

EMS • Eignungstest für das Medizinstudium in der Schweiz • 2002

*Bericht 8 über die Durchführung und Ergebnisse 2002
Trends und Schlussfolgerungen aus fünf Jahren Testanwendung*

avec un résumé en français

K.-D. Hänsgen und B. Spicher

Bericht 8 (2002)

Hänsgen, Klaus-Dieter; Spicher, Benjamin (2002).

EMS Eignungstest für das Medizinstudium in der Schweiz 2002

Berichte des Zentrums für Testentwicklung, Nummer 8, im Auftrag der Schweizerischen Universitätskonferenz (SUK)

Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik, 2002

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 2002:

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).

Inhaltsverzeichnis

1	ZUSAMMENFASSUNG	5
2	RÉSUMÉ	8
3	AKTUELLE DISKUSSIONEN ZUM EMS.....	11
4	NUMERUS CLAUSUS UND MEDIZINSTUDIUM	13
5	ANMELDUNG ZU MEDIZINSTUDIUM UND TEST 2002.....	15
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 2001	26
5.7	Erneute Testteilnahme 2002	26
5.8	Sprachgruppen.....	27
5.9	Alter und Maturitätsjahr	30
6	BESCHREIBUNG DES VERWENDETEN EIGNUNGSTESTS.....	31
6.1	Aufbau des Tests.....	31
6.2	Berechnung der Werte	33
7	TESTANWENDUNG IN DER SCHWEIZ 2002.....	34
7.1	Organisation der Testabnahme.....	34
7.2	Verteilungsprüfung	34
7.3	Vergleich der Testfassungen 1998 bis 2002.....	37
7.4	Äquivalenz der Sprachversionen	40
7.4.1	Sprachvergleich für die Untertests	42
7.4.2	Darstellung des Korrekturverfahrens	45
7.4.3	Leistungen in identischen Testaufgaben 1998 – 2002	59
7.5	Vergleichbarkeit der Testlokale	60
7.6	Vergleich der Geschlechter	62
7.7	Vergleich der Kantone	67
7.7.1	Vergleich 2002 innerhalb der Deutschschweiz	68
7.8	Vergleiche für Altersgruppen.....	71
7.9	Vergleiche nach Wunschuniversitäten	74
8	ERGEBNISSE ZUR TESTGÜTE	75
8.1	Zuverlässigkeit	75

8.2	Binnenstruktur	77
8.3	Item-Trennschärfen	84
8.4	Konzentriertes und sorgfältiges Arbeiten.....	88
9	ZUSAMMENFASSUNG KANTONE 1998 BIS 2002	92
10	LITERATUR.....	96
10.1	Originaltest zur Information und Vorbereitung.....	97
10.2	Frühere Berichte des ZTD.....	98
11	ANHANG ZUM BERICHT 2002	100
11.1	Beispielaufgaben für die Untertests.....	101
11.1.1	Untertest: Muster zuordnen.....	101
11.1.2	Untertest: Medizinisch-naturwissenschaftliches Grundverständnis	101
11.1.3	Untertest: Schlauchfiguren.....	103
11.1.4	Untertest: Quantitative und formale Probleme	103
11.1.5	Untertest: Konzentriertes und sorgfältiges Arbeiten.....	104
11.1.6	Untertest: Figuren lernen	104
11.1.7	Untertest: Fakten lernen	105
11.1.8	Untertest: Textverständnis	105
11.1.9	Untertest: Diagramme und Tabellen	106
11.2	Weitere Detailergebnisse	109

1 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2002 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Human- und Veterinärmedizin ab Wintersemester 2002/2003. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2001 herangezogen.

Die Ergebnisse sind nicht repräsentativ für Schlussfolgerungen hinsichtlich einer gesamthaften oder regionalen Evaluation von Bildungseinrichtungen und Bildungswegen.

Statistik

Im Jahr 2002 gilt – wie seit 1999 – ein Numerus Clausus für Personen, die sich für ein Studium der Human- oder Veterinärmedizin an den Universitäten Basel, Bern, Freiburg oder Zürich angemeldet haben. Für die Zahnmedizin ist weiterhin kein Numerus Clausus notwendig, weil die Anfängerkapazität dort ausreicht.

Bei einer Anfängerkapazität der vier Universitäten von insgesamt 833 Studienplätzen (Humanmedizin: 648, Veterinärmedizin: 185) lagen der CRUS im Februar 2002 Anmeldungen von **1334** Bewerberinnen und Bewerbern (Humanmedizin: 1080, Veterinärmedizin 254) für beide Disziplinen vor.

Insgesamt 1033 Personen haben sich bis zur gesetzten Frist im Mai 2002 bei der CRUS zum Test angemeldet, tatsächlich angetreten sind am Testtag 973 Personen. 60 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen – was mit den Vorjahren vergleichbar ist. **973 Personen haben somit den EMS vollständig und mit gültigem Ergebnis absolviert.** Dies ist ein neuer Teilnahmerecord.

Organisation und Ablauf

Der Test fand am 5.7.2002 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 keine Testabbrüche, Verwarnungen oder Testausschlüsse. Dies spricht für eine disziplinierte und motivierte Testteilnahme aller Kandidatinnen und Kandidaten.

Testgüte

Auch im fünften Jahr der Testanwendung kann die hohe Testqualität konstant beibehalten werden. Bekanntlich wird jedes Jahr eine neue Testform aus bereits umfangreich erprobten Aufgaben des deutschen Tests für Medizinische Studiengänge (TMS) zusammengestellt und in die drei Testsprachen adaptiert. Die faktorielle Struktur der Untertests bleibt identisch. Die Zuverlässigkeit des Testwertes ist mit 0.80 (Konsistenz) bzw. 0.91 (Stabilität) sehr hoch und im Bereich der Vorjahre. Einzelne Unterschiede in den Untertests wirken sich nicht zuverlässigkeitsmindernd aus.

Gleichbehandlung der Sprachgruppen

Aufgrund der unterschiedlichen Zusammensetzung der drei Sprachgruppen und der über die Jahre zu beobachtenden Dynamik sind Unterschiede in den mittleren Testwerten auch in diesem Jahr zu erwarten. Von besonderer Bedeutung bei der Bewertung der Unterschiede ist der nunmehr gesicherte Befund, dass in allen drei Sprachgruppen gleiche Testwerte eine gleiche Studieneignung bedeuten – die mittleren Testwerte für die Sprachgruppen nach dem Bestehen und Nichtbestehen der ersten Vorprüfung unterscheiden sich nicht.

Es gilt 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.

Ein Ausgleich war 2002 notwendig in vier (französisch) bzw. zwei (italienisch) der möglichen fünf Untertests. Maximal wurden pro Person nur 2.9 (französisch) bzw. 1.8 (italienisch) Punkte gutgeschrieben. Die insgesamt geringe Zahl der betroffenen Aufgaben spricht dafür, dass das Übersetzungsverfahren weiter verbessert werden konnte. Die abschliessende Kontrolle zeigt, dass die Unterschiede in den nicht sprachabhängigen Tests wiederum deutlich grösser ausfallen als in den sprachabhängigen Tests.

Es kann daher 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 unterscheiden sich die jüngsten Bewerber (geboren 1981-1985) mit einem mittleren Testwert von 100.8 nicht von den ältesten mit einer frühen Maturitätsprüfung (geboren 1954-1977) mit einem mittleren Testwert von 100.4. Die Gruppe der 1978-1980 Geborenen ist statistisch davon ebenfalls nicht signifikant zu unterscheiden (Testwert 99.4). Erneut weisen die 1954-1977 Geborenen mit später Matur einen geringeren mittleren Testwert auf (93.0). Der Trend der Vorjahre wird für diese Gruppe einmal mehr bestätigt.

Differenzierung nach Geschlecht

Der Anteil der Bewerberinnen für die Human- und Veterinärmedizin hat auch in diesem Jahr weiter zugenommen (1998: 55.2% - 2001: 61.5% - 2002: 68.5%). 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 1.3 - 1999: 1.5; 2000: 1.4). 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

Erneut hat die 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.

Evaluation: Prognose der Studieneignung

Die Ergebnisse zur Evaluation des EMS bleiben weiter gültig: Der Test regelt die Zulassung nach der Eignung zum Studium der Human- und Veterinärmedizin, indem diejenigen bevorzugt einen Studienplatz erhalten, die das Studium in kürzerer Zeit (nachgewiesen für die erste und zweite Vorprüfung) und mit besseren Leistungen (nachgewiesen für die erste Vorprüfung) absolvieren. Es kann davon ausgegangen werden, dass gleiche Testwerte in jeder Teilgruppe (nach Sprache, Alter, Geschlecht, Disziplin, Universität) eine gleiche Bestehenswahrscheinlichkeit der Prüfungen bedeuten. Wenn der Testwert für Eignung steht, dann wird die Forderung nach Chancengleichheit erfüllt: Gleiche Chancen bei gleicher Eignung.

Schlussfolgerungen

Die Session des EMS 2002 konnte erfolgreich durchgeführt werden. Im Vergleich zu den Vorjahren bleibt die Testqualität konstant.

Der Test erweist sich nach fünf Jahren Einsatz in der Schweiz weiterhin als geeignetes Instrument für die Zulassung. Mögliche eignungswirksame Änderungen der Studienanforderungen sollten genutzt werden, um den Test weiter daran anzupassen.

2 Résumé

Le présent rapport expose les résultats du test d'aptitudes pour les études de médecine en Suisse (AMS), un test nécessaire pour être admis en médecine humaine ou vétérinaire au semestre d'hiver 2002/2003 dans les Universités de Bâle, Berne, Fribourg ou Zurich. Certaines questions ont été clarifiées à l'aide de données comparatives des tests effectués de 1998 à 2001.

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

Comme depuis 1999, un numerus clausus a été appliqué en 2002 pour gérer l'admission aux études de médecine humaine ou vétérinaire aux Universités de Bâle, Berne, Fribourg et Zurich. A l'instar des années précédentes, cette mesure n'a pas été nécessaire en médecine dentaire puisque la capacité en première année était suffisante.

En effet, alors que la capacité des quatre universités en première année est de 833 places d'études au total (médecine humaine: 648, médecine vétérinaire: 185), la CRUS enregistrait **1334** inscriptions pour les deux disciplines (1080 pour la médecine humaine et 254 pour la médecine vétérinaire) en février 2002.

Au total, 1033 personnes se sont inscrites au test dans les délais (mai 2002) auprès de la CRUS: 973 personnes se sont présentées au test tandis que 60 personnes se sont retirées ou ne se sont pas présentées, une situation comparable aux années précédentes. Cette année est marquée par un nouveau record de participation avec **973 personnes qui ont passé l'AMS en obtenant un résultat valide.**

Organisation et déroulement

Le test s'est déroulé simultanément dans sept villes et en trois langues le 5 juillet 2002 (en allemand à Bâle, Berne, Coire, St-Gall et Zurich; en français à Fribourg; en italien à Bellinzone). Il faut noter qu'on peut choisir le lieu de test indépendamment du futur lieu d'études. Aucun incident n'est à signaler: ni abandon, ni avertissement, ni exclusion, signe que tous les participants étaient disciplinés et motivés.

Qualité du test

Il a été possible de maintenir le test à un haut niveau de qualité pour la 5^e année consécutive. Rappelons qu'un nouveau test est préparé chaque année à partir du vaste corpus d'exercices éprouvés que fournit le Test pour les études de médecines (TMS) de l'Allemagne et adapté dans les trois langues nationales. La structure factorielle des sous-tests reste identique. La fiabilité des résultats au test est très élevée: 0,80 (consistance) et 0,91 (stabilité) et se situe au même niveau que les années précédentes. Les quelques écarts constatés dans les sous-tests ne portent pas préjudice à la validité de l'ensemble.

Egalité de traitement entre groupes linguistiques

La composition différente des trois groupes linguistiques ainsi que la dynamique observée au fil des ans, laissent présager que les moyennes des résultats au test diffèrent d'un groupe linguistique à l'autre. Il ne faut cependant pas perdre de vue lors de l'évaluation de ces différences un résultat désormais acquis, à savoir qu'un même résultat au test signifie une même aptitude à étudier la médecine quelque soit le groupe linguistique, car il n'y a pas de différence entre les moyennes des groupes linguistiques répartis en fonction de la réussite au premier propédeutique.

Conformément à la décision que la Conférence universitaire suisse (CUS) a prise en 1999, on utilise une procédure de compensation (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, Mémorisation de faits, Compréhension de textes, Diagrammes et tableaux; et (2) que lorsqu'on y constate des moyennes différentes entre groupes linguistiques.

En 2002, il a fallu apporter un correctif dans quatre sous-tests en français et deux sous-tests en italien. Sa valeur a été au maximum de 2,9 points (tests francophones) et de 1,8 point (tests italophones) par personne. Le fait que cette procédure n'a été appliquée qu'à un nombre restreint de tâches indique qu'on a pu encore apporter des améliorations au niveau de la traduction. Le contrôle final met en évidence le fait 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 qu'il n'existe aucune différence au niveau des résultats due au test lui-même et qui désavantagerait le groupe francophone ou italoophone.

Différenciation en fonction de l'âge

Cette année, la moyenne des plus jeunes candidats (nés entre 1981 et 1985) est de 100,8. Elle n'est guère différente de celle des candidats plus âgés ayant passé tôt leur maturité (nés entre 1954 et 1977), qui est de 100,4. Le groupe des candidats nés entre 1978 et 1980 ne s'en distingue pas non plus de manière significative d'un point de vue statistique (moyenne: 99,4). Mais le groupe des plus âgés ayant passé tardivement leur maturité obtient une moyenne inférieure (93,0), la tendance observée les années précédentes se confirme donc une fois encore.

Différenciation en fonction du sexe

La proportion des candidates à la médecine humaine ou vétérinaire a encore augmenté cette année (1998: 55,2% - 2001: 61,5% - 2002: 68,5%). On ne peut dès lors pas raisonnablement 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 (différence de 1,3 contre 1,5 en 1999 et 1,4 en 2000). En effet, le fait que les moyennes au test pour les groupes répartis en fonction du 1^{er} 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 prétendant le contraire.

Différenciation par taux de maturité et par canton

Le taux 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).

Evaluation: pronostics quant à l'aptitude aux études

Les résultats de l'évaluation de l'AMS restent valables: le test règle l'admission aux études en fonction de l'aptitude à étudier la médecine humaine ou vétérinaire, dans le sens où les personnes qui obtiennent en priorité une place sont celles qui achèvent leurs études le plus vite (avéré pour les 1^{er} et 2^e propédeutiques) et le mieux (avéré pour le 1^{er} propédeutique). On peut dire que, dans chaque sous-groupe (défini par la langue, l'âge, le sexe, la discipline et l'université), un même résultat signifie un même pronostic quant aux chances de réussir l'examen et que, lorsque le résultat au test reflète l'aptitude, le principe de l'égalité de traitement est respecté: à aptitude égale, chance égale.

Conclusions

La session 2002 de l'AMS a été menée avec succès. La qualité du test demeure constante face aux années précédentes.

Le test a fait ses preuves en Suisse: après avoir été appliqué durant cinq années, il s'avère un instrument de sélection approprié pour l'admission aux études de médecine. Des modifications éventuelles des exigences devraient être mises à profit afin de poursuivre son adaptation.

