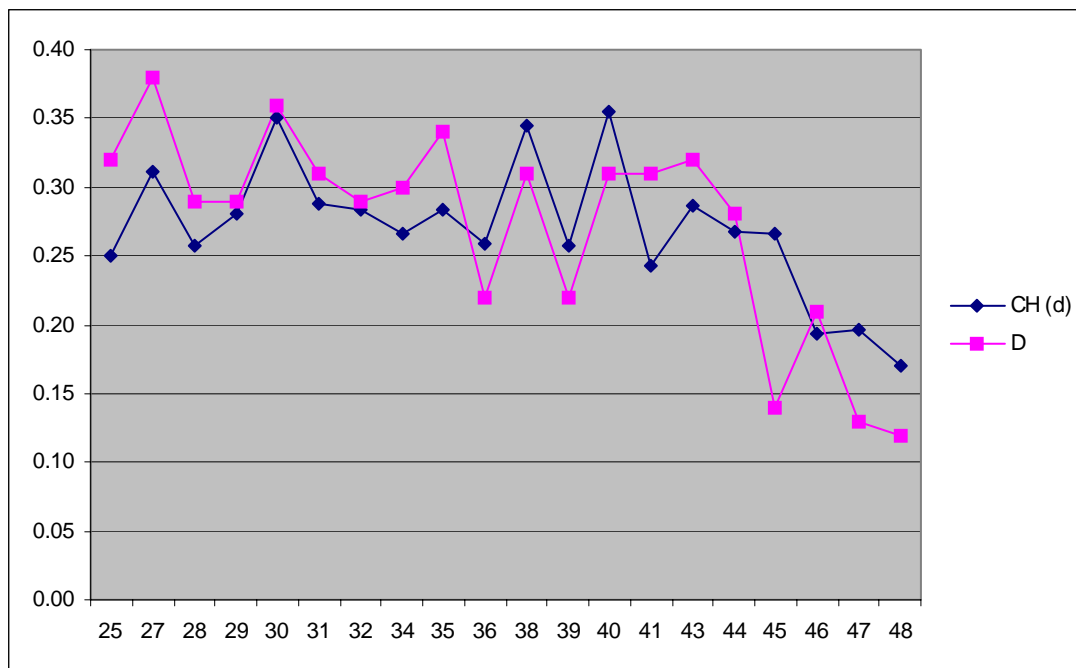


Abbildung 52: Trennschärpen für den Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“

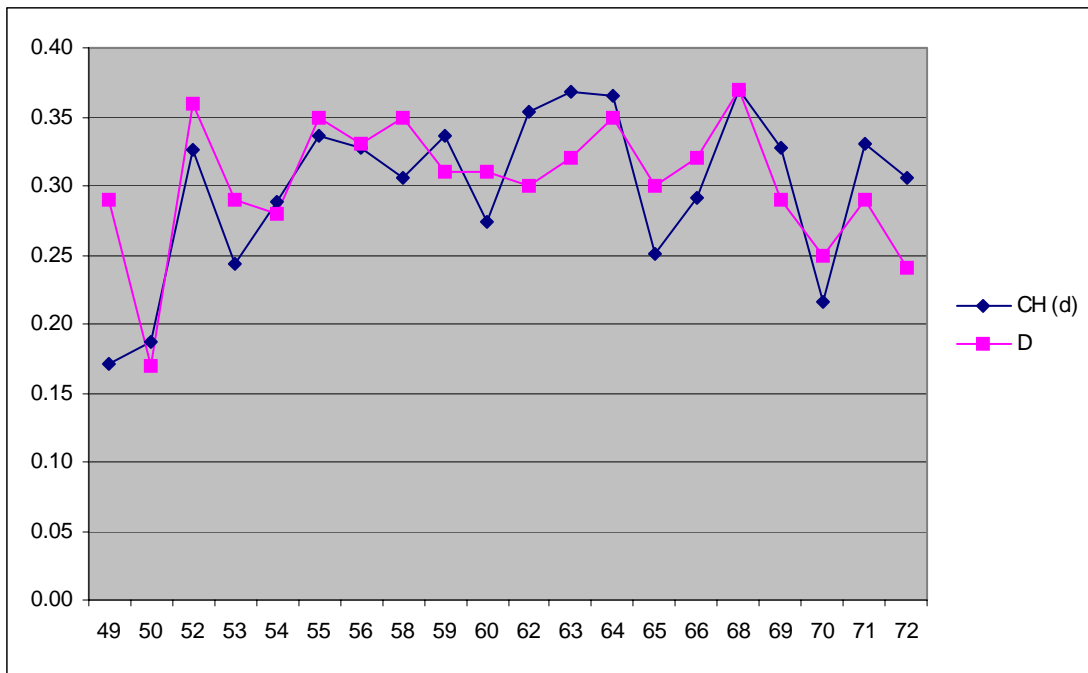


Abbildung 53: Trennschärfen für den Untertest „Schlauchfiguren“

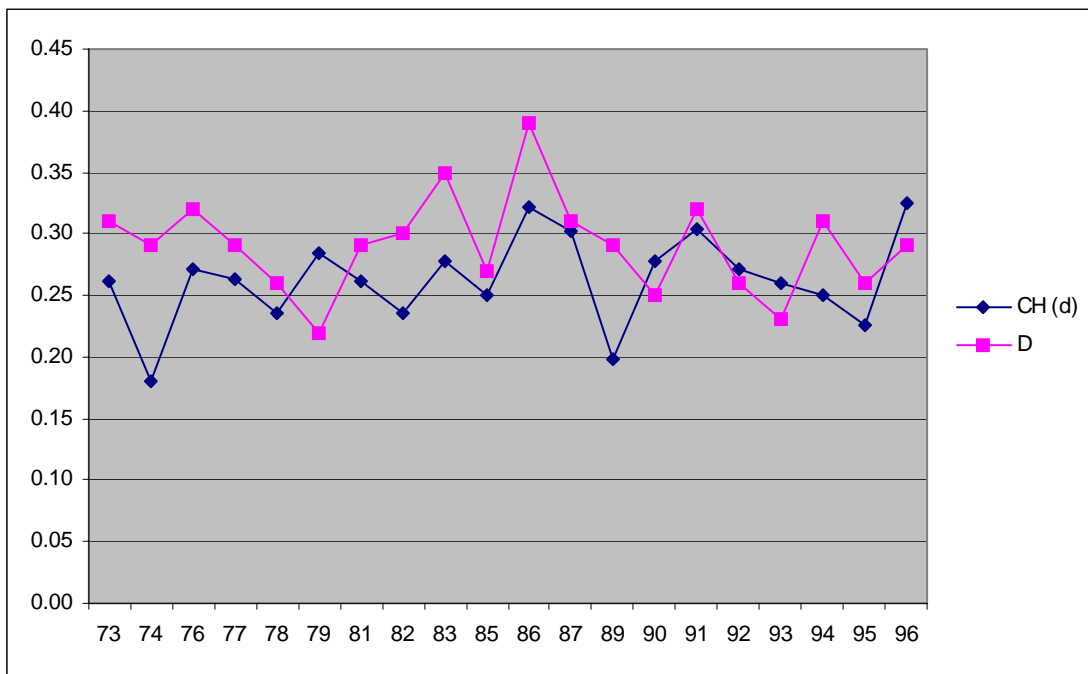


Abbildung 54: Trennschärfen für den Untertest „Quantitative und formale Probleme“