3 Aktuelle Diskussionen zum EMS

Seit 1998 wird ein Eignungstest für das Medizinstudium (EMS) in der Schweiz als Zulassungskriterium für den Numerus Clausus (NC) eingesetzt. Die sogenannten "Doppelmaturitäten" in zahlreichen Kantonen (zwei Jahrgänge schliessen wegen der Ausbildungszeit-Verkürzung im gleichen Jahr ab) bewirkten 2001 und 2002 eine zusätzliche Erhöhung der Bewerbungszahlen. Durch zeitlich beschränkte Kapazitätserhöhungen konnte dies jedoch ausgeglichen werden.

Studien zur Evaluation des EMS haben ausreichend belegt, dass durch die Testwerte sowohl Studiendauer, als auch Studienqualität in der Schweiz sehr gut vorhergesagt werden können (vgl. Berichte 6 und 7 des ZTD sowie Hänsgen & Spicher 2002) – dies für alle Sprachgruppen, Geschlechter, Universitäten oder Disziplinen. Die Zulassung erfolgt damit wie von der Gesetzgebung gefordert unter Berücksichtigung der Eignung für das Studium.

Diskussionen zur Lage der Medizin in der Schweiz beschäftigen sich gegenwärtig vor allem mit zwei Fragen: der Kostenexplosion im Gesundheitswesen und einem befürchteten Ärztemangel. Beide Dinge widersprechen sich scheinbar: Die Schweiz hat eine im internationalen Massstab recht hohe und ansteigende Ärztedichte und mehr Ärzte bedeuten in der Regel auch höhere Kosten. Es sei denn, dass die Kosten pro Arzt oder Ärztin begrenzt werden könnten. Einerseits glaubt man also, vor allem in Krankenhäusern, vermehrt Ärztinnen und Ärzte zu benötigen – andererseits ist ziemlich wahrscheinlich, dass die Kostenexplosion weitergeht, wenn es nicht gelingt, die Zulassungen zur freien Praxistätigkeit zu begrenzen.

Es wäre aber sehr kurzsichtig, im Numerus Clausus die Ursache für eine befürchtete Abnahme von Absolventenzahlen zu sehen, denn nach bisher vorliegenden Zahlen wird sogar das Gegenteil erwartet: Ausbildungskapazitäten werden von den Kantonen festgelegt und diese berücksichtigen neben einem Bedarf vor allem die Kosten der Ausbildung, die der Kanton zu tragen bereit oder in der Lage ist. Die Zahl wird vom Bedarf also nur indirekt beeinflusst. Nach wie vor gehören die Kosten für einen Medizin-Studienplatz mit zu den höchsten und man sieht sich auch in anderen Fächern (etwa Psychologie) einer verstärkten Studienplatznachfrage ausgesetzt, die ausgeglichen werden muss. Vorhandene Ausbildungskapazitäten müssen deshalb effektiv genutzt werden.

Weil es nachweislich gelingt, mittels EMS weniger geeignete Personen vom Studium auszuschliessen, ist die Aussicht auf Studienerfolg bei den Zugelassenen grösser. Die vorhandenen Studienplätze werden optimal genutzt, die Studiendauer verkürzt sich und es sollten insgesamt mehr Absolventen abschliessen als ohne NC.

In Humanmedizin entsprach die Kapazität in den letzten Jahren immer ziemlich konstant 80% der Personen mit gültigem Testwert. Würde man alle Personen zum Medizinstudium zulassen, die sich angemeldet haben, würden auch solche mit niedrigem Testwert verstärkt zugelassen. Diese haben aber nachweislich eine unvermeidbar geringe Chance, das Studium überhaupt abzuschliessen. Es würden dadurch die Kapazitäten der Universitäten insgesamt mehr in Anspruch genommen, was sich auch negativ auf Studienzeit und Studienqualität der übrigen Studierenden auswirken dürfte. Die Einführung des NC erfolgte ja bekanntlich, weil an den Universitäten durch den Zwang, alle Personen zuzulassen, teilweise katastrophale Zustände herrschten.

Ein Problem ist viel mehr, dass mögliche Dissuasionseffekte jetzt reduziert werden könnten. Wenn sich mehr Personen für ein Medizinstudium bewerben, die mittels EMS vor

Studienaufnahme eine Rückmeldung über ihre Erfolgsaussichten erhalten (und damit auch ihre berufliche Zukunft sicherer planen können), würde die Eignung der Zugelassenen sich insgesamt erhöhen. Rückzüge aufgrund von Umleitungen werden durch Nachrückende aufgefüllt – deren Eignung ist aber nachweislich bereits deutlich geringer.

Immer noch wird über das Verhältnis von Studien- zu Berufseignung diskutiert. Auch hier ist Annäherung in Sicht. Zum einen wird seitens der wissenschaftlichen Eignungsdiagnostik weiter über Möglichkeiten nachgedacht, Motivationen, Einstellungen und Haltungen sowie Persönlichkeitsvoraussetzungen auch in Eignungstests zu erfassen. Die Erwartungen sind allerdings wenig optimistisch, weil alle für diese Eignungsmerkmale vorhandenen Diagnostik-Methoden sehr trainingsabhängig (niemand will eine kostenintensive Trainingszene) oder sehr aufwändig sind (mehrere geschulte Beurteiler, längere Beobachtungszeiträume).

Zum anderen ist auch deutlich geworden, dass der NC die Vielfalt der Berufsgerechten eben nicht einengt, weil er vergleichsweise mild ist. Es werden letztendlich nur diejenigen nicht zugelassen, die bereits sehr geringe Chancen haben, das Studium überhaupt zu absolvieren. Der Studienabschluss bleibt auch weiterhin eine Voraussetzung für die Berufsausübung – in allen Anforderungsmodellen zum Arztberuf spielen Wissen und Kenntnisse weiterhin eine wichtige Rolle.

Die meisten der Motivationen, Einstellungen und Haltungen zum Arztberuf entwickeln sich im übrigen erst während des Studiums durch die Auseinandersetzung mit den verschiedenen Anforderungen und Tätigkeiten, auch ein Vorpraktikum kann allenfalls eine grobe Orientierung sein. Ausserdem wird es kaum DIE Berufseignung geben, sie wird so vielfältig und facettenreich sein wie der Arztberuf selber.

Insgesamt liegt die aktuelle Strategie des NC weiterhin voll im Trend:

- Vorhandene Kapazitäten, die nicht unbegrenzt steigerbar sind, optimal ausnutzen und nicht überlasten – gute Studienbedingungen für alle schaffen (auch als Voraussetzung für Studienreformen);
- Verstärkte Zulassung von Personen, die in optimaler Zeit das Studium beenden und die Kapazitäten dadurch rechtzeitig wieder freigeben;
- Personen rechtzeitig signalisieren, dass ihre Chancen, ein Studium erfolgreich abzuschliessen, deutlich geringer sind und so auch dafür sorgen, dass diese Personen sich in ihrer Biografie anders und damit auch für sie selbst optimal orientieren können.

4 Numerus Clausus und Medizinstudium

Die Abbildung 1 verdeutlicht die Anmeldeverläufe 1994 bis 1997 (ohne NC) sowie 1998 bis 2002 (mit NC). Die letzte Säule zeigt die Kapazitäten und das 20%-Überschreitungskriterium, ab welchem ein NC zur Anwendung kommt. In diesem Jahr waren im Kanton Zürich Doppeljahrgänge der Maturitäten vorhanden, was sich allerdings nicht so deutlich wie erwartet zeigt (etwa wie aufgrund der letztjährigen Situation in Bern). Hier können mehrere Faktoren eine Rolle spielen: Z.B. dass viele Personen das Jahr des "Ansturmes" meiden wollen oder dass auf die Einführung der Reformstudiengänge gewartet wird.

Die Zahlen von Bern belegen in diesem Jahr, dass sich der Anstieg aufgrund von Doppelmaturitäten mindestens auf 2 Jahre verteilt und im zweiten Jahr sogar höher ist als im ersten. Bekanntlich war der Doppeljahrgang dort im Jahre 2001. Es wäre auch möglich, dass die Einführung der Reformstudiengänge sich auf die Nachfrage auswirkt. Da Bern in diesem Prozess schon relativ weit fortgeschritten ist, könnte dies für die sehr deutliche Zunahme der Bewerbungen verantwortlich sein.

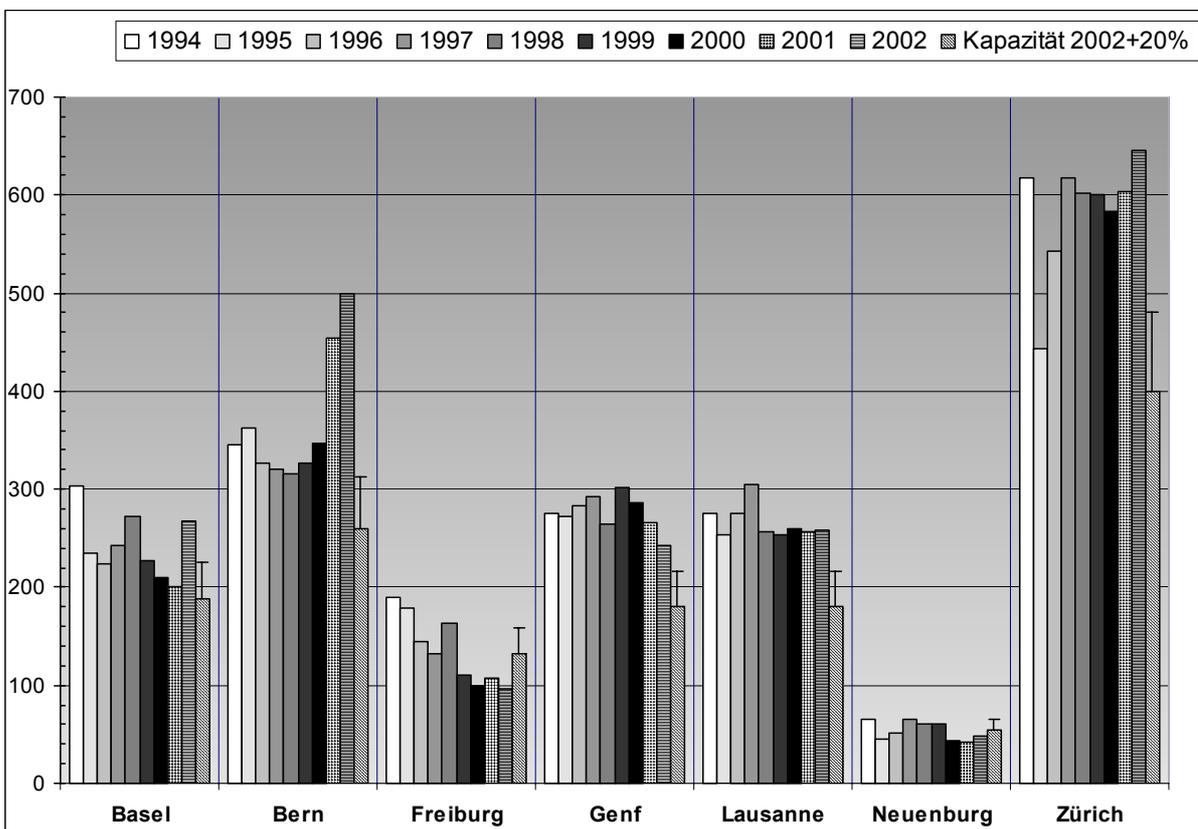


Abbildung 1: Anmeldungen zum Medizinstudium 1994 bis 2002, Anfängerprognose und Kapazitäten 2002 (letzte Säule – mit Markierung des 20%-Überschreitungskriteriums für 2002 als Grenze für die Notwendigkeit einer Zulassungsbeschränkung)

Unsere Ergebnisse bestätigen die vom BfS bereits publizierte Annahme, dass sich diese Doppeljahrgänge tatsächlich auf mehrere Jahre verteilen¹ - dass der Zeitpunkt der

¹ Bundesamt für Statistik (BfS) (2001). Doppelte Maturitätsjahrgänge und Zahl der Studienanfänger. Nur ein Fehlalarm? Vision – Das Schweizer Magazin für Wissenschaft und Innovation 2 S. 60-61.

Maturitätsprüfung für die Studienaufnahme entscheidend ist und die Wartezeiten zum Teil auch bei der Bewerbung bewusst so eingerichtet werden, dass ein Studienbeginn in den "überlasteten" Anfangszeiten vermieden wird.

Der Numerus Clausus hat wiederum nicht zu einer Abwanderung in die Westschweiz geführt (Abbildung 2). Die Bewerberzahl in der Westschweiz bleibt konstant. In den am NC beteiligten Universitäten ist der Anstieg vor allem auf die Situation in Bern und Zürich zurückzuführen. Auch in der Westschweiz wäre das 20%-Kriterium erfüllt. Hier erfolgt die Selektion allerdings intrauniversitär im Verlauf des ersten Studienjahres.

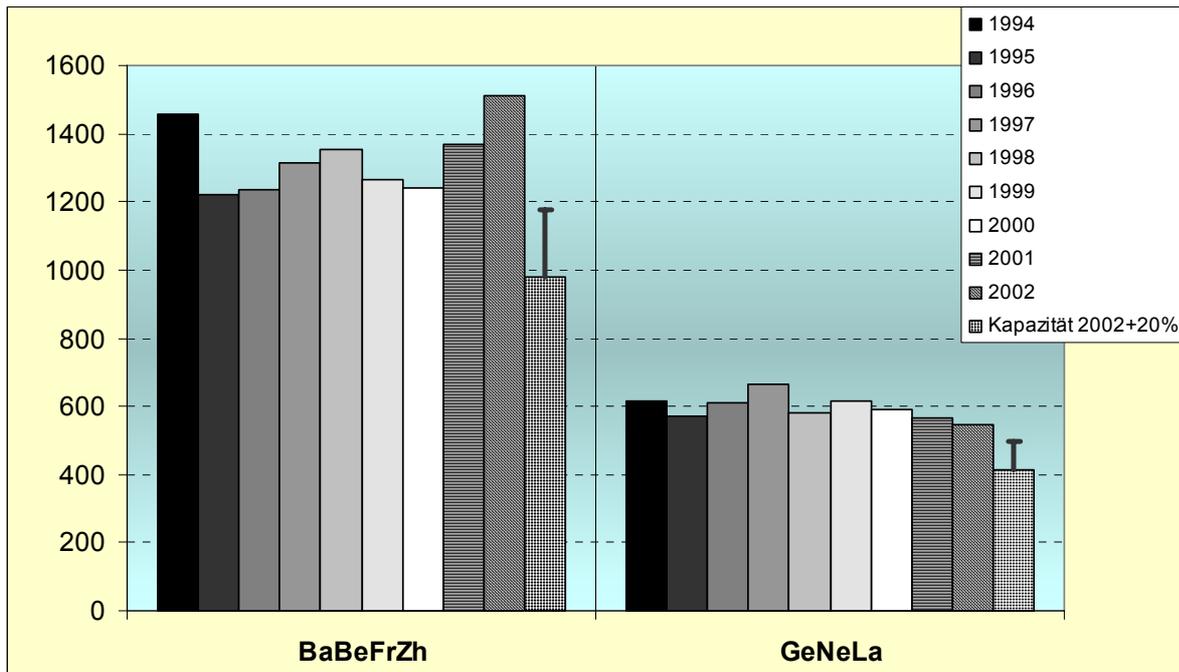


Abbildung 2: Anmeldungen zum Medizinstudium 1994 bis 2002 nach Gruppen (**B**asel, **B**ern, **F**reiburg, **Z**ürich vs. **G**enf, **N**euenburg, **L**ausanne)

5 Anmeldung zu Medizinstudium und Test 2002

5.1 Anmeldungen, Testantritte und Zulassungen

Tabelle 1 fasst den jährlichen Anmeldeverlauf bis zur Testabsolvierung zusammen. Im Fach Humanmedizin ist der Bewerbungsindex weiter sehr stabil (siehe auch Abbildung 3). Wie in den Vorjahren können etwa 80% der Testabsolventen einen Studienplatz erhalten. Seit dem Jahr 2000 hat auch die Zahl der Personen, die ihr Testergebnis aus dem Vorjahr übernehmen, eine beachtenswerte Grösse. Rückzüge in einem Jahr müssen also nicht bedeuten, dass kein Studium aufgenommen wird.