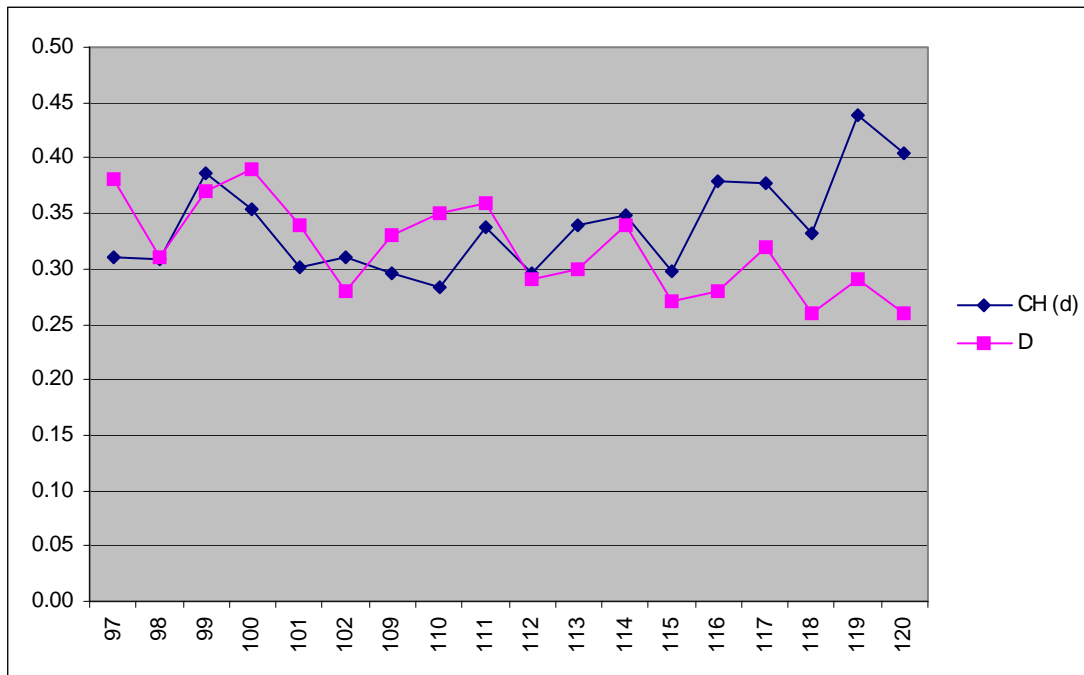


Abbildung 55: Trennschärfen für den Untertest „Textverständnis“

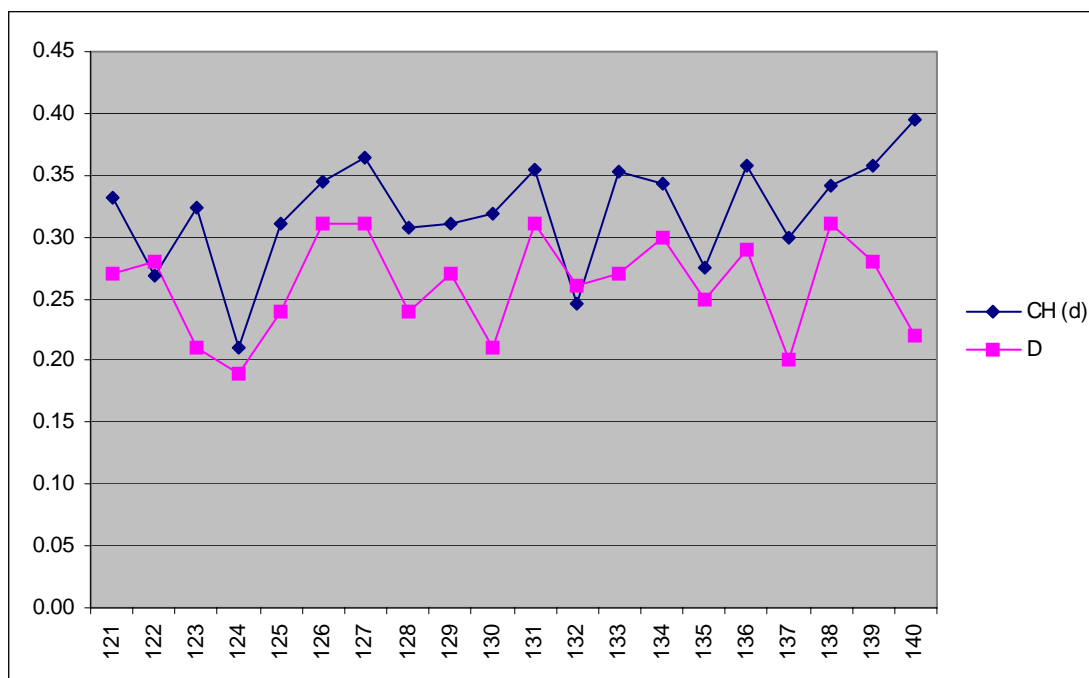


Abbildung 56: Trennschärfen für den Untertest „Figuren lernen“



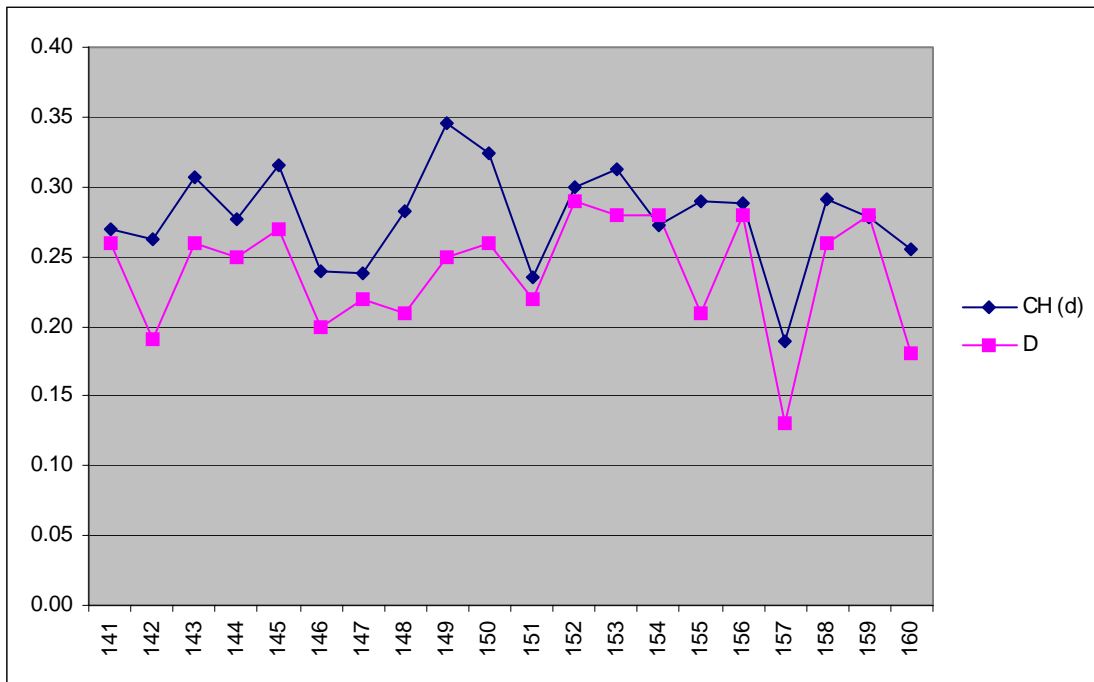


Abbildung 57: Trennschärfen für den Untertest „Fakten lernen“

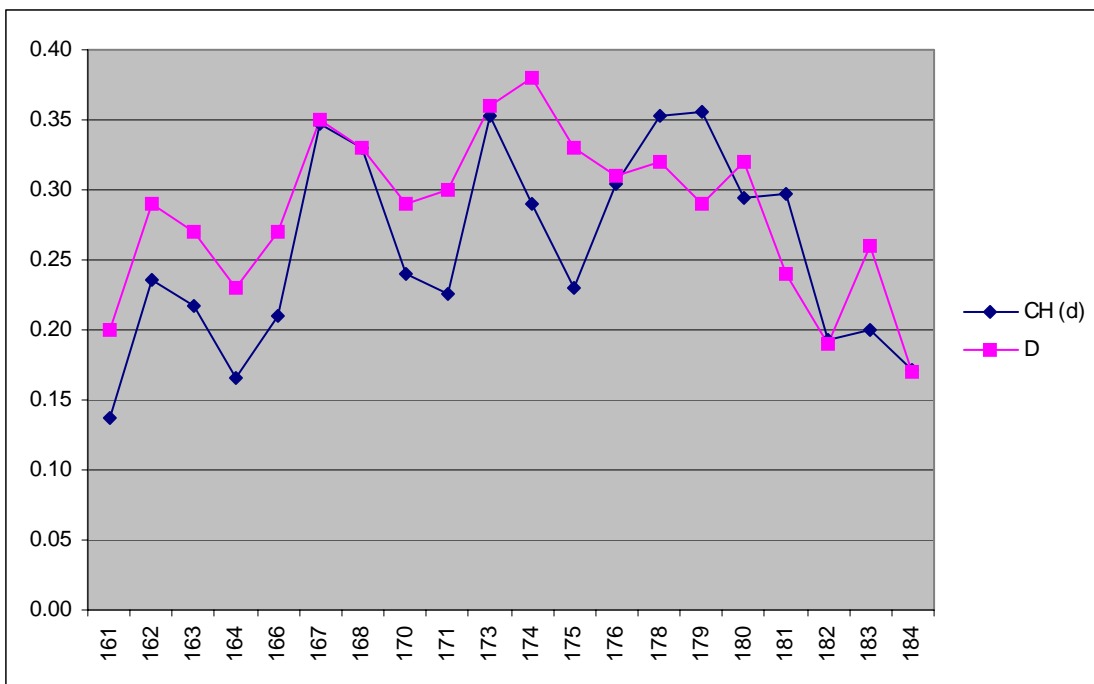


Abbildung 58: Trennschärfen für den Untertest „Diagramme und Tabellen“

## 8.4 Konzentriertes und sorgfältiges Arbeiten

In diesem Untertest kann die Zuverlässigkeit nicht nach dem Itemkonzept eingeschätzt werden. Der Untertest besteht aus 1200 Zeichen, die innerhalb einer vorgegebenen Zeit bearbeitet werden müssen. Jedes Zeichen kann dabei als ein Item aufgefasst werden. Der Punktwert wird aufgrund der Menge der korrekt bearbeiteten Zeichen (total bearbeitete minus falsch bearbeitete Zeichen) berechnet. Dabei können zwei Fehlertypen auftreten: Übersehene oder fälschlicherweise markierte Zeichen (Tabelle 53).

	Deutsch		Französisch		Italienisch	
	m	s	m	s	m	s
<b>2003</b>						
Richtig	427.46	106.85	339.61	77.07	390.79	86.23
Falsch	19.22	20.89	13.43	10.41	15.74	13.16
... übersehen	10.83	17.16	7.86	9.07	6.21	6.05
... fälschlich markiert	8.39	9.26	5.58	4.56	9.52	9.97
<b>2002</b>						
Richtig	424.85	103.41	325.03	102.16	358.75	82.61
Falsch	22.64	23.90	27.91	45.97	16.21	15.08
... übersehen	14.27	20.68	18.86	38.80	7.64	8.88
... fälschlich markiert	8.37	8.48	9.05	10.31	8.57	10.36
<b>2001</b>						
Richtig	420.37	109.99	338.72	83.54	350.98	87.43
Falsch	20.05	20.81	13.18	10.45	15.95	17.49
... übersehen	12	17.72	8.69	9.28	10.24	15.13
... fälschlich markiert	8.05	9.01	4.49	4.49	5.71	5.77
<b>2000</b>						
Richtig	404.70	105.37	322.26	84.44	355.59	103.73
Falsch	16.41	20.61	12.87	36.50	9.03	14.92
... übersehen	12.95	19.27	10.65	36.43	6.59	12.88
... fälschlich markiert	3.46	4.14	2.22	2.56	2.45	2.67
<b>1999</b>						
Richtig	391.24	98.11	308.12	93.35	352.16	75.76
Falsch	12.72	13.03	20.99	63.18	9.23	7.99
<b>1998</b>						
Richtig	393.38	95.55	345.74	97.57	359.48	73.05
Falsch	11.77	16.92	13.82	17.64	8.55	7.86

Tabelle 53: Kennwerte für den Untertest „Konzentriertes und sorgfältiges Arbeiten“

Ein signifikanter, allerdings vernachlässigbar geringer Zusammenhang besteht dieses Jahr zwischen der Anzahl richtig markierter und übersehener Zeichen. Die Zusammenhänge mit

der Gesamtzahl falscher Zeichen sind nicht erstaunlich – es handelt sich hier um die beiden Fehler-Unterkategorien. Die beiden Fehlerkategorien ihrerseits korrelieren ebenfalls mit 0.18 signifikant untereinander.

	Falsch	...übersehen	...fälschlich durchgestrichen
Richtig	.03	<b>-.07</b>	<b>.19</b>
Falsch	<b>.35</b>	<b>.03</b>	<b>.00</b>
...übersehen		<b>.90</b>	<b>.60</b>
		<b>.00</b>	<b>.00</b>
			<b>.18</b>
			<b>.00</b>

Tabelle 54: Korrelationen der Parameter im Konzentrationstest; kursiv: Signifikanzniveau

Seit 1999 steigt die durchschnittliche Anzahl richtig bearbeiteter Zeichen in der deutschsprachigen Teilnehmergruppe stetig an (Abbildung 59). Die Leistung der italienischsprachigen Kandidaten blieb bis 2002 konstant, in diesem Jahr ist eine Leistungssteigerung zu verzeichnen. Die französischsprachigen Teilnehmer erreichen wieder die Werte von 2001. Die beste Leistung wurde dort 1998 erreicht.

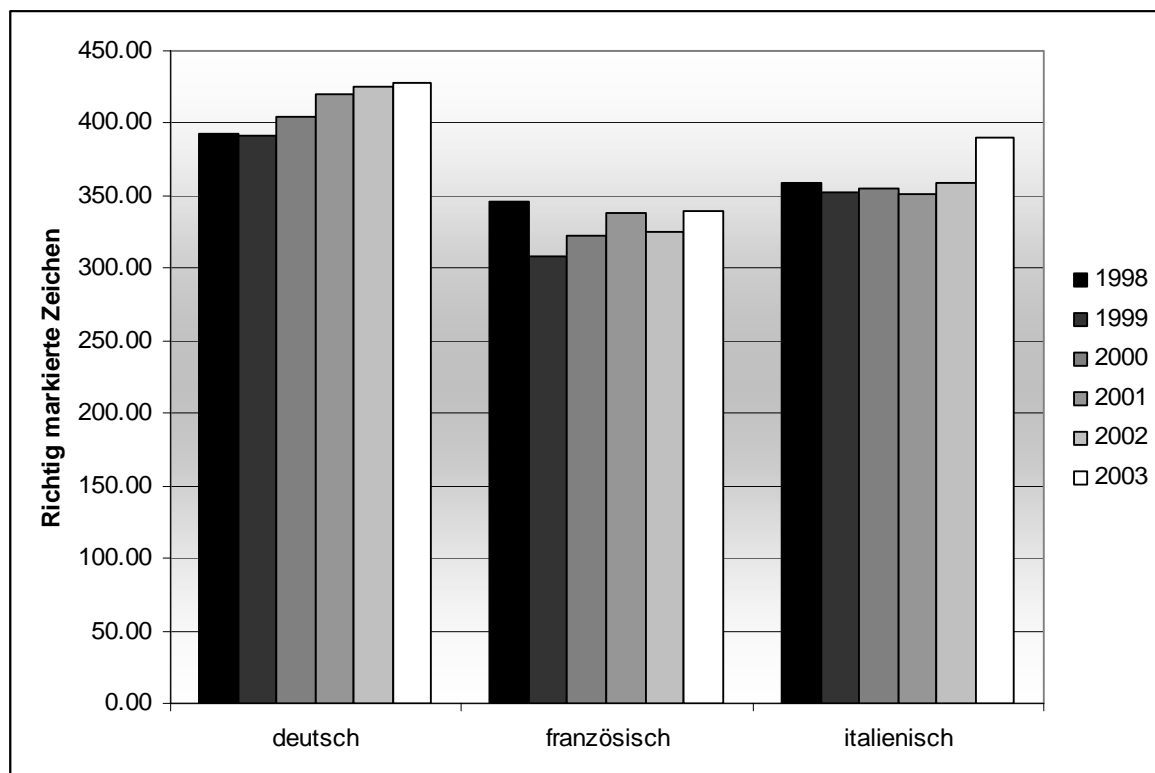
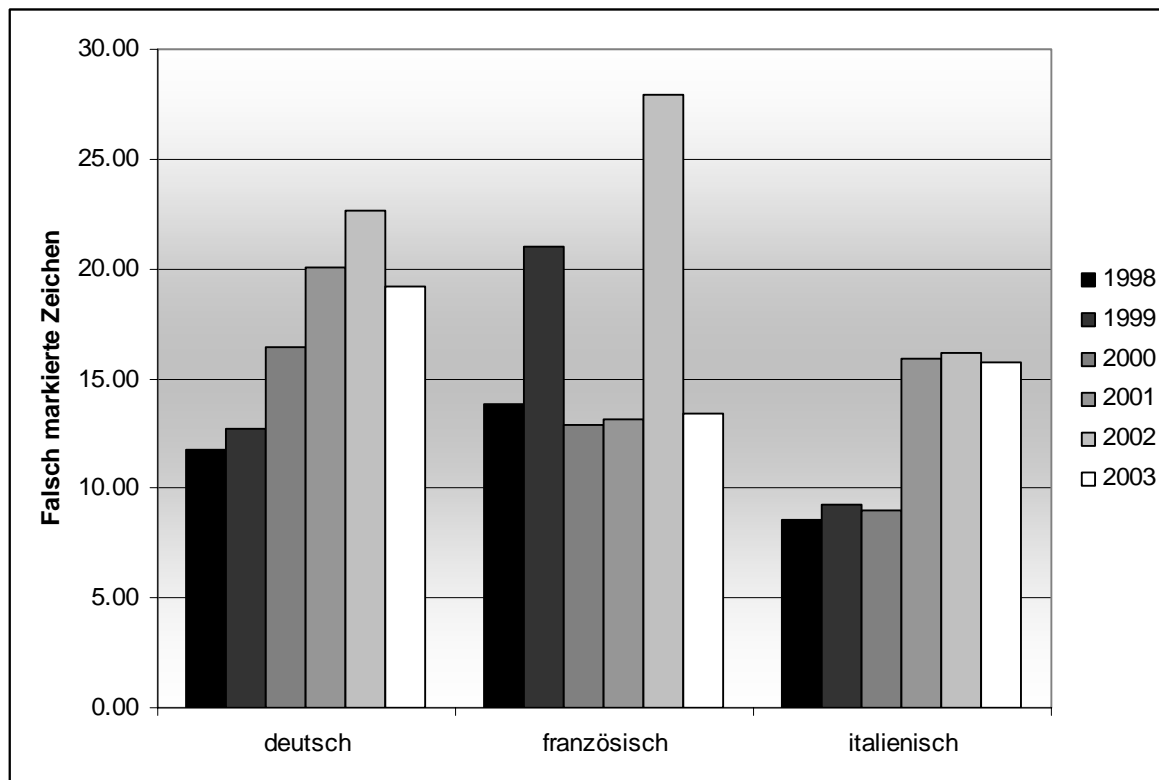


Abbildung 59: Anzahl richtig bearbeiteter Zeichen im Untertest „Konzentriertes und sorgfältiges Arbeiten“, nach Jahren und Sprachgruppen

Betrachtet man die durchschnittliche Anzahl falsch bearbeiteter Zeichen, so zeigt sich für die deutschsprachige Gruppe von 1998 bis 2002 ein stetiger Anstieg. 2003 ist diese Anzahl erstmals wieder gesunken. Die französische Gruppe erreicht nach einem schlechten Jahrgang wieder das übliche Niveau– wobei hier starke Schwankungen beispielsweise durch einzelne Kandidaten zustande kommen können, welche die Instruktionen nicht korrekt verstanden haben.