	Humanmedizin					Veterinärmedizin			
	1998	1999	2000	2001	2002	1999	2000	2001	2002
Kapazitäten	620	546	546	583	648	150	150	165	185
Anmeldungen Mitte Februar d.J.	1020	845	844	974	1080	231	213	217	254
... in % zu Kapazität	186%	155%	154%	167%	185%	153%	142%	132%	137%
+ Anmeldungen zum Test im Mai d.J.	797	717	678	752	827	194	176	179	206
+ Übernahme Test- ergebnis aus Vorjahr	0	5	33	27	40	0	5	5	3
= Total der Bewerbungen im Mai	797	722	711	779	867	194	181	184	209
... in % der Kapazität	129%	132%	130%	134%	134%	129%	121%	112%	113%
Absolvierung Test	750	673	637	709	770	182	164	171	203
Zuteilung von Plätzen an Personen mit gültigem Testwert	750	678	670	736	805	182	169	176	205
Rückzugsquote zwischen Februar und Zuteilung	26,5%	19,8%	20,6%	24,4%	25,5%	21,2%	20,7%	18,9%	19,3%
Personenanteil mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann	82,7%	80,5%	81,4%	79,2 %	80,5%	82,4%	88,7%	93,7%	90,2%
Bewerbungsindex Bewerber pro Studienplatz	1,21	1,24	1,23	1,25	1,24	1,21	1,13	1,06	1,11

Tabelle 1: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; NC 1998 nur für Humanmedizin – 1999 bis 2002 für Human- und Veterinärmedizin

Rückzüge zwischen Februar und Testtermin sind vermutlich darauf zurückzuführen, dass sich viele Personen bis zur Anmeldefrist im Februar noch alle Optionen offen halten wollten. Trotz aller Schwankungen der Anmeldezahlen und Kapazitäten pendelt der Prozentsatz der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann, in Humanmedizin um 80%. In Veterinärmedizin schwankt dieser Prozentsatz stärker (siehe Abbildung 3).

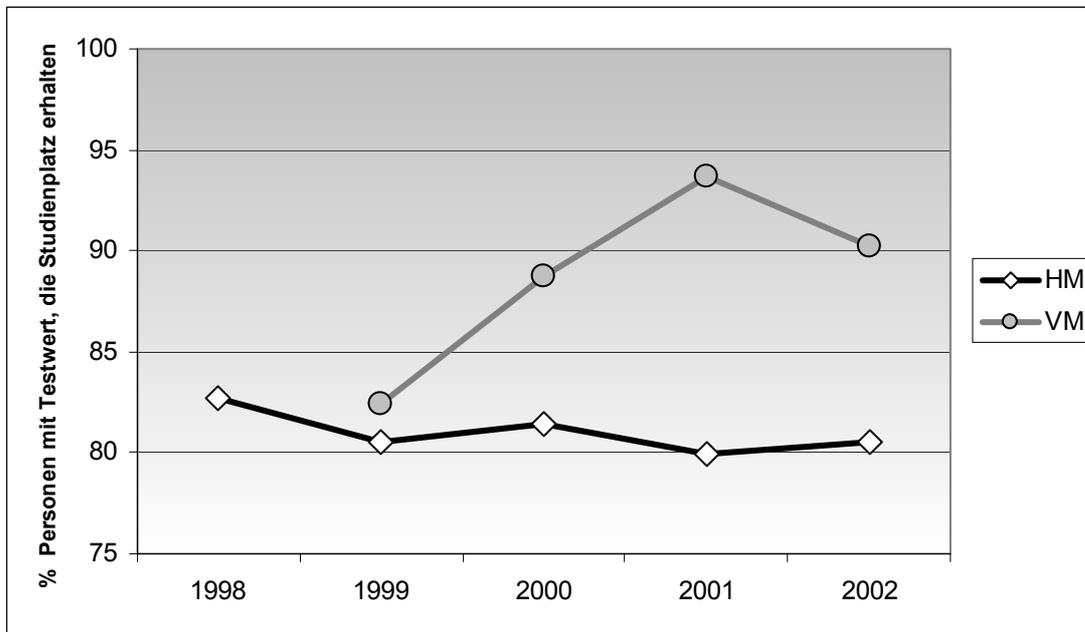


Abbildung 3: Prozent der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann – HM: Humanmedizin; VM: Veterinärmedizin 1998-2002

Die folgenden Abbildungen stellen Anmeldungen (Februar), Testanmeldungen und die Absolvierung des Tests im Verhältnis zur Kapazität für die bisherigen Testanwendungen dar.

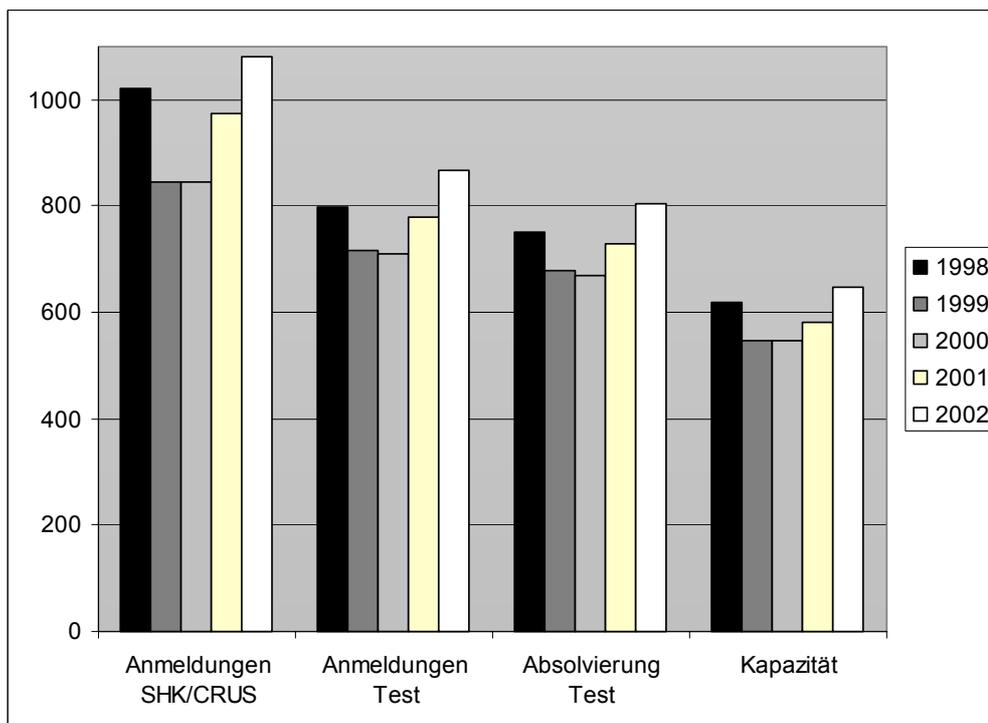


Abbildung 4: Anmeldeverlauf für Humanmedizin 1998 bis 2002 im Vergleich (mit NC)

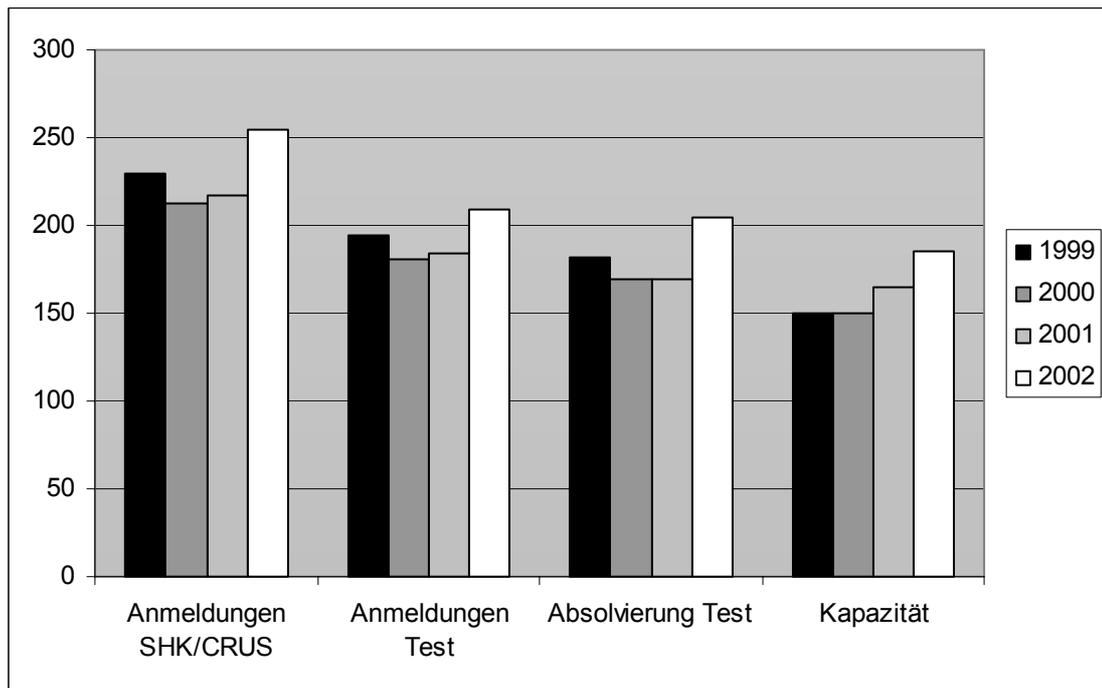


Abbildung 5: Anmeldeverlauf Veterinärmedizin 1999 bis 2002 im Vergleich (mit NC)

In beiden Disziplinen zeigen sich vergleichbare Verhältnisse über die einzelnen Jahre. Spezielle Dissuasionseffekte, die vom Jahr 1 zum Jahr 2 des NC aufgetreten sein können, sind ausgeglichen – möglicherweise auch durch die Doppelmaturitäten. In der Abbildung 6 werden die Anteile der einzelnen Gruppen bezogen auf die Gesamtzahl der Anmeldungen dargestellt.

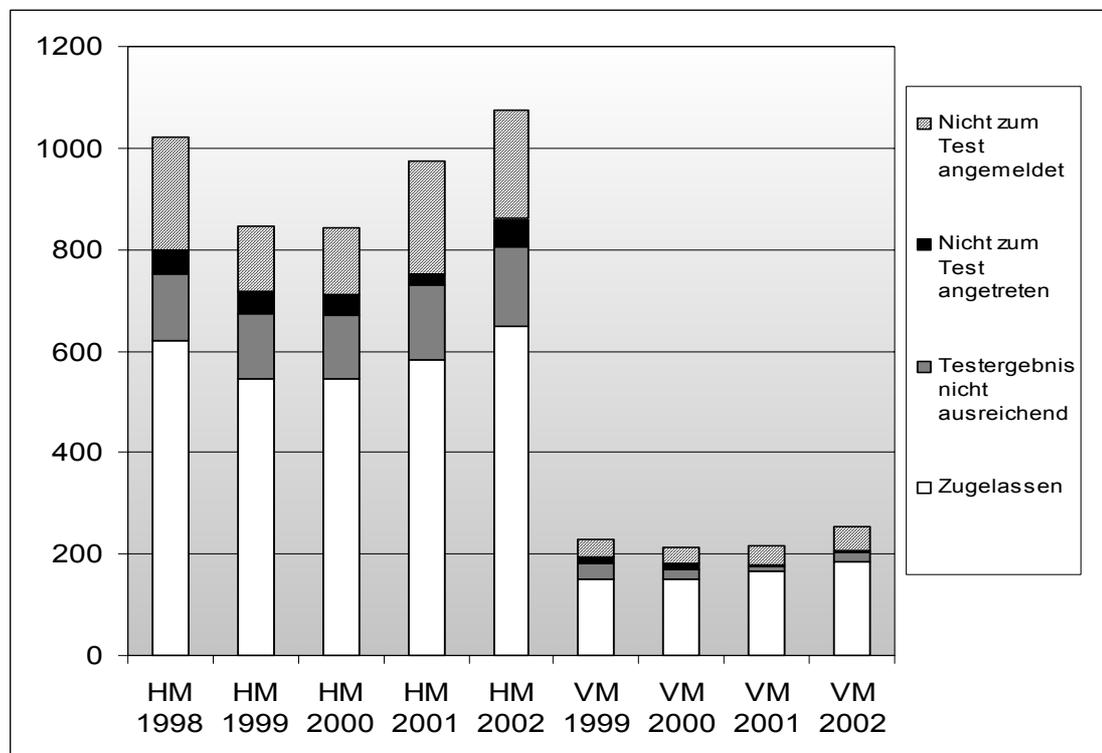


Abbildung 6: Kapazität, Anmeldungen und Testteilnahme für Human- (HM) und Veterinärmedizin (VM). Aufteilung der Voranmeldungen (Februar des jeweiligen Jahres)

In Tabelle 2 und Abbildung 7 (Quelle: CRUS) werden Rückzüge zwischen Anmeldung und Testabsolvierung differenziert. Die um ca. 5% erhöhten Rückzüge in den Jahren 2001 und 2002 begegnen uns in Personen wieder, die ihr Testergebnis ins Folgejahr übernehmen. Die Zunahme der Rückzüge 2002 um ein Prozent und der hohe Anteil in Zürich würden belegen, dass auch in 2003 mit einer grossen Zahl von Testübernahmen zu rechnen sein wird.