Der Wert für die italienischsprachige Gruppe entspricht denjenigen der beiden Vorjahre.



**Abbildung 60:** Anzahl falsch bearbeiteter Zeichen im Untertest „Konzentriertes und sorgfältiges Arbeiten“, nach Jahren und Sprachgruppen

Abbildung 61 gibt den prozentualen Anteil falsch bearbeiteter Zeichen an der durchschnittlichen Anzahl korrekt bearbeiteter Zeichen für die Sprachgruppen wieder. Alle drei Sprachgruppen erreichen dieses Jahr fast identische Werte um 4 bis 4.5%

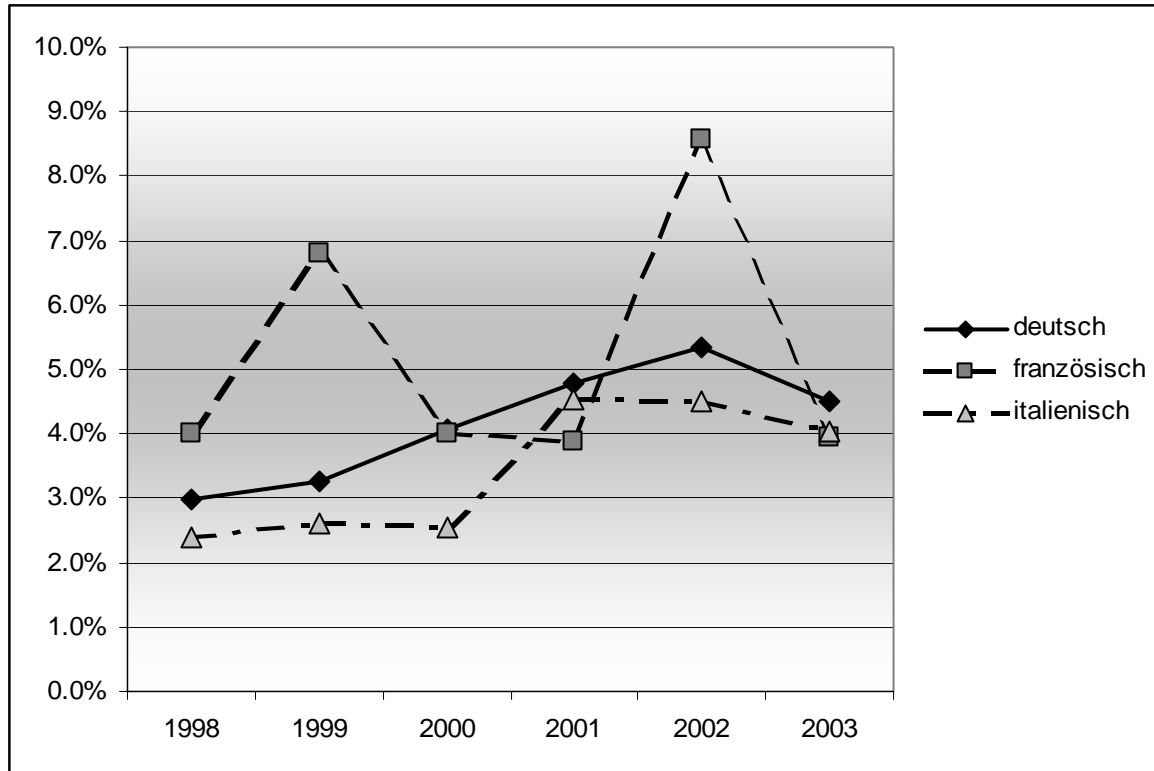


Abbildung 61: Prozentualer Anteil falsch bearbeiteter Zeichen

## 9 Zusammenfassung Kantone 1998 bis 2003

Seit 1998 haben nunmehr 5347 Personen am Test teilgenommen. In der Tabelle 55 sind alle Teilnehmer am EMS 1998 bis 2003 nach Kantonen aufgeführt. Zu beachten ist, dass 1998 nur Humanmedizin, seit 1999 Human- und Veterinärmedizin einem NC unterlagen.

Wohnkanton/ Wohnort	Disziplin		Geschlecht		Sprache			Total
	Human- medizin	Veterinär- medizin	männlich	weiblich	deutsch	franzö- sisch	italie- nisch	
AG	338	60	142	256	398			398
AI	8	4	4	8	12			12
AR	34	4	14	24	38			38
BE	698	153	304	547	835	15	1	851
BL	262	29	117	174	289	2		291
BS	210	17	81	146	226		1	227
FR	202	40	94	148	86	156		242
GE	4	43	13	34	2	45		47
GL	23	4	6	21	27			27
GR	151	33	71	113	177		7	184
JU	2	9	5	6		11		11
LU	262	47	112	197	309			309
NE	5	26	4	27	3	28		31
NW	27	4	9	22	31			31
OW	26		9	17	26			26
SG	305	43	131	217	348			348
SH	51	8	18	41	59			59
SO	155	32	80	107	186	1		187
SZ	90	10	32	68	100			100
TG	138	32	52	118	170			170
TI	153	37	88	102	3		187	190
UR	17	3	10	10	20			20
VD	10	75	34	51	5	80		85
VS	75	28	36	67	69	34		103
ZG	63	22	33	52	84	1		85
ZH	1035	164	416	783	1189	3	7	1199
FL	20	3	10	13	23			23
Ausland	46	7	21	32	52	1		53
<b>Total</b>	<b>4410</b>	<b>937</b>	<b>1946</b>	<b>3401</b>	<b>4767</b>	<b>377</b>	<b>203</b>	<b>5347</b>

**Tabelle 55:** Gesamtzahl der Teilnehmer am EMS 1998 bis 2003. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

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	338	101.7	9.0	74	127	60	99.4	8.7	78	118
AI	8	103.8	7.0	96	116	4	102.0	7.4	95	111
AR	34	100.9	9.3	88	127	4	97.0	7.8	91	108
BE	698	100.5	10.0	72	125	153	96.8	8.9	76	121
BL	262	100.2	10.0	74	125	29	100.0	9.3	74	115
BS	210	96.8	10.2	71	122	17	96.0	9.9	76	117
FR	202	94.9	8.8	70	121	40	94.0	7.9	73	106
GE	4	100.3	10.8	85	110	43	93.4	8.3	76	117
GL	23	102.9	9.4	90	130	4	97.8	6.0	89	102
GR	151	99.8	10.0	76	128	33	97.6	8.4	77	113
JU	2	87.0	2.8	85	89	9	91.8	7.5	85	109
LU	262	101.1	9.5	72	126	47	97.7	9.4	76	118
NE	5	100.2	10.0	91	116	26	90.8	8.2	76	106
NW	27	101.2	11.1	78	121	4	96.5	8.9	88	109
OW	26	100.6	7.4	92	126	0	.	.	.	.
SG	305	101.8	10.1	70	128	43	100.7	8.4	79	118
SH	51	100.8	9.0	82	120	8	97.3	10.9	86	114
SO	155	101.4	9.9	70	127	32	98.9	8.8	80	119
SZ	90	98.9	9.9	72	127	10	100.3	11.5	80	121
TG	138	101.0	9.6	77	122	32	101.6	9.0	82	122
TI	153	96.7	9.3	73	120	37	94.2	7.7	82	112
UR	17	102.5	7.7	87	120	3	101.7	10.5	91	112
VD	10	104.4	8.6	91	119	75	93.6	8.0	78	110
VS	75	97.5	9.6	70	120	28	97.8	10.3	75	116
ZG	63	102.2	10.1	78	128	22	97.7	7.0	84	108
ZH	1035	102.8	10.2	66	131	164	99.0	10.2	76	126
FL	20	98.8	10.4	82	120	3	99.3	7.0	92	106
Ausland	46	97.1	9.7	81	119	7	87.9	9.5	76	99
<b>Total</b>	<b>4410</b>	<b>100.6</b>	<b>10.0</b>	<b>66</b>	<b>131</b>	<b>937</b>	<b>97.2</b>	<b>9.3</b>	<b>73</b>	<b>126</b>

Tabelle 56: Kennwerte für den Testwert von Kantonen 2003 nach Disziplinen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Nachfolgend werden getrennt für Human- und Veterinärmedizin die Testwertverteilungen und die theoretischen (!) Zulassungsquoten verglichen. Letztere sind nach der genauen Kapazität festgelegt und berücksichtigen nicht die Rückzüge und erneute Vergabe dieser Plätze an Personen mit niedrigeren Testwerten als dem theoretischen Zulassungskriterium. Bei der Interpretation sind die Stichprobengrößen der Tabelle 55 unbedingt zu beachten.



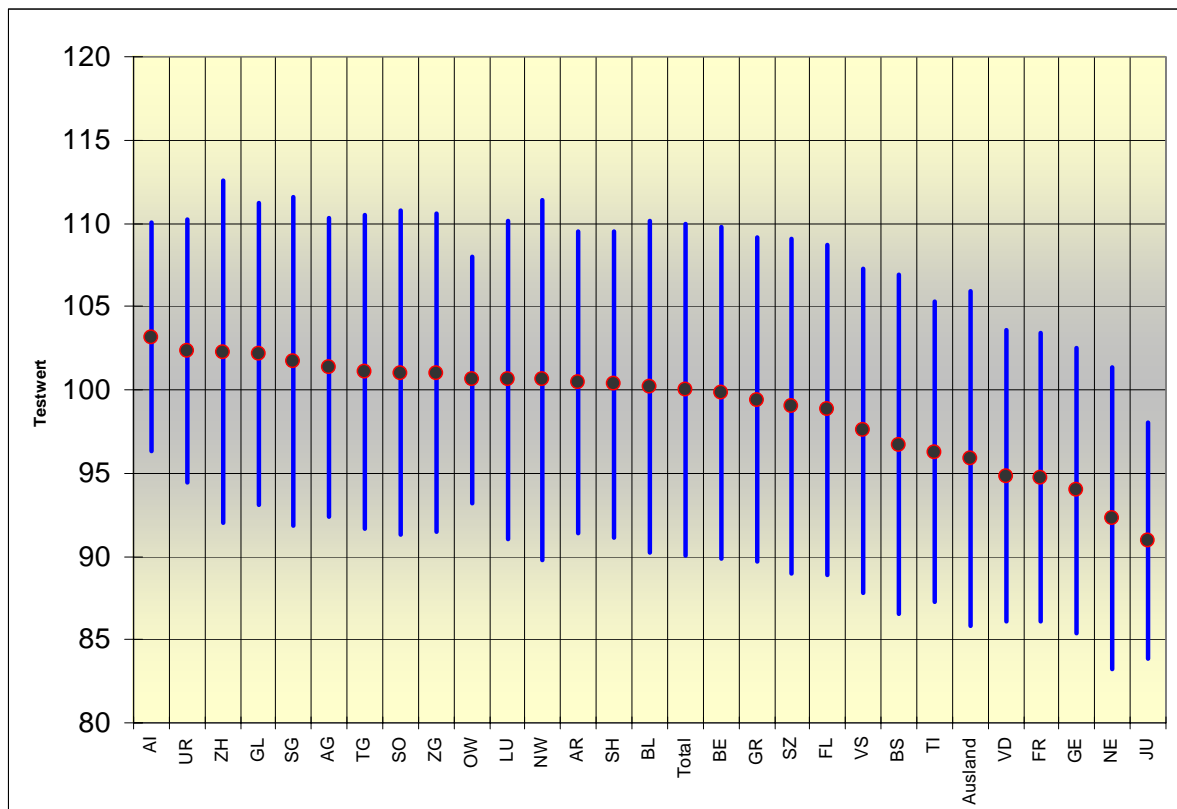


Abbildung 62: Mittelwerte und Streubereiche ( $m \pm s$ ) der Testwerte für Kantone 1998 bis 2003 zusammengefasst; **Humanmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

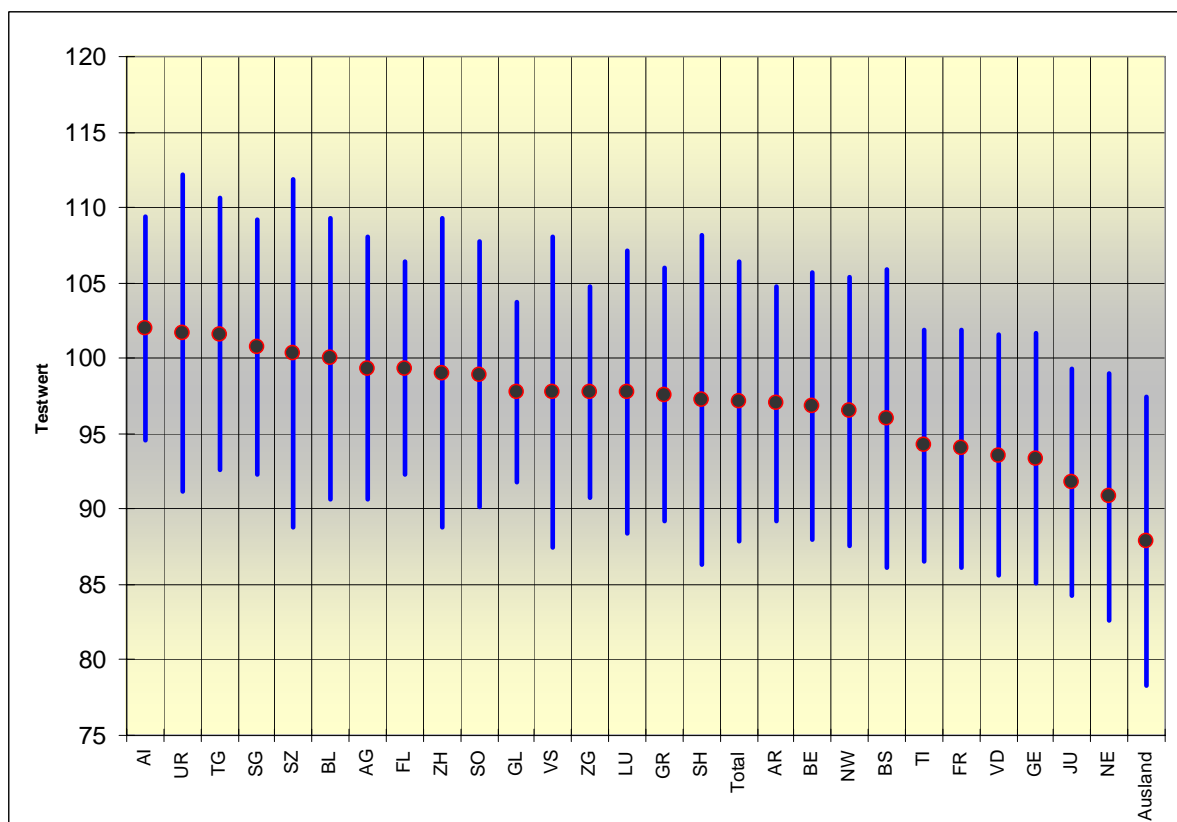


Abbildung 63: Mittelwerte und Streubereiche ( $m \pm s$ ) der Testwerte für Kantone 1999 bis 2003 zusammengefasst; **Veterinärmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