	Studienbeginn 1999			Studienbeginn 2000			Studienbeginn 2001			Studienbeginn 2002		
	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	189	151 (2)	20.1	174	133 (9)	23.6	175	136 (7)	22.3	239	181 (6)	24.3
Bern HM	164	132 (2)	19.5	177	144 (5)	18.6	282	232 (5)	17.7	305	245 (16)	19.7
Freiburg HM	98	79 (0)	19.4	88	66 (2)	25.0	97	69 (1)	28.9	87	66 (5)	24.1
Zürich HM	394	316 (1)	19.8	405	327 (17)	19.3	420	299 (14)	28.8	449	313 (13)	30.3
Total	845	678	19.8	844	670	20.6	974	736	24.4	1080	805	25.5
Bern VM	112	85 (0)	24.1	117	99 (0)	15.4	118	104 (4)	11.9	132	106 (2)	19.7
Zürich VM	119	97 (0)	18.5	96	70 (0)	21.7	99	72 (1)	27.3	122	99 (1)	18.9
Total	231	182	21.2	213	169	20.7	217	176	18.9	254	205	19.3

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

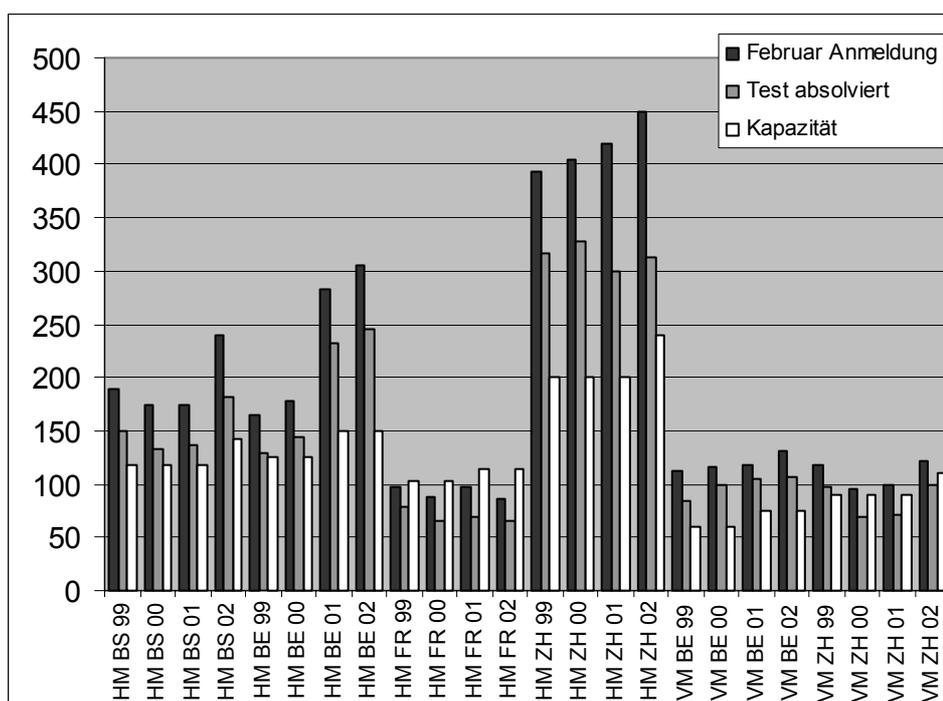


Abbildung 7: Kapazität, Anmeldungen und Testteilnahme für Fächer und Universitäten

5.2 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme in 14 Testlokalen in 7 Testorten (diese wie seit 1998). 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 unterschiedliche Raumgrössen sich auf die Ergebnisse auswirken. In den grossen Räumen ist der zur Verfügung stehende Platz teilweise grosszügiger (wenn es sich um grössere Hörsäle handelt).

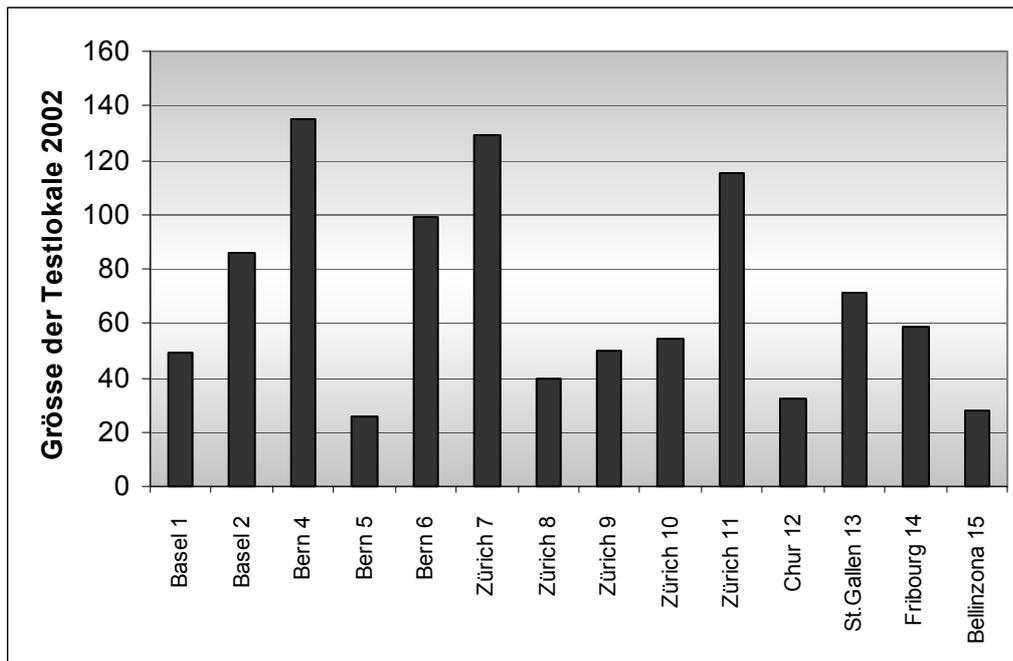


Abbildung 8: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2002

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 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	Voranmeldung Universität				Total
	Bern	Basel	Freiburg	Zürich	
Basel	10 7.4%	116 85.9%	1 0.7%	8 5.9%	135
Bellinzona	2 7.1%	3 10.7%	7 25.0%	16 57.1%	28
Bern	222 85.7%	14 5.4%	15 5.8%	8 3.1%	259
Chur	7 21.9%	5 15.6%	4 12.5%	16 50.0%	32
Freiburg	40 66.7%	1 1.7%	17 28.3%	2 3.3%	60
St. Gallen	13 18.3%	16 22.5%	5 7.0%	37 52.1%	71
Zürich	40 10.3%	22 5.7%	12 3.1%	314 80.9%	388
Total	334 34.3%	177 18.2%	61 6.3%	401 41.2%	973

Tabelle 3: Testorte und Wunschuniversität (erste Wahl) 2002

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

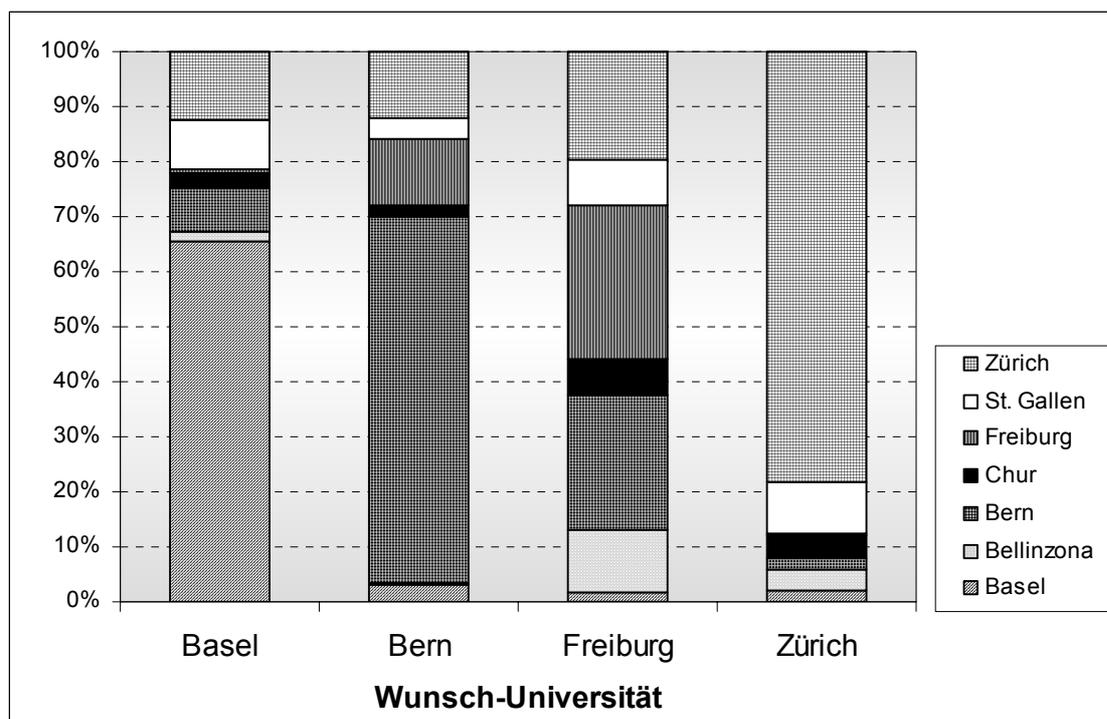


Abbildung 9: Aufteilung der Kandidaten nach Wunschuniversität auf die Testorte

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.

Wohnkanton /-ort	Testort 2002							Total				
	BASEL	BELLIN-ZONA	BERN	CHUR	FREI-BURG	ST. GALLEN	ZÜRICH	2002	2001	2000	1999	1998 (HM)
AG	25		2			1	32	<u>60</u>	49	71	84	66
AI						4	1	<u>5</u>	1	<u>4</u>	0	2
AR						6		<u>6</u>	10	4	6	4
BE			165		3		1	<u>169</u>	<u>186</u>	114	113	100
BL	41					1	2	<u>44</u>	45	40	59	50
BS	39							<u>39</u>	39	28	39	40
FR			15		20		1	<u>36</u>	27	44	47	34
GE					7			<u>7</u>	10	11	13	1
GL				1			3	<u>4</u>	3	5	4	4
GR	2	1		16			3	<u>22</u>	36	19	29	19
JU					2			<u>2</u>	1	1	2	0
LU	5		32	1			37	<u>75</u>	47	41	44	43
NE	1				6			<u>7</u>	3	6	5	2
NW	1		3				2	<u>6</u>	3	2	3	6
OW	2		1				4	<u>7</u>	3	<u>9</u>	2	3
SG			1	8		38	12	<u>59</u>	48	55	61	<u>64</u>
SH			1				9	<u>10</u>	<u>18</u>	8	4	7
SO	16		24		1		2	<u>43</u>	25	20	28	26
SZ							15	<u>15</u>	15	18	18	<u>21</u>
TG						21	13	<u>34</u>	38	23	20	22
TI		26					1	<u>27</u>	38	27	29	30
UR			2				3	<u>5</u>	5	2	1	3
VD					13		1	<u>14</u>	23	17	11	3
VS			11		8		1	<u>20</u>	19	18	15	22
ZG							19	<u>19</u>	<u>23</u>	9	10	6
ZH		1		1			222	<u>224</u>	148	194	191	164
FL				4				<u>4</u>	7	4	2	3
Übriges Ausland	3		2	1			4	<u>10</u>	10	11	17	8
	135	28	259	32	60	71	388	973	880	801	855	750

Tabelle 4: Wohnkanton und Testort, Vergleich zu 1998 bis 2002 (1998 nur Humanmedizin). Unterstrichen: Doppeljahrgänge in den Kantonen. Das Fürstentum Liechtenstein (FL) wird als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wegen der geringen Absolutzahl ist in kleineren Kantonen mit Doppeljahrgängen allenfalls ein Trend zur Zunahme zu erkennen. Auch in den anderen Kantonen finden sich starke Fluktuationen der Bewerberzahlen zwischen den Jahren. In diesem Jahr ist die Zunahme der Bewerberzahlen in allen Kantonen mit Doppelmaturität allerdings recht deutlich.

Die Präferenz des gewünschten Studienortes in Abhängigkeit vom Wohnort wird in der Tabelle 5 dargestellt. Die 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 2002				Total 2002	Total 2001	Total 2000	Total 1999
	Basel	Bern	Freiburg	Zürich				
AG	33	7	3	17	60	49	71	84
AI		2	1	2	5	1	4	0
AR	4	1		1	6	10	4	6
BE	5	158	4	2	169	186	114	113
BL	31	5	1	7	44	45	40	59
BS	34	4		1	39	39	28	39
FR		13	21	2	36	27	44	47
GE	1	6			7	10	11	13
GL	1	1		2	4	3	5	4
GR	5	3	1	13	22	36	19	29
JU		2			2	1	1	2
LU	9	35	7	24	75	47	41	44
NE	1	6			7	3	6	5
NW	2	3		1	6	3	2	3
OW	2	4		1	7	3	9	2
SG	9	8	4	38	59	48	55	61
SH	2	2	1	5	10	18	8	4
SO	20	18	1	4	43	25	20	28
SZ		4	1	10	15	15	18	18
TG	7	3	2	22	34	38	23	20
TI	3	2	7	15	27	38	27	29
UR		4		1	5	5	2	1
VD		12		2	14	23	17	11
VS		15	2	3	20	19	18	15
ZG	2	3		14	19	23	9	10
ZH	2	9	1	208	224	148	194	191
FL	1	2		1	4	7	4	2
Übriges Ausland	3	1	1	5	10	10	11	17
Total	177	333	62	401	973	880	801	855

Tabelle 5: Gewünschte Studienorte nach Wohnkantonen, Vergleich zu 1999 bis 2002. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Bezogen auf die Doppelmaturitäten kann festgehalten werden: Im Jahr 1 nach der Doppelmaturität sinkt die Zahl der Bewerber aus Bern geringfügig ab – der weitere Zuwachs der Bewerbungen in Bern rekrutiert sich aus anderen Kantonen. Hierfür können die bereits fortgeschrittenen Reformbemühungen eine Ursache sein. In Zürich ist eine Zunahme vorhanden, die jedoch relativ gesehen geringer ausfällt als die Zunahme in Bern von 2000 zu 2001.

Tabelle 6 zeigt, dass im Fach Veterinärmedizin mehr als eine Verdoppelung der Bewerbungen aus dem Kanton Zürich aufgetreten ist.

Wohnkanton/ Wohnort	Studienort		Total	
	Bern	Zürich	2002	2001
AG	3	5	8	7
AI	1	2	3	0
AR			0	2
BE	29	1	30	39
BL	3	4	7	4
BS	3	1	4	7
FR	9		9	3
GE	6		6	9
GL		1	1	0
GR		4	4	6
JU	2		2	1
LU	6	7	13	5
NE	6		6	3
NW	2		2	0
SG	1	8	9	7
SH			0	3
SO	7	1	8	4
SZ	2	2	4	0
TG	1	5	6	8
TI	2	7	9	7
UR	1		1	1
VD	12	1	13	21
VS	7		7	7
ZG		3	3	4
ZH	1	45	46	19
FL		1	1	2
Übriges Ausland		1	1	2
	104	99	203	171

Tabelle 6: Bewerbungen 2002 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 2002 in den Fächern Human- und Veterinärmedizin – mit weiter ansteigender Tendenz. In Humanmedizin sind es vor allem die Jüngeren, wo nunmehr das Verhältnis 1/3 Männer zu 2/3 Frauen herrscht. Bei den Älteren (die allerdings weniger Personen umfassen), ist das Geschlechterverhältnis ausgeglichener. Eine mögliche Erklärung wäre, dass Männer eher eine Wartezeit zwischen Schulabschluss und Studienbeginn einlegen. Es können allerdings auch unterschiedlich wirkende Dissuasionseffekte dafür verantwortlich sein. In Veterinärmedizin ist das Verhältnis bereits auf 4/5 Frauen (und mehr) zu 1/5 Männer (und weniger) angewachsen.

	Geburtsjahre			Total
	1954 – 1978	1979 – 1981	1982 – 1984	
Humanmedizin				
männlich	25 37.9%	100 41.8%	144 31.0%	269 34.9% von Gesamt 1998 44.8%
weiblich	41 49.0%	139 55.5%	321 65.5%	501 65.1% von Gesamt 1998 55.2%
Gesamt	66 8.5% v. Gesamt 1998 10.4%	239 31.0% v. Gesamt 1998 29.5%	465 60.4% v. Gesamt 1998 60.1%	770
Veterinärmedizin				
männlich	3 18.8%	13 24.1%	21 15.8%	37 18.2% von Gesamt 1999 24.2%
weiblich	13 81.3%	41 75.9%	112 84.2%	166 81.8% von Gesamt 1999 75.8%
Gesamt	16 7.9% v. Gesamt 1999 9.9%	54 26.6% v. Gesamt 1999 26.9%	133 65.5% v. Gesamt 1999 63.2%	203

Tabelle 7: Testabsolventen nach Geburtsjahr und Geschlecht 2002 und im Vergleich zum ersten Jahr des NC pro Disziplin

Die Abbildung 10 verdeutlicht, dass in Humanmedizin auch 2002 vor allem der Frauenanteil die Absolut-Unterschiede der Bewerberzahlen bedingt – die Zahl der Männer ist weiter geringfügig kleiner geworden. Die Anteile der Älteren sind 2002 mit 2001 vergleichbar (Abbildung 11), wobei auch zwischen 1999 und 2000 ein Rückgang bei den Männern zu verzeichnen war.

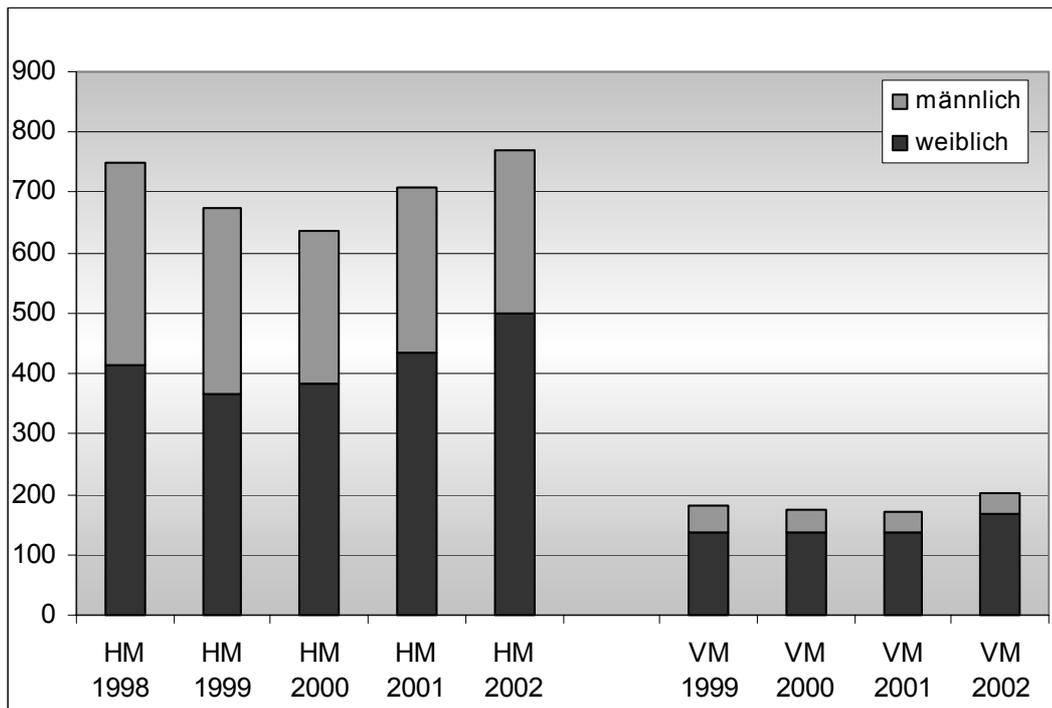


Abbildung 10: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM) 1998 bis 2002 und Veterinärmedizin (VM) 1999 bis 2002

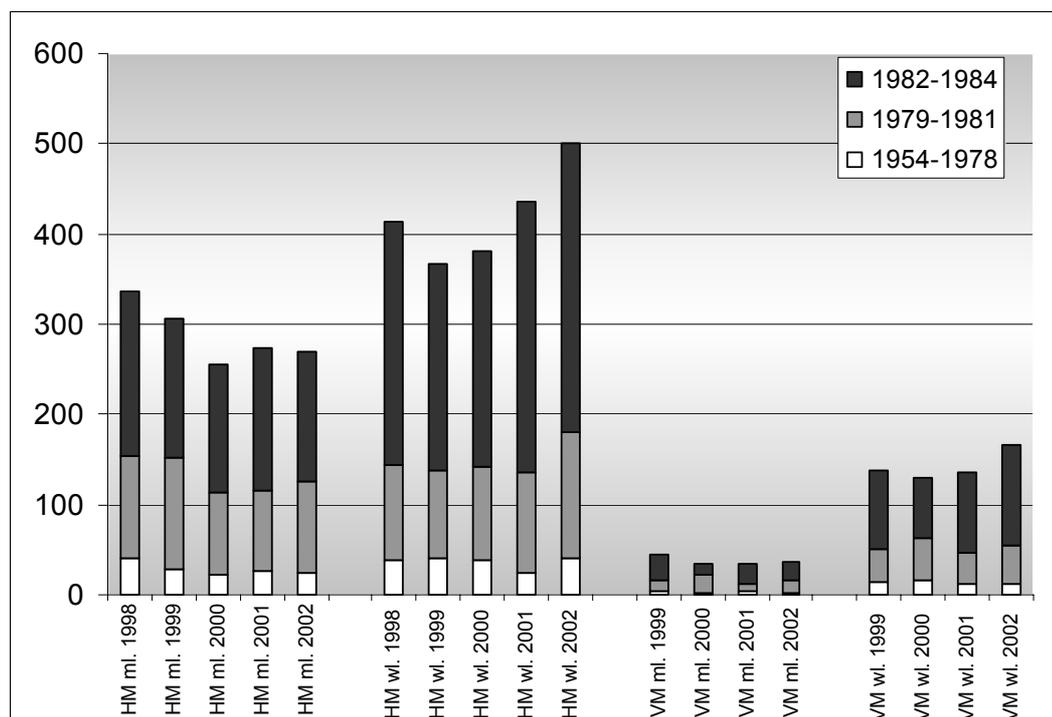


Abbildung 11: Zusammensetzung der Bewerbergruppen nach Geschlecht ml: männlich wl: weiblich und Alter (Geburtsjahre für Kohorte 2002 – 1998 bis 2001 jeweils entsprechend versetzt)

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.

5.6 Übernahme des Testergebnisses aus 2001

Im Jahr 2002 übernahmen 43 Personen ihr Testergebnis aus dem Vorjahr. Allen Personen konnte ein Studienplatz angeboten werden.

Disziplin	N	Minimum	Maximum	Mittelwert	Stand.-abw.
Humanmedizin 2000	33	85	119	102.85	7.25
Humanmedizin 2001	27	94	119	102.96	7.28
Humanmedizin 2002	40	83	118	104.03	9.16
Veterinärmedizin 2000	5	97	107	102.00	3.81
Veterinärmedizin 2001	5	89	112	102.80	9.52
Veterinärmedizin 2002	3	94	109	101.33	7.51

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

5.7 Erneute Testteilnahme 2002

Es ist auch 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.

Erste Testteilnahme	Zweite Testteilnahme			
	1999	2000	2001	2002
1998	12	2	1	1
1999		19	3	0
2000			12	0
2001				11

Tabelle 9: Wiederholte Testteilnahme: Anzahl der Personen

Insgesamt 61 Personen haben seit 1998 zweimal am Test teilgenommen. Die Mehrzahl davon (89%) traten bereits im Folgejahr noch einmal an. Dieser Personenkreis könnte das Testergebnis auch übernehmen.

Betrag der mittlere Testwert bei der ersten Testung 87.3, ist er bei der zweiten Testung 96.8 – also eine mittlere Verbesserung um den Betrag einer ganzen Standardabweichung. Die Korrelation zwischen beiden Messungen beträgt 0.66.

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.

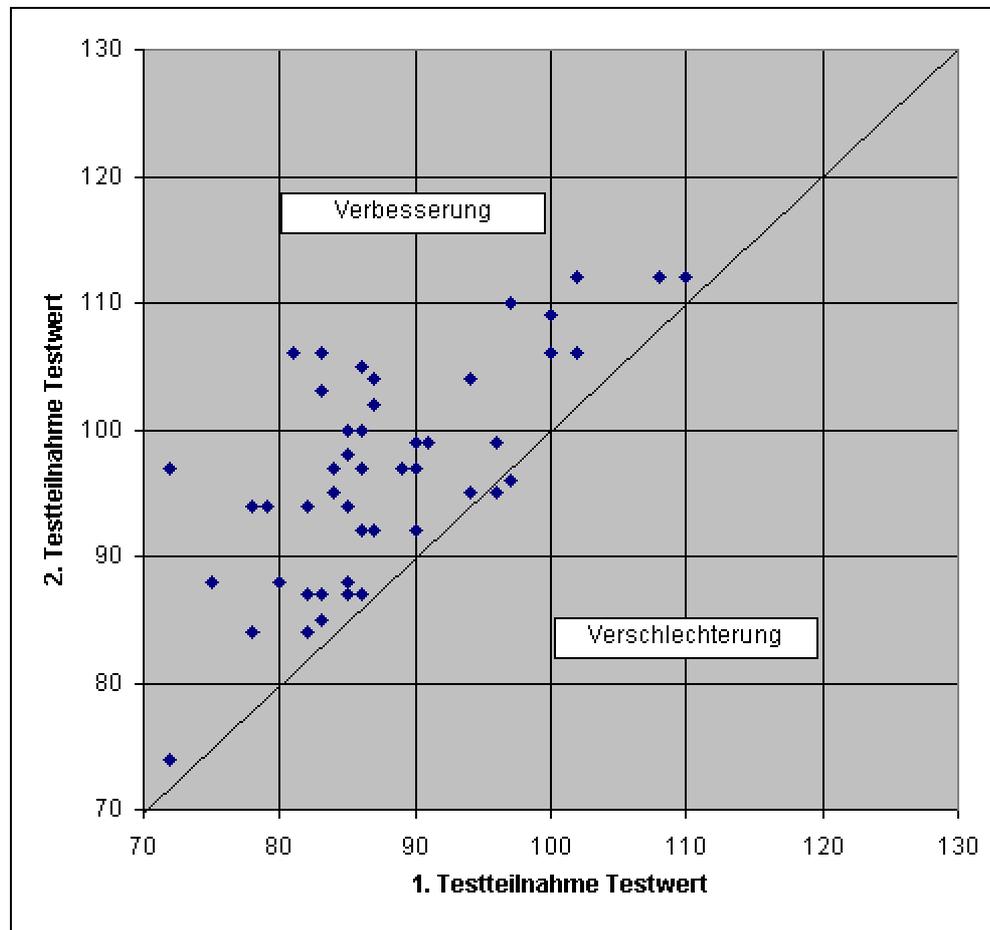


Abbildung 12: Testwerte für Personen, die zweimal am Test teilnahmen (1998 bis 2002)

5.8 Sprachgruppen

Die Zusammensetzung der Sprachgruppen unterlag in den letzten Jahren einer beständigen Dynamik: für die französisch- und italienischsprachigen Personen existieren Alternativen für ein Studium in der Romandie bzw. in Italien.

Insgesamt bleiben diese beiden Gruppen aber weiter klein – entsprechende Vorsicht ist bei der Bewertung aller statistischen Vergleiche geboten.

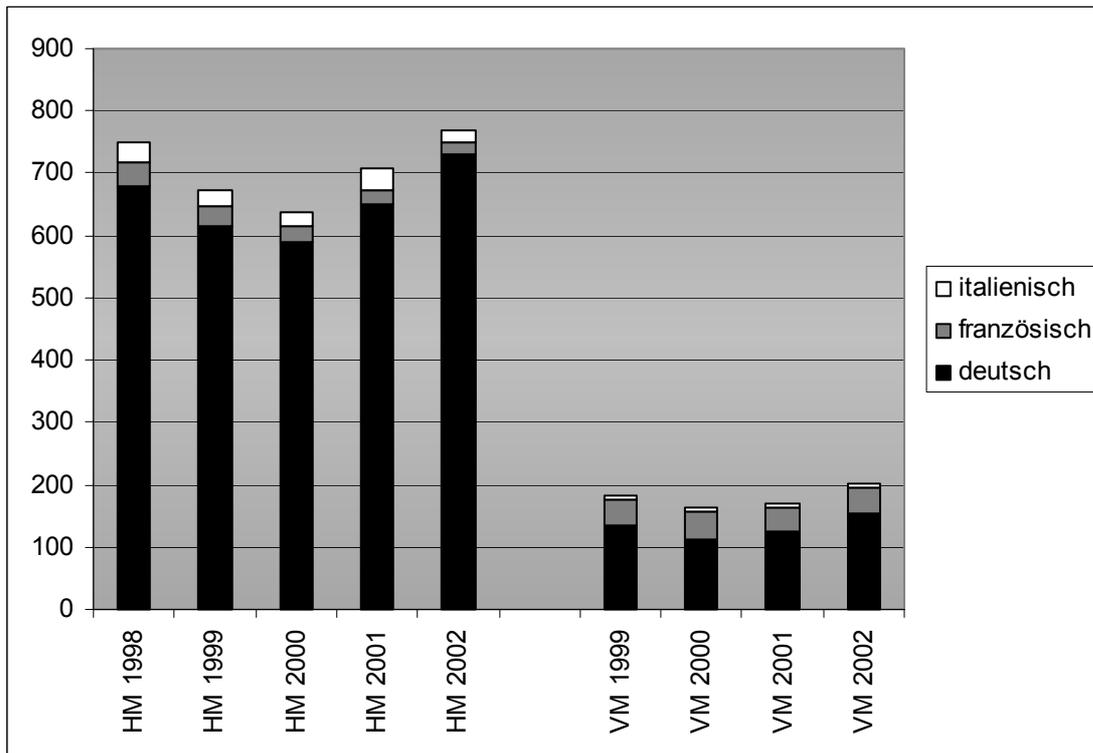


Abbildung 13: Vergleich der Sprachgruppen 1998 bis 2002, Absolutzahlen

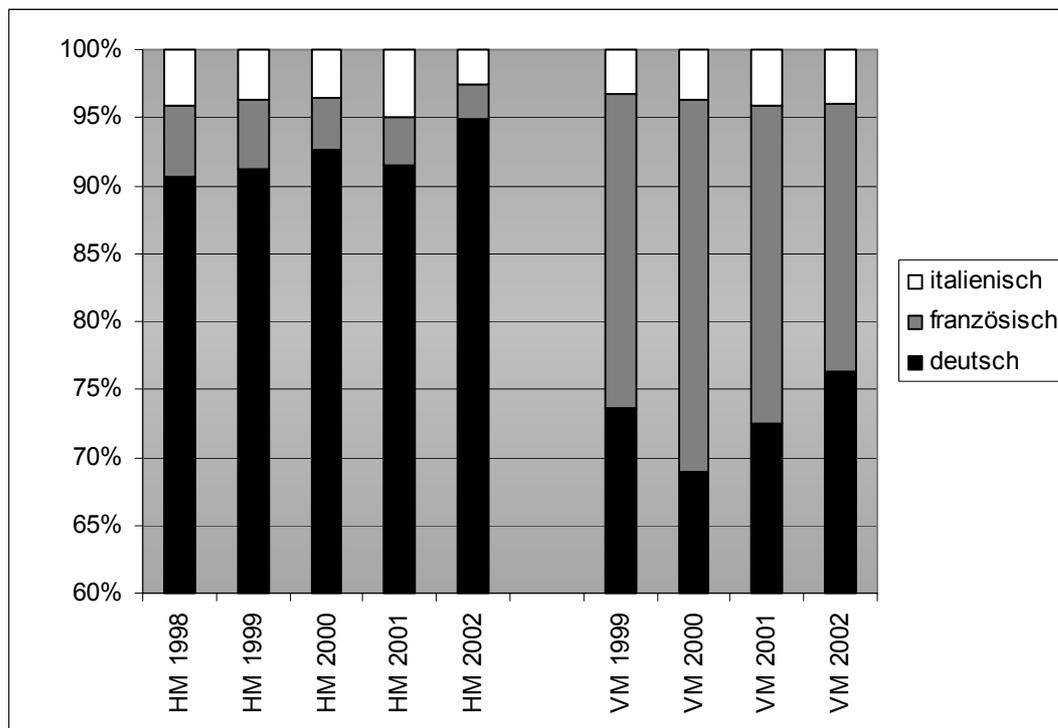


Abbildung 14: Vergleich der Sprachgruppen 1998 bis 2002, Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%)