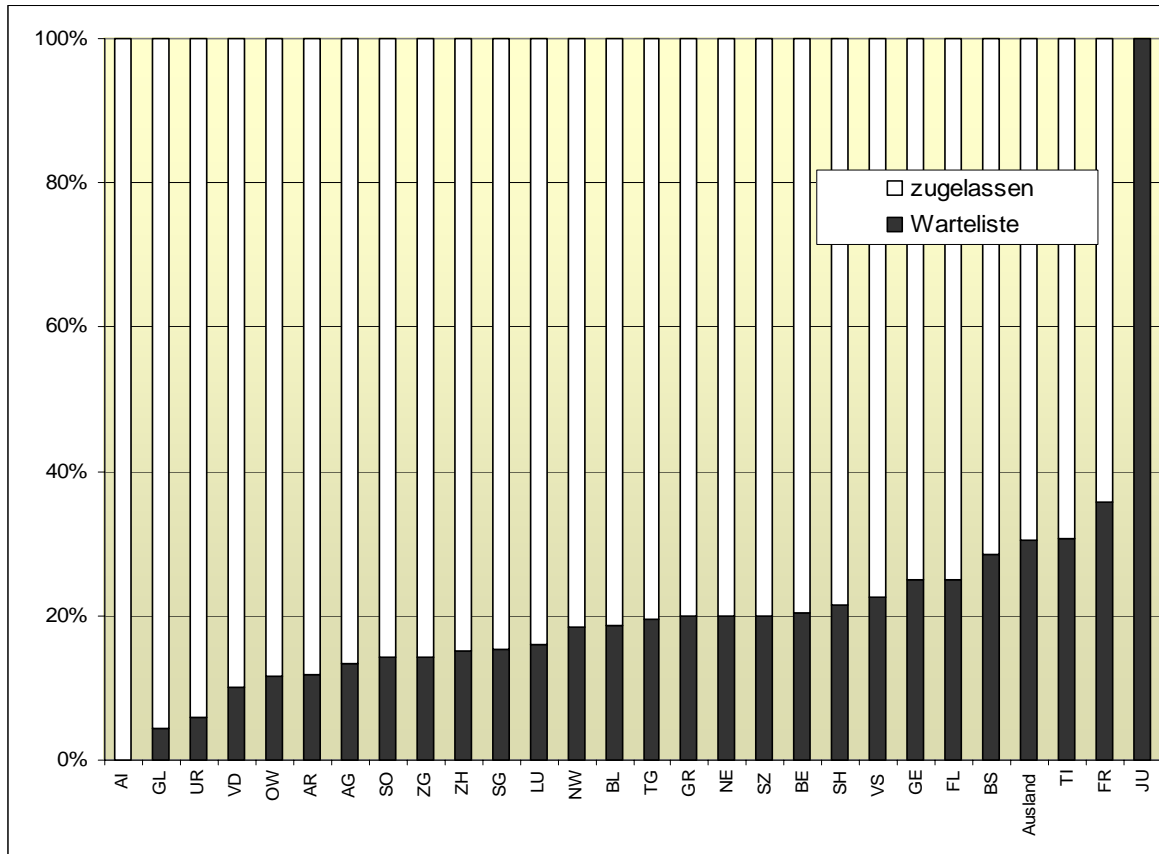


Abbildung 64: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1998 bis 2003; **Humanmedizin**. (JU: nur 2 Personen!)

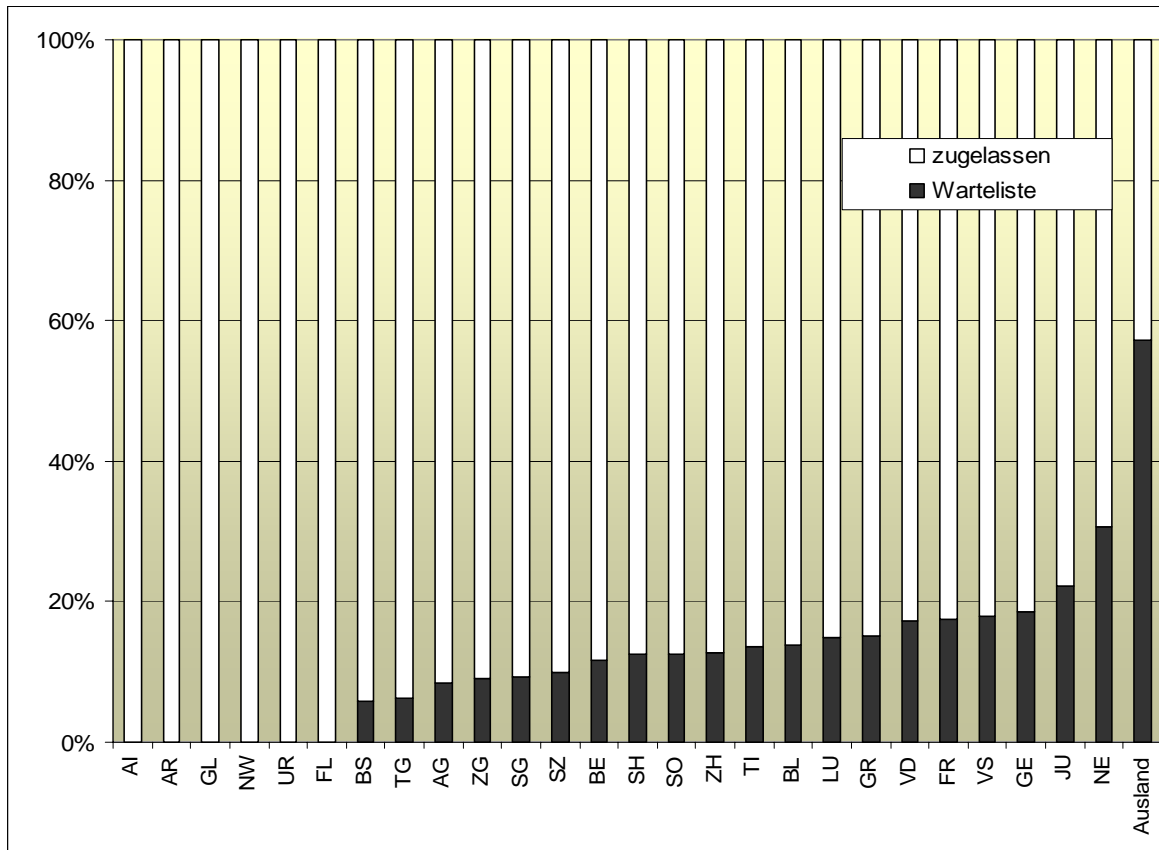


Abbildung 65: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1999 bis 2003; **Veterinärmedizin**

## 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.
- Celio, M.R. (1999) Ausbau des Medizinstudiums in Freiburg? Reformdruck als Chance der zweisprachigen Universität. Neue Zürcher Zeitung. 29.9. 1999.  
[http://www.nzz.ch/online/01\\_nzz\\_aktuell/sonderbeilagen/studium98/studium98celio.htm](http://www.nzz.ch/online/01_nzz_aktuell/sonderbeilagen/studium98/studium98celio.htm)
- 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 Friburgo (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>

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

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

## **11 Anhang zum Bericht 2003**

## 11.1 Beispielaufgaben für die Untertests

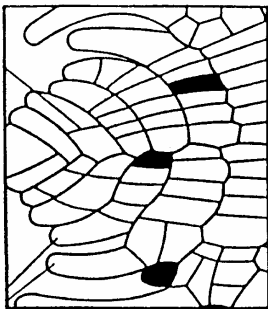
Nachfolgend wird pro Untertest eine Beispielaufgabe dargestellt. So können lediglich die Prinzipien der Aufgabenstruktur verdeutlicht werden – die Aufgaben unterscheiden sich innerhalb jedes Untertests bezüglich des Schwierigkeitsgrades und der Anforderung.

### 11.1.1 Untertest: Muster zuordnen

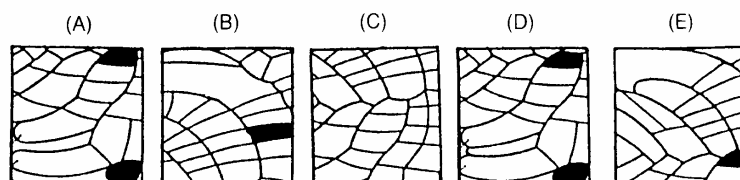
In diesem Untertest wird die Fähigkeit geprüft, Ausschnitte in einem komplexen Bild wiederzuerkennen. Dazu werden pro Aufgabe ein Muster und je fünf Musterausschnitte (A) bis (E) vorgegeben. Die Testteilnehmerin oder der Testteilnehmer soll herausfinden, welcher dieser fünf Musterausschnitte an irgendeiner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

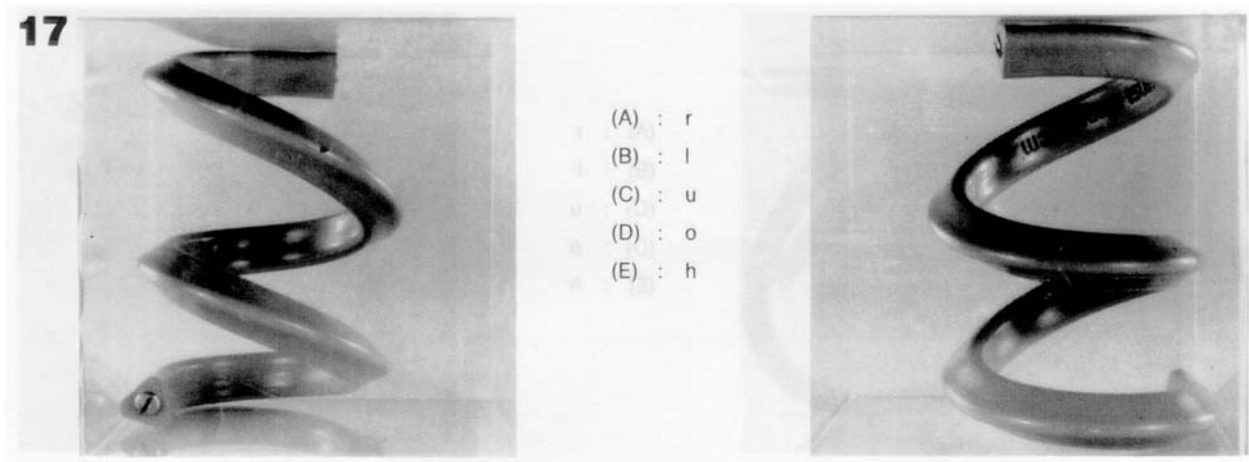
- I. Da bei der Mehrzahl der Linkshänder die Sprachregion in der linken Hemisphäre liegt, müssen sie also mit einer Sprachstörung rechnen, weshalb Aussage I falsch ist.
- II. Da es im Kindesalter noch offen ist, in welcher Hälfte des Gehirns die Sprachregion angelegt wird, besteht für ein Vorschulkind 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 Broteinheit (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 Querstreichen 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 – im Test-Info wird darauf hingewiesen, diesen Untertest vor der Testabnahme mehrfach zu üben.

b    $\overline{\overline{b}}$     $\overline{\underline{b}}$

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.