Disziplin x Sprache x Geschlecht 2002			Klassen Geburtsjahr			
			1954-1978	1979-1981	1982-1985	
Humanmedizin	deutsch	männlich	23	94	129	
		weiblich	40	136	309	
	französisch	männlich	2	3	7	
		weiblich	1	1	5	
	italienisch	männlich		3	8	
		weiblich		2	7	
	Veterinärmedizin	deutsch	männlich	3	9	10
			weiblich	12	38	83
französisch		männlich		4	11	
		weiblich		1	24	
italienisch		männlich				
		weiblich	1	2	5	

Tabelle 10: Sprachgruppen 2002 nach Alter und Geschlecht aufgeschlüsselt

5.9 Alter und Maturitätsjahr

In diesem Jahr tritt wegen der Doppelmaturität erneut eine Häufung jüngerer Jahrgänge auf. Wir teilen später die Personen analog der Vorjahre zu Vergleichszwecken in drei Altersgruppen auf. 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																	Total
	74	83	84	86	89	90	91	92	94	95	96	97	98	99	00	01	02	
1954	1																	1
1959																	1	1
1961																	1	1
1962																	1	1
1963		1																1
1966			1	1														2
1967														1			1	2
1968								1										1
1969				1	1													2
1970					1												1	2
1971						1	1										3	5
1972									1			1					2	4
1973										1					1	2	2	6
1975									2		1			1	1	2	1	8
1976										5	2	1		1			6	15
1977											1	3	2	1			5	12
1978										2	1	2	1	5	2		4	17
1979											1	2	13	6	6		5	33
1980														4	16	28	26	74
1981														1	15	80	90	186
1982															1	65	213	279
1983																20	244	264
1984																1	54	55
	1	1	1	2	2	1	1	1	3	1	8	6	8	23	47	206	660	972

Tabelle 11: Geburtsjahr und Maturitätsjahr

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 12 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Untertests. Aufgabenbeispiele finden Sie im Anhang ab Seite 101.

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
MITTAGSPAUSE 60 Minuten			
Lernphase zu den Gedächtnistests: Figuren lernen Fakten lernen			4 6
Textverständnis	Verständnis und Interpretation medizinischer und naturwissensch. Texte	24 (18)*	60
Reproduktionsphase Gedächtnistests: Figuren lernen Fakten lernen	Behalten von figuralem Material Behalten von verbalem Material	20 20	5 7
Diagramme und Tabellen	Interpretation von Diagrammen und Tabellen	24 (20)*	60
Gesamttest		204 (178)*	5 Std. 7 Min.

Tabelle 12: 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.

Seine Vorteile 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 erfolgt¹);
- Ökonomisches und genau kapazitätsentsprechendes Zulassungsverfahren möglich, die Rechtsfähigkeit des Verfahrens hat sich bei mehreren Überprüfungen (auch gerichtlich) bestätigt.

Wie wurden die Aufgaben entwickelt? Ausgangspunkt war eine differenzierte Anforderungsanalyse des Medizinstudiums, an der zahlreiche Lehrbeauftragte und Experten mitarbeiteten. Das Ziel war die Vorhersage des Studienerfolges, um ein faires und wissenschaftlich begründetes Zulassungsverfahren zu erhalten. Gemäss diesen Anforderungen wurden einzelne Aufgabengruppen (Untertests) konstruiert. Neun davon erfüllten alle notwendigen Anforderungen – das Resultat ist also bereits eine Auswahl bewährter Aufgabentypen aus mehreren möglichen Alternativen. Jedes Jahr wurden neue Aufgaben für die Untertests entwickelt und in mehreren Schritten überarbeitet. 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 12) erfolgte in Deutschland im Rahmen sogenannter "**Einstreuaufgaben**". Nur bei ausreichender Bewährung wurden solche Aufgaben in nachfolgenden Testversionen für die Werteberechnung verwendet. Vier neue Aufgaben pro Untertest wurden in jeder Testform probeweise mitbearbeitet – ihr Ergebnis wurde jedoch nicht gezählt. Da in Deutschland acht verschiedene Testformen bei jedem Durchführungstermin zusammengestellt worden sind, konnten jeweils 32 neue Aufgaben pro Untertest an ausreichend grossen Stichproben erprobt werden. Maximal 20 davon wurden pro Jahr gebraucht – dieser Überschuss war nach Aussagen der Entwickler auch notwendig, da nicht alle Aufgaben die Kriterien zufriedenstellend erfüllten. Es versteht sich, dass die Einstreuaufgaben nicht besonders gekennzeichnet worden sind – jede Aufgabe des Tests konnte eine solche sein.

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

6.2 Berechnung der Werte

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

Beim Test "Konzentriertes und sorgfältiges Arbeiten" müssen insgesamt 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 15). 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.

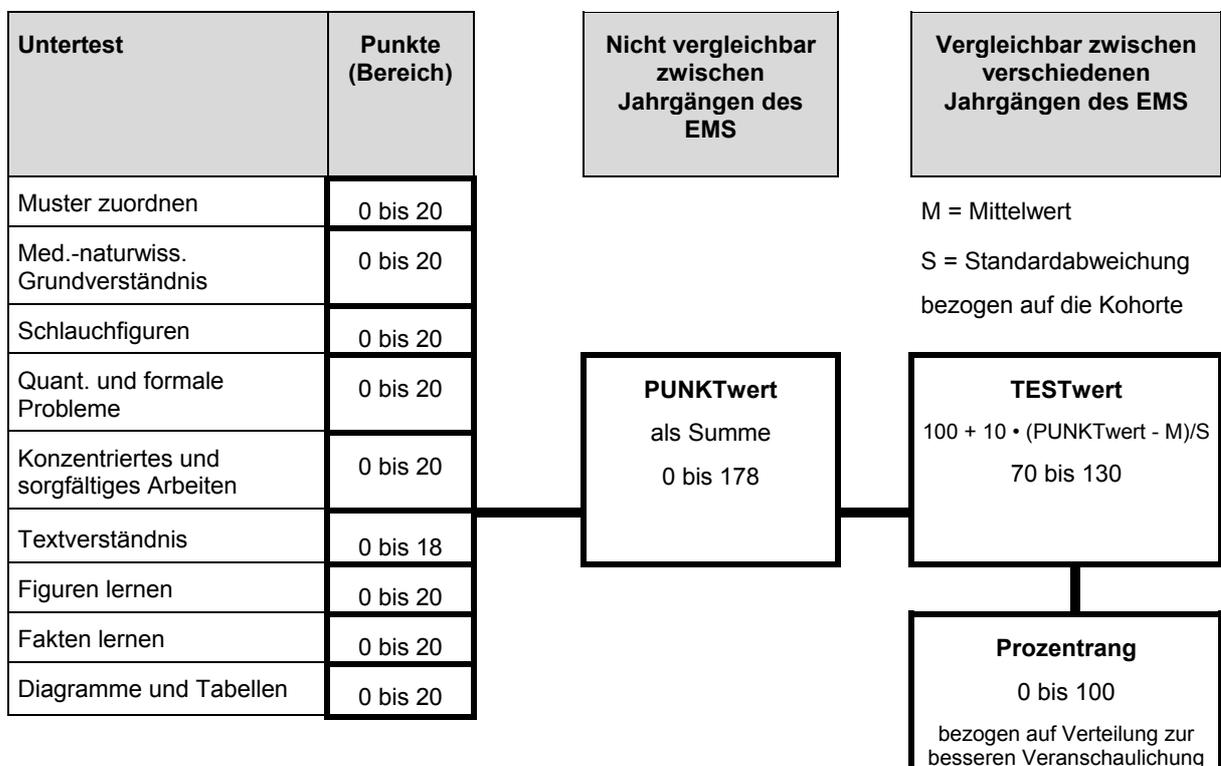


Abbildung 15: Punktwerte der einzelnen Untertests und ihre Zusammenführung über den Punktwert zum Testwert

7 Testanwendung in der Schweiz 2002

7.1 Organisation der Testabnahme

Mit dem Jahrgang 2002 wurde der EMS bereits zum fünften Mal in der Schweiz eingesetzt. Erneut konnte auf die bewährten Strukturen zurückgegriffen werden. Unter Testleitern und Assistenten kam es zu einigen wenigen personellen Veränderungen – allerdings wurden diese so gestaltet, dass weiterhin in jedem Testlokal erfahrenes Personal anwesend war.

In diesem Jahr sind keine Testabbrüche aufgetreten. Alle Personen, die den Test begannen, konnten ihn auch mit einem gültigen Testergebnis beenden. Alle Testabläufe wurden genauestens protokolliert – beispielsweise auch die genaue Einhaltung der Zeiten. Aus den Protokollen ergeben sich keine Hinweise darauf, dass irreguläre Bedingungen vorhanden gewesen wären. Die Testabnahme kann daher in allen Testlokalen als uneingeschränkt gültig angesehen werden.

7.2 Verteilungsprüfung

Der Testwert ist erneut normalverteilt (Tabelle 13), eine parametrische Prüfung ist daher zulässig. Die Punktwerte der Einzeltests sind, trotz der optischen Ähnlichkeit (Abbildung 16 und Abbildung 17) mit der eingezeichneten Normalverteilung von der Normalverteilung unterschiedlich.

Besonders auffällig sind in diesem Jahr die Punktwerte für den Untertest "Figuren lernen". Eine grosse Anzahl Teilnehmer hat dort die maximal mögliche Wertung erreicht. Eine Kontrolle dieses Untertests nach Testformen, Testorten und Testlokalen hat gezeigt, dass dieser Effekt überall auftritt. Offenbar war der Untertest in diesem Jahr leichter als in anderen Jahren, so dass hier ein Deckeneffekt zum Tragen kam. Da der Eignungstest insgesamt eine hohe Zulassungsquote hat und deshalb die Kandidaten in erster Linie im unteren Leistungsbereich zuverlässig differenzieren muss, sind von einem solchen Einzelfall keine wichtigen negativen Effekte zu erwarten (vorausgesetzt der Untertest zeigt eine gute Trennschärfe bezüglich des Gesamtprofils).

	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	103.01	11.21	9.78	11.98	11.06	9.13	15.39	10.55	10.71	13.21
Stand.- abweich.	19.53	2.86	3.13	3.50	3.73	3.40	3.88	3.74	3.15	3.51
extremste Differenz	0.02	0.07	0.08	0.07	0.08	0.10	0.12	0.09	0.08	0.07
Kolmog.- Smirnov Z	0.74	2.31	2.37	2.14	2.45	2.97	3.73	2.85	2.44	2.08
Asymp. Sig.	0.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tabelle 13: Verteilungsprüfung auf Normalverteilung (Kolmogorov-Smirnov Test)