$\overline{\underline{b}}$     $\overline{\overline{q}}$     $\underline{\underline{b}}$     $\overline{\underline{b}}$     $\underline{\underline{b}}$     $\overline{\underline{b}}$     $\overline{\overline{b}}$     $\underline{\underline{q}}$     $\underline{\underline{b}}$     $\underline{\underline{b}}$     $\overline{\underline{b}}$     $\overline{\overline{q}}$

### 11.1.6 Untertest: Figuren lernen

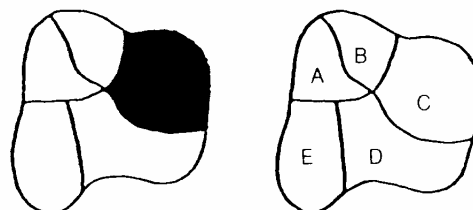
Für beide nachfolgende Gedächtnistests wird nach der Mittagspause das Material zum Einprägen ausgeteilt. Vor der Abfrage des Gelernten wird der Untertest „Textverständnis“ bearbeitet, damit liegt die Zeit des Behaltens der gelernten Inhalte über einer Stunde. Gedächtnisleistungen sind wichtige Voraussetzungen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

<b>Gesamt</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
Muster zuordnen	1			3	6	11	32	65	90	151	151	181	147	100	55	43	24	16	10	2	
Med.-nat. Grundv.	1	2	6	20	35	48	56	75	98	111	107	103	109	95	72	52	50	23	15	8	2
Schlauchfiguren				5	10	18	30	42	75	64	97	95	102	116	90	84	76	67	46	44	27
Qua.form Probl.		4	4	14	20	45	63	77	113	129	101	111	121	78	56	50	40	28	19	11	4
Textverständnis	2	9	8	19	37	51	74	103	100	104	107	88	79	64	67	67	47	41	21		
Figuren lernen		1	1	3	3	20	19	26	54	63	72	89	93	93	92	107	85	89	48	59	71
Fakten lernen		1	5	11	12	27	50	86	90	100	110	111	93	82	67	68	56	41	42	30	6
Diagr. und Tabellen		2	3	4	10	19	40	56	80	104	126	87	123	129	92	82	55	32	23	13	8
Kon.sorgf. Arbeiten						6	8	30	51	65	88	105	122	118	99	78	75	51	50	88	54

Tabelle A 1: Häufigkeitsverteilung der Punkte Gesamt 2003

<b>Deutsch</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
Muster zuordnen	1			1	6	7	26	61	76	134	138	158	129	94	54	42	22	16	10	2	
Med.-nat. Grundv.		2	5	15	30	44	49	66	85	93	97	93	103	89	64	48	47	23	15	7	2
Schlauchfiguren				4	10	15	23	35	66	54	83	80	92	102	83	77	73	65	44	44	27
Qua.form Probl.		4	3	12	19	36	53	66	103	115	88	100	111	71	53	48	38	27	19	7	4
Textverständnis	2	7	6	19	35	42	63	91	85	84	97	83	68	59	64	66	46	40	20		
Figuren lernen		1		3	3	19	15	21	47	53	59	74	83	86	80	96	75	86	47	59	70
Fakten lernen		1	5	10	11	24	42	75	76	88	98	101	84	74	57	61	55	39	41	29	6
Diagr. und Tabellen		2	3	3	9	16	35	41	63	90	117	78	109	122	86	75	53	32	23	13	7
Kon.sorgf. Arbeiten						6	7	25	44	49	72	95	108	98	88	76	73	48	47	87	54

Tabelle A 2: Häufigkeitsverteilung der Punkte deutsche Sprachgruppe 2003

<b>Franzö- sisch</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
Muster zuordnen				2		3	3	2	10	10	7	13	10	5	1	1	2				
Med.-nat. Grundv.	1			2	4	3	3	6	10	11	5	7	5	2	8	1	1				
Schlauch- figuren				1		1	5	5	5	7	6	10	8	7	3	5	2	2	2		
Qua.form Probl.				1	1	1	5	7	4	4	12	6	8	8	6	2		2	1		1
Textver- ständnis				1		1	7	9	6	9	14	4	3	8	5	1			1		
Figuren lernen							3	4	5	5	10	5	4	7	9	8	4	3	1		1
Fakten lernen						3	4	9	11	5	6	7	5	7	6	4	1	1			
Diagr. und Tabellen						2	5	8	9	10	6	7	11	1	4	4	2				
Kon.sorgf. Arbeiten							1	4	6	12	12	6	8	11	6				1	2	

Tabelle A 3: Häufigkeitsverteilung der Punkte französische Sprachgruppe 2003

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

Tabelle A 4: Häufigkeitsverteilung der Punkte italienische Sprachgruppe 2003



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					5	8	24	51	72	119	112	147	125	81	45	37	20	15	9	1	
Med.-nat. Grundv.		2	6	14	25	38	43	57	80	82	87	82	89	84	62	40	41	19	12	7	1
Schlauch- figuren				5	6	14	24	30	56	44	74	76	83	98	73	65	66	55	41	38	23
Qua.form Probl.		3	1	11	13	36	46	52	91	103	80	94	102	56	49	43	36	25	17	9	4
Textver- ständnis	2	7	5	14	30	37	60	82	80	75	83	75	62	51	59	53	39	38	19		
Figuren lernen		1	1	1	3	19	15	17	42	55	52	67	78	76	78	82	70	68	39	50	57
Fakten lernen		1	4	11	11	20	39	66	68	76	96	88	72	60	52	56	47	35	36	27	6
Diagr. und Tabellen		2	3	4	9	14	29	38	64	73	105	62	102	101	79	68	47	29	23	11	8
Kon.sorgf. Arbeiten						6	6	19	37	45	59	84	102	102	83	61	63	43	40	74	47

Tabelle A 5: Häufigkeitsverteilung der Punkte Humanmedizin 2003

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	1			3	1	3	8	14	18	32	39	34	22	19	10	6	4	1	1	1	
Med.-nat. Grundv.	1			6	10	10	13	18	18	29	20	21	20	11	10	12	9	4	3	1	1
Schlauch- figuren					4	4	6	12	19	20	23	19	19	18	17	19	10	12	5	6	4
Qua.form Probl.		1	3	3	7	9	17	25	22	26	21	17	19	22	7	7	4	3	2	2	
Textver- ständnis		2	3	5	7	14	14	21	20	29	24	13	17	13	8	14	8	3	2		
Figuren lernen				2		1	4	9	12	8	20	22	15	17	14	25	15	21	9	9	14
Fakten lernen			1		1	7	11	20	22	24	14	23	21	22	15	12	9	6	6	3	
Diagr. und Tabellen					1	5	11	18	16	31	21	25	21	28	13	14	8	3		2	
Kon.sorgf. Arbeiten							2	11	14	20	29	21	20	16	16	17	12	8	10	14	7