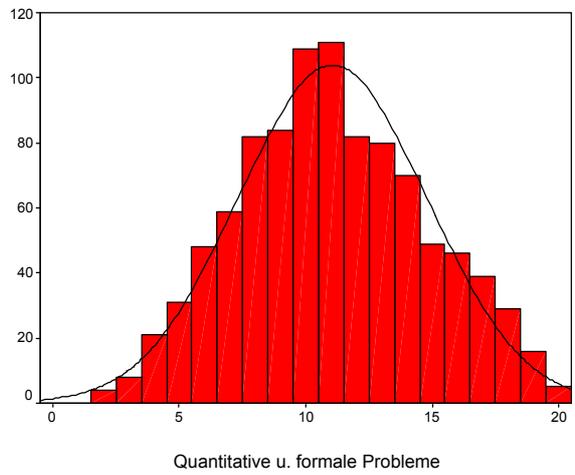
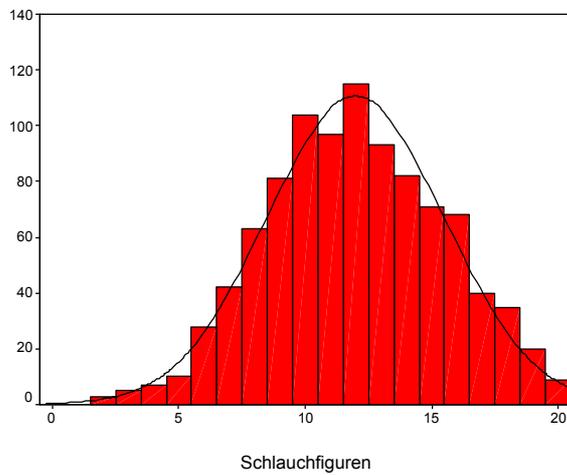
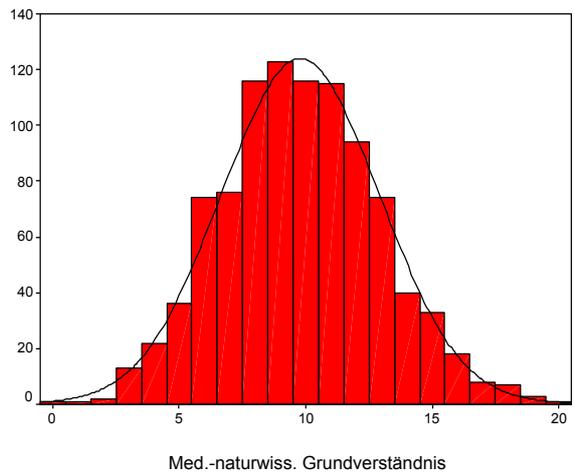
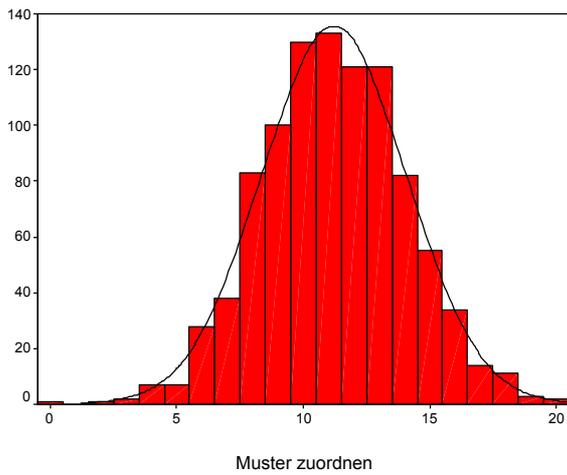
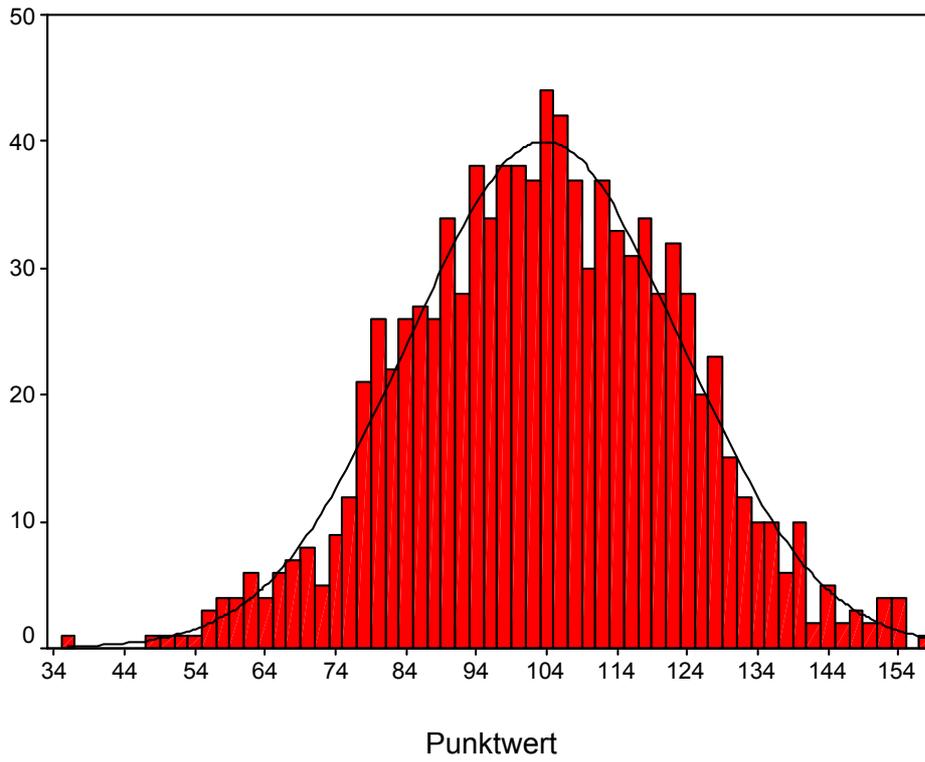


Abbildung 16: Häufigkeitsverteilungen für Punktwert und Punkte der Untertests 1-4

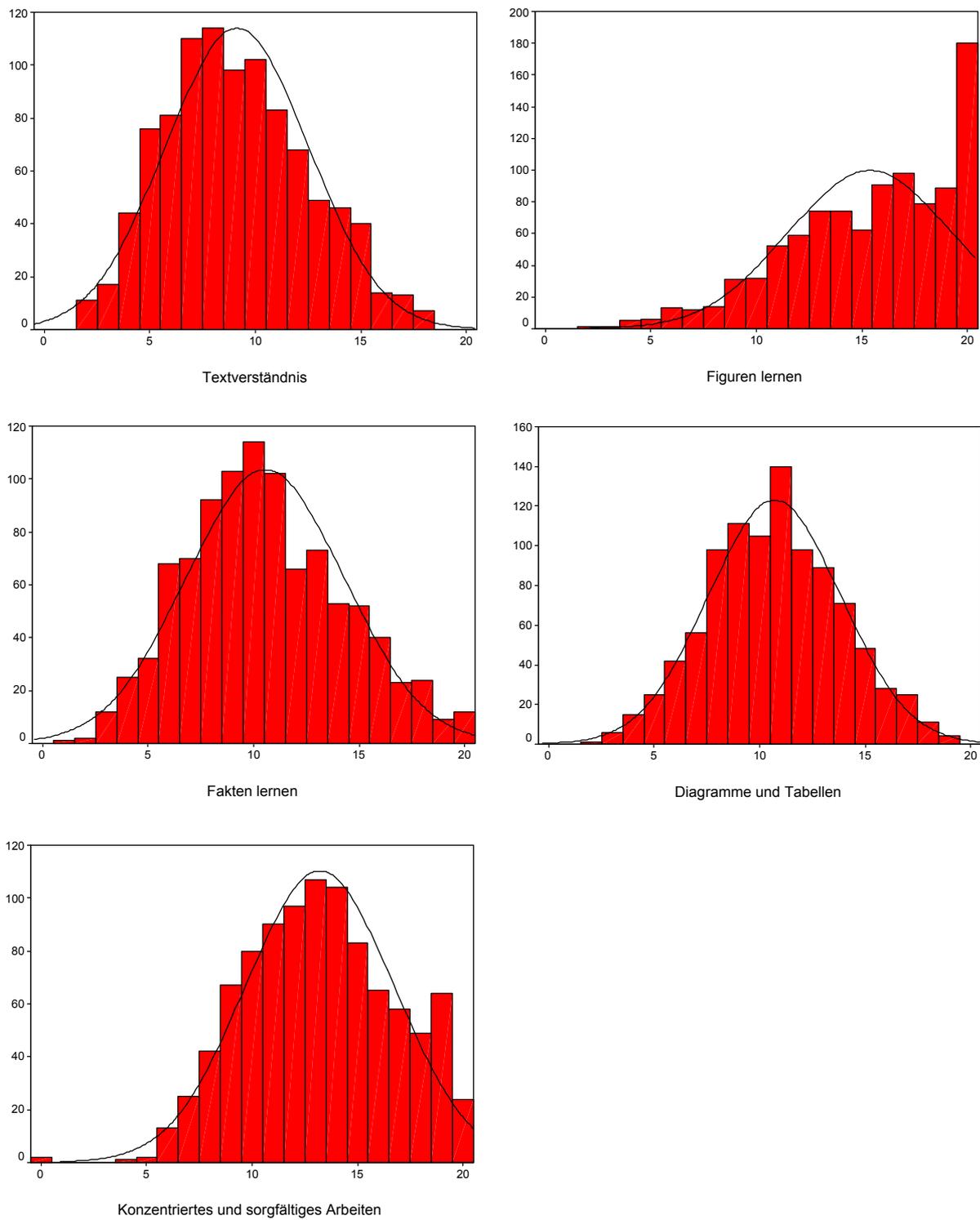


Abbildung 17: Häufigkeitsverteilungen für Punkte der Untertests 5-9

7.3 Vergleich der Testfassungen 1998 bis 2002

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. Da jedes Jahr nahezu identische Gütekriterien ermittelt werden konnten, wird dies 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 14 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.

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

Tabelle 14: Mittelwerte (Punkte) der Untertests 1998 bis 2002 deutsche Sprachgruppe Schweiz und Schätzung aufgrund der Item-Schwierigkeitswerte aus Deutschland (HM: Humanmedizin)

Erstmals seit der ersten Testdurchführung kann in diesem Jahr ein Anstieg der Gesamtpunktzahlen beobachtet werden, wobei die hohen Ergebnisse im Untertest "Figuren lernen" eine wichtige Rolle spielen. Aber auch in den Untertests "Muster zuordnen", "Quantitative und formale Probleme", "Textverständnis" sowie "Diagramme und Tabellen" sind leichte Anstiege zu verzeichnen.

Der Vergleich mit den Ergebnissen aus Deutschland spricht ebenfalls dafür, dass der Test 2002 etwas leichter war als derjenige der letzten beiden Jahrgänge.

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

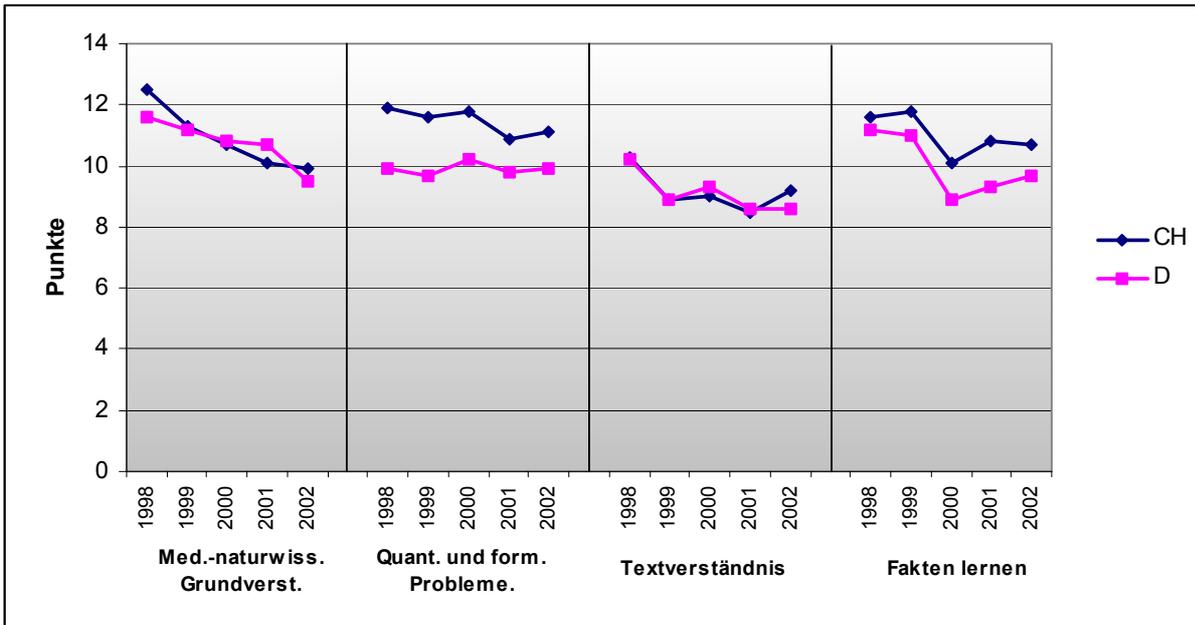


Abbildung 18: Mittelwerte eher sprachabhängiger Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2002

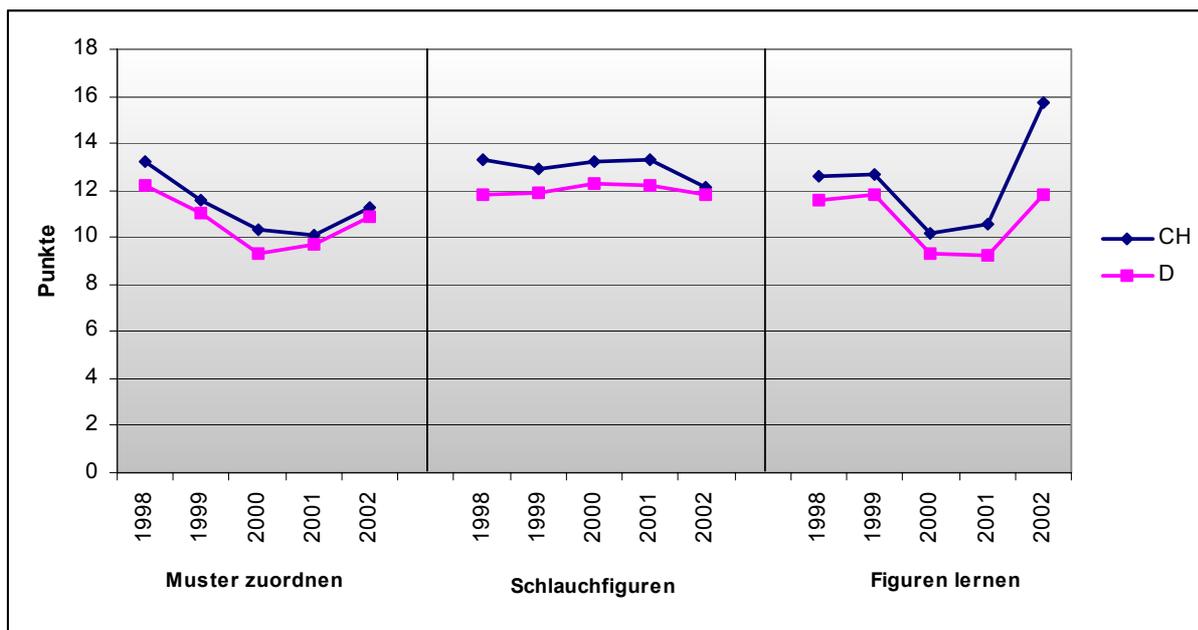


Abbildung 19: Mittelwerte wenig sprachabhängiger Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2002

Erstmals seit 1999 liegen die Ergebnisse im Untertest "Medizinisch-naturwissenschaftliches Grundverständnis" leicht über den deutschen Vergleichswerten. Im Untertest "Textverständnis" zeigt sich neu eine wahrnehmbare Differenz in dieser Richtung.

Im Untertest "Schlauchfiguren" fällt die Differenz geringer aus als in den Vorjahren. Erstmals liegen die Leistungen der beiden Vergleichsgruppen hier fast gleichauf.

Die Kurve für den Untertest "Figuren lernen" zeigt, dass die diesjährige Fassung dieses Untertests auch aufgrund der deutschen Daten als vergleichsweise leicht einzustufen war, allerdings nicht in dem Ausmass wie dies nun in den Schweizer Ergebnissen beobachtet werden konnte. Für die übrigen drei Untertests bleiben die Differenzen stabil.

Nach wie vor ist der Untertest "Textverständnis" der schwierigste Untertest. Erfreulich ist aber dort die Trendwende und erstmalige Verbesserung gegenüber den deutschen Vergleichswerten. Ähnliches gilt für den Untertest "Medizinisch-naturwissenschaftliches Grundverständnis", der aufgrund der deutschen Daten eigentlich als bedeutend schwieriger als in den Vorjahren einzustufen gewesen wäre. Sollte sich die Entwicklung dieser beiden Untertests, die sich sehr nahe an klassischen schulischen Anforderungen bewegen, in Zukunft bestätigen, wäre es möglicherweise lohnend, den Ursachen dieser Entwicklung nachzugehen. Eine mögliche Hypothese wäre, dass auf diese in der Vergangenheit eher schulisch vernachlässigten Bereiche (siehe auch Oswald, 1999) wieder stärkeres Augenmerk gerichtet wird.

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 solchen Unterschieden gerechnet werden. Die Abbildung 20 fasst zusammen, welche Ursachen für Unterschiede in den drei Sprachgruppen verantwortlich sein können:

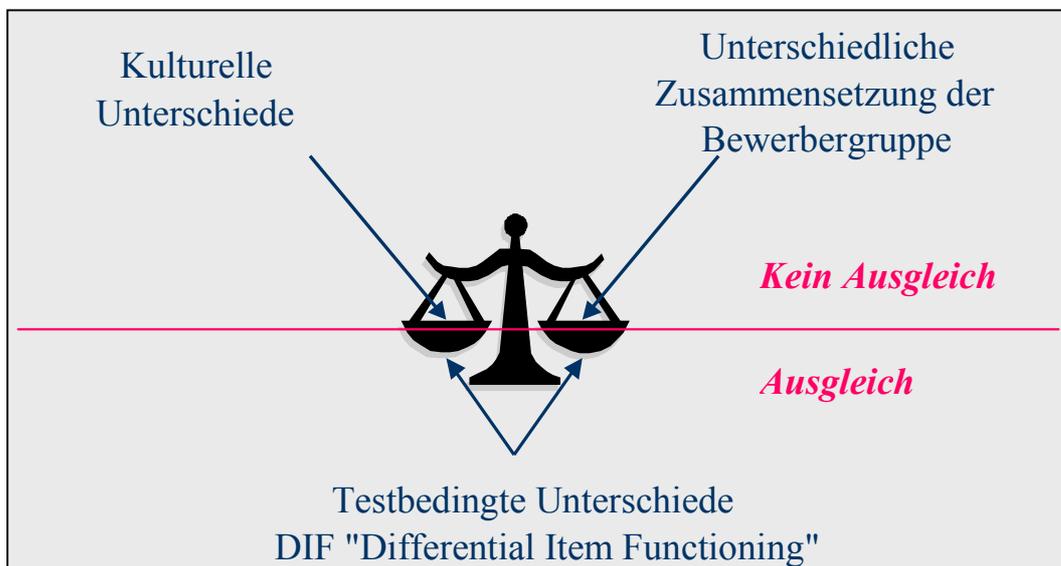


Abbildung 20: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen

Bei den ersten beiden Ursachen handelt es sich um "echte" Personen-Unterschiede, die nicht auf den Test oder seine Adaptation zurückgeführt werden können. Sie dürfen dementsprechend nicht ausgeglichen werden.

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

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

¹ DIF als Abkürzung für Differential Item Functioning

Kulturelle Unterschiede zwischen den Sprachgruppen

- Mögliche generelle Fähigkeitsunterschiede oder unterschiedliche Fähigkeitsprofile (anderes Verhältnis der einzelnen Fähigkeiten zueinander) im Vergleich der Kulturen können vorhanden sein, die beispielsweise auf Unterschiede im Schulsystem zurückzuführen sein können. 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.

Für das Jahr 2002 ist der Beschluss der Schweizerischen Hochschulkonferenz aus dem Jahre 1999 weiter 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 kaum betroffen sind.

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

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

Es ergibt sich für das Ausgleichsverfahren die nachfolgende Zielstellung:

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

7.4.1 Sprachvergleich für die Untertests

Wie in den Vorjahren treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Untertests gleichermassen auf. 1998 und 1999 erzielte die italienischsprachige Gruppe noch den besten Wert beim Untertest "Quantitative und formale Probleme", was seit der Durchführung im Jahr 2000 nicht mehr der Fall ist. 2002 liegen die Punktzahlen der deutschsprachigen Teilnehmer durchwegs über denen der französisch- und italienischsprachigen Kandidaten. Die Ergebnisse der italienischsprachigen Gruppe liegen, mit Ausnahme des Untertests "Diagramme und Tabellen", über jenen der französischsprachigen Teilnehmer (Abbildung 21).

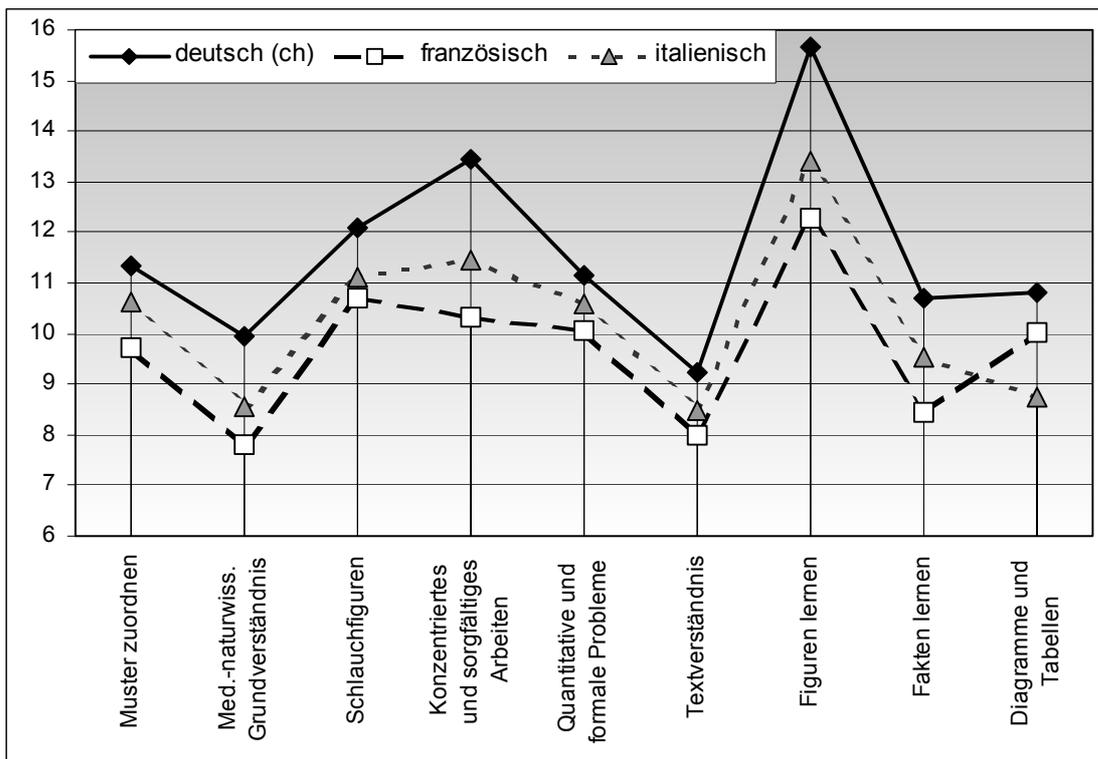


Abbildung 21: Mittelwerte der Untertests für die Sprachgruppen 2002 (unkorrigiert)

Die Abbildung 22 zeigt einen Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern über die Durchführungen. 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 hohe Punktzahl der deutschsprachigen Teilnehmer im Untertest "Figuren lernen" führte zu einer Ausdehnung der entsprechenden Spannweite. In drei weiteren Untertests wurden ebenfalls neue Höchstwerte (grosse Differenzen zwischen deutsch- und französischsprachigen Teilnehmern) erreicht. Bereits hier wird ersichtlich, dass die Differenzen in den sprachabhängigen Untertests zumindest nicht höher ausfallen als in den wenig sprachabhängigen.

Vergleichsweise (zu anderen Jahren) gute Ergebnisse erzielten die französischsprachigen Kandidaten in den Untertests "Schlauchfiguren", "Textverständnis" sowie "Diagramme und Tabellen".

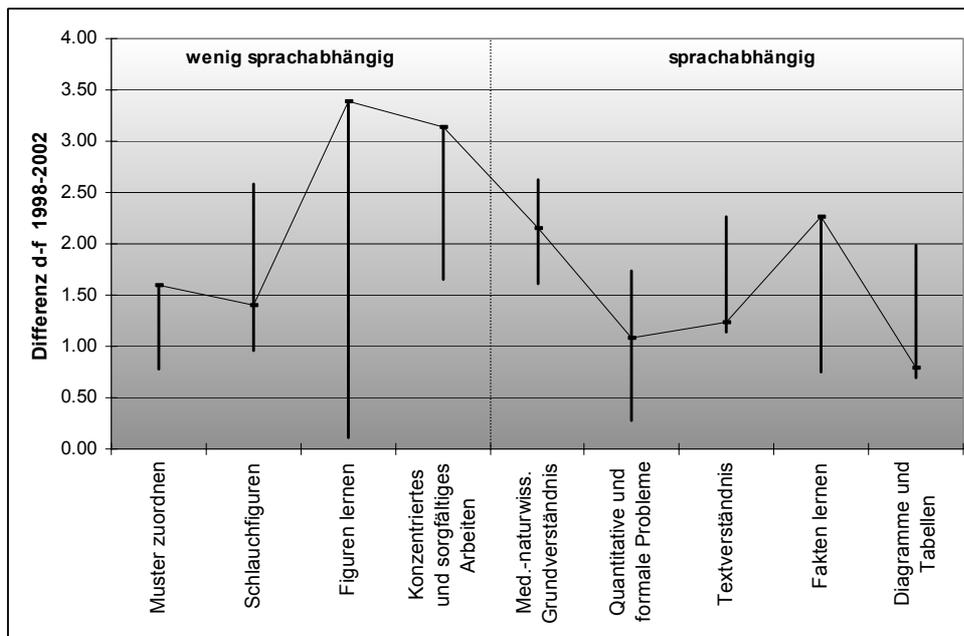


Abbildung 22: Differenz Testwert deutsch-französisch 2002 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2002

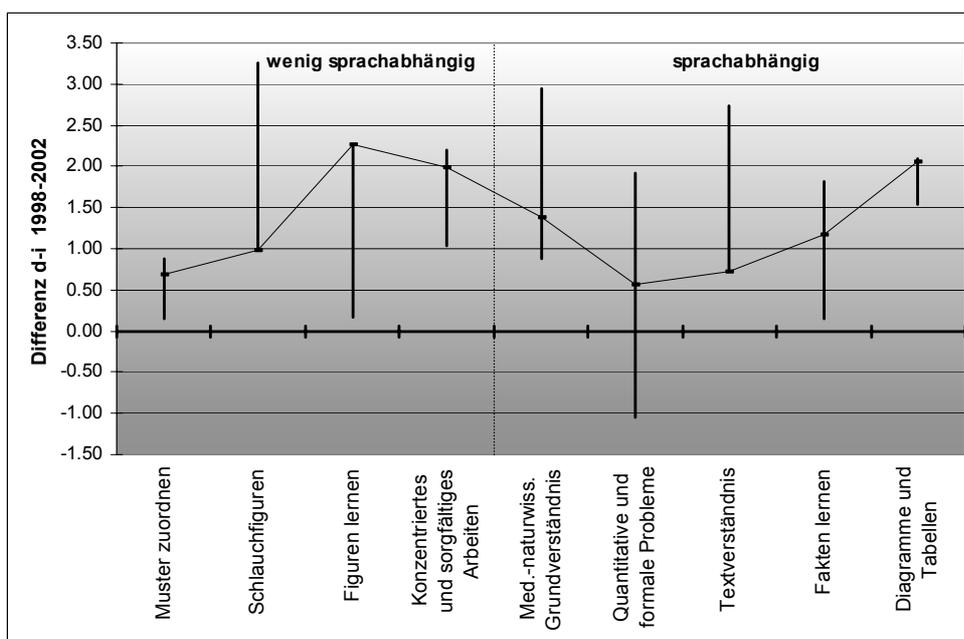


Abbildung 23: Differenz Testwert deutsch-italienisch 2002 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2002

Im Vergleich zwischen deutsch- und italienischsprachigen Kandidaten wird im Untertest "Figuren lernen" ebenfalls ein neuer Höchstwert erreicht (Abbildung 23). Vergleichsweise gute Ergebnisse wurden andererseits in den Untertests "Schlauchfiguren", "Medizinisch-naturwissenschaftliches Grundverständnis" und "Textverständnis" erreicht.

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Testwert ** <i>D > (F,I)</i>	deutsch	886	100.63	9.85	72	128
	französisch	59	93.10	9.34	70	112
	italienisch	28	95.32	8.96	81	112
Muster zuordnen ** <i>D > F</i>	deutsch	886	11.33	2.80	0	20
	französisch	59	9.73	2.98	2	18
	italienisch	28	10.64	3.39	4	16
Med.-naturwiss. Grundverständnis ** <i>D > F</i>	deutsch	886	9.95	3.09	0	20
	französisch	59	8.41	2.86	4	16
	italienisch	28	8.83	2.81	3	14
Schlauchfiguren ** <i>D > F</i>	deutsch	886	12.10	3.50	2	20
	französisch	59	10.69	3.11	2	17
	italienisch	28	11.11	3.63	4	18
Quantitative und formale Probleme <i>homogen</i>	deutsch	886	11.14	3.70	2	20
	französisch	59	10.39	3.91	2	18
	italienisch	28	10.57	3.95	4	20
Textverständnis <i>homogen</i>	deutsch	886	9.22	3.41	2	18
	französisch	59	8.54	3.03	3	16
	italienisch	28	8.50	3.18	5	17
Figuren lernen ** <i>D > F</i>	deutsch	886	15.66	3.79	4	20
	französisch	59	12.27	4.00	2	19
	italienisch	28	13.39	2.94	7	18
Fakten lernen ** <i>D > F</i>	deutsch	886	10.72	3.75	1	20
	französisch	59	9.16	2.85	3	16
	italienisch	28	9.54	3.35	4	18
Diagramme und Tabellen <i>homogen</i>	deutsch	886	10.81	3.14	2	19
	französisch	59	10.02	3.09	5	17
	italienisch	28	9.91	2.58	4	14
Konzentriertes und sorgf. Arbeiten ** <i>D > (F,I)</i>	deutsch	886	13.46	3.46	0	20
	französisch	59	10.32	3.12	0	17
	italienisch	28	11.46	2.85	6	19

Tabelle 15: Testwert und Punktwerte der Untertests für die Sprachgruppen. **: signifikanter Einfluss des Faktors "Sprache" bei varianzanalytischer Prüfung (1%-Niveau); daneben sind signifikante Unterschiede bei multiplem Mittelwertvergleich angegeben. In französischer und italienischer Sprachgruppe korrigierte Punktwerte

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

Da die Differenzen und ihre Schwankungen zwischen den einzelnen Jahren sowohl in sprachabhängigen wie auch wenig sprachabhängigen Untertests auftreten, sind grössere adaptionsbedingte Textunterschiede als massgebliche Ursache dafür sehr unwahrscheinlich. Durch das Korrekturverfahren werden vorhandene testbedingte Differenzen zwischen den Sprachversionen ausgeglichen – die Erwartung, dass danach alle Sprachgruppen die gleichen Mittelwerte erreichen, ist dennoch gering. Wegen des hierarchischen Faktormodells mit hohem Generalfaktoranteil an der Varianzaufklärung werden die Unterschiede wahrscheinlich mindestens so hoch wie in den wenig sprachabhängigen Untertests bleiben.

Die Mittelwertsdifferenzen für jeden Untertest **nach** dem Sprachausgleich sind der Tabelle 15 zu entnehmen. Die varianzanalytische Prüfung zeigt nach dem Ausgleichsverfahren