Tabelle A 6: Häufigkeitsverteilung der Punkte Veterinärmedizin 2003

	Humanmedizin		Veterinärmedizin			Humanmedizin		Veterinärmedizin	
	Punktwert	Testwert	Punktwert	Testwert		Punktwert	Testwert	Punktwert	Testwert
42	1				101	16	29	6	13
46	1				102	11	34	5	7
48	2				103	9	32	6	8
50				1	104	19	32	5	10
53	2				105	17	40	5	8
54	2			1	106	12	23	6	3
55	1			1	107	17	33	2	
58	1			1	108	18	23	5	6
60	2				109	16	23	5	7
61	3			1	110	15	25	3	2
62				3	111	13	22	6	
63	1			1	112	18	19	4	3
64	3				113	18	16	5	1
65	5			1	114	11	23	1	3
66	5				115	11	9	2	6
67	3				116	10	20	2	
68	5			2	117	14	13	1	3
69	3				118	16	10		2
70	2	1	1		119	16	6		
71	5			2	120	14	2	3	
72	7	1	2		121	10	3	5	1
73	2	2	1		122	12	3	2	1
74	6		2	1	123	11	2	3	
75	5	2	5		124	15	1	1	
76	6	3	2	2	125	9	2	1	
77	7		1		126	9			
78	8	1	4		127	13	1		
79	10	5	2	1	128	11	1	2	
80	8	1	4	5	129	8		1	
81	12	8	1		130	10		1	
82	4	8	4	1	131	6			
83	12	9	6	3	132	16		1	
84	11	12	5	2	133	7		2	
85	7	9	3	5	134	4		3	
86	18	11	4	4	135	5		3	
87	13	15	3	8	136	10			
88	12	16	6	6	137	10			
89	7	17	2	4	138	5		1	
90	17	24	4	9	139	8		2	
91	15	21	7	6	140	3		2	
92	16	28	3	10	141	7			
93	17	24	3	8	142	2			
94	15	29	4	11	143	4			
95	8	33	3	6	144	2			
96	16	25	3	7	146			1	
97	15	33	2	4	147	3			
98	18	32	3	8	148	2			
99	12	27	3	13	149	1		1	
100	19	27	6	9	150	2			
					153	1			
					154	2			
					159	1			
					161	1			

Tabelle A 7: Punkt- und Testwertverteilung nach Disziplinen 2003

	Punkt- wert	Test- wert	Mus- ter zu- ord- nen	Med.- naturw. Grundv.	Schlauch- figuren	Quantit. formale Proble- me	Text- verständ- nis	Figur- ren lernen	Fak- ten ler- nen	Dia- gramme und Tabellen	Konz. und sorgf. A.
Total											
Mittelwert	103.78	100.01	10.72	10.37	12.52	10.40	10.09	13.43	11.26	11.49	13.41
N	1088	1088	1088	1088	1088	1088	1088	1088	1088	1088	1088
Standardabw.	20.459	10.005	2.687	3.664	3.799	3.586	3.879	3.922	3.848	3.416	3.626
Kurtosis	-.368	-.378	.315	-.465	-.630	-.272	-.659	-.595	-.588	-.267	-.806
Skewness	-.048	-.052	.160	.001	-.058	.183	.061	-.212	.147	-.023	.116
Median	104.00	100.00	11.00	10.00	13.00	10.00	10.00	14.00	11.00	12.00	13.00
Deutsch											
Mittelwert	105.06	100.63	10.82	10.50	12.71	10.50	10.24	13.63	11.37	11.66	13.62
N	977	977	977	977	977	977	977	977	977	977	977
Standardabw.	20.527	10.035	2.696	3.655	3.820	3.581	3.926	3.945	3.894	3.415	3.655
Kurtosis	-.345	-.353	.255	-.497	-.636	-.298	-.719	-.593	-.608	-.211	-.831
Skewness	-.124	-.127	.189	-.005	-.106	.142	.014	-.262	.115	-.080	.042
Median	106.00	101.00	11.00	11.00	13.00	10.00	10.00	14.00	11.00	12.00	13.00
Französisch											
Mittelwert	92.61	94.59	9.88	9.16	10.90	9.59	8.70	11.94	10.14	9.97	10.90
N	69	69	69	69	69	69	69	69	69	69	69
Standardabw.	14.605	7.218	2.709	3.333	3.313	3.358	2.846	3.280	3.064	2.828	2.550
Kurtosis	-.266	-.328	.326	-.183	-.385	.131	-.040	-.733	-.976	-.685	.410
Skewness	.112	.100	-.285	-.193	.079	.203	.322	.040	.228	.328	.588
Median	92.00	94.00	10.00	9.00	11.00	9.00	9.00	12.00	10.00	10.00	10.00
Italienisch											
Mittelwert	92.45	94.43	9.79	9.40	10.64	9.45	8.79	11.33	10.50	10.00	12.55
N	42	42	42	42	42	42	42	42	42	42	42
Standardabw.	18.252	8.936	1.994	3.945	2.827	3.833	3.619	3.221	3.617	3.436	2.698
Kurtosis	1.193	1.138	-.332	-.217	-.667	.890	.693	.520	-.083	.403	-.044
Skewness	.455	.468	-.658	.231	-.202	1.032	.140	-.482	.307	.474	.274
Median	93.00	95.00	10.00	9.00	10.50	8.50	9.00	11.00	10.00	9.00	13.00
Humanmedizin											
Mittelwert	104.97	100.59	10.84	10.48	12.70	10.61	10.23	13.46	11.34	11.66	13.60
N	871	871	871	871	871	871	871	871	871	871	871
Standardabw.	20.582	10.061	2.659	3.638	3.797	3.582	3.919	3.928	3.927	3.486	3.590
Kurtosis	-.363	-.367	.046	-.469	-.593	-.297	-.692	-.569	-.606	-.221	-.747
Skewness	-.104	-.106	.245	-.037	-.112	.176	.040	-.228	.123	-.087	.066
Median	105.00	101.00	11.00	11.00	13.00	10.00	10.00	14.00	11.00	12.00	13.00
Veterinärmedizin											
Mittelwert	99.01	97.69	10.27	9.94	11.80	9.58	9.52	13.32	10.93	10.77	12.64
N	217	217	217	217	217	217	217	217	217	217	217
Standardabw.	19.280	9.451	2.756	3.744	3.727	3.489	3.672	3.905	3.501	3.028	3.675
Kurtosis	-.186	-.232	1.093	-.363	-.628	-.154	-.504	-.683	-.587	-.519	-.870
Skewness	.126	.121	-.104	.163	.148	.208	.107	-.148	.210	.144	.354
Median	100.00	98.00	10.00	10.00	12.00	9.00	9.00	13.00	11.00	11.00	12.00

**Tabelle A 8:** Verteilungskennwerte für Punktwert, Testwert und Untertest nach Sprachgruppen und Disziplinen 2003


## Das Zentrum für Testentwicklung und Diagnostik im Internet:

[www.unifr.ch/ztd](http://www.unifr.ch/ztd)

The screenshot shows a Microsoft Internet Explorer browser window. The address bar displays <http://www.unifr.ch/ztd/ems/welcome.htm>. The page content is as follows:

### Kandidatinnen & Kandidaten

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



### Forschung zum EMS

- deutsch
- français

### Information des ZTD für Teilnehmende am EMS

#### Eignungstest 2003 für das Medizinstudium am 4.Juli erfolgreich absolviert

[Vollständiger Statistischer Bericht 2003 hier](#)

[Rapport statistique 2003 ici](#)

Erneut ist wie seit 1999 ein Numerus clausus für Personen notwendig, die sich für ein Studium der Human- oder Veterinärmedizin an den Universitäten Basel, Bern, Fribourg oder Zürich angemeldet hatten.

Bei einer Anfängerkapazität der vier Universitäten von insgesamt 768 Studienplätzen (Humanmedizin: 598, Veterinärmedizin: 170) lagen der CRUS im Februar 2003 Anmeldungen von 1448 Bewerberinnen und Bewerbern (Humanmedizin: 1177, Veterinärmedizin 271) für beide Disziplinen vor.

1088 Personen (871 Humanmedizin, 217 Veterinärmedizin) haben davon den EMS vollständig und mit gültigem Ergebnis absolviert. 58 weitere Personen (47 Humanmedizin, 11 Veterinärmedizin) übernehmen ihr Testergebnis aus dem Vorjahr, diese Zahl hat ebenfalls zugenommen. 11 Personen haben am EMS teilgenommen, um später an eine Universität bzw. Disziplin zu wechseln, für die der Numerus clausus gilt.

Der Test fand am 4.7.2003 gleichzeitig an sieben Testorten in drei Sprachen (deutsch: Basel, Bern, Chur, St. Gallen, Zürich; französisch: Freiburg; italienisch: Bellinzona) statt.

Gegenwärtig werden die Antwortbelege im ZTD ausgewertet und Testwerte berechnet. Für alle Bewerberinnen und Bewerber werden auf der Grundlage der erreichten Ergebnisse Testbescheide erstellt und der CRUS bzw. den zulassenden Universitäten übergeben. Alle Testteilnehmerinnen und Testteilnehmer werden entsprechende Verfügungen von der Universität erhalten, bei der sie sich angemeldet haben. Diesen ist der Testbescheid mit den Ergebnissen im Eignungstest beigefügt.

[Hier finden Sie Hinweise](#) zum Ablauf. Ein statistischer Bericht folgt in Kürze, der ausführliche Evaluations-Bericht im Herbst 2003.

Beachten Sie auch die Hinweise zum kommerziellen Training des Tests, die wir in diesem Jahr präzisiert haben (siehe unten).

**Weitere Informationen des ZTD**

**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)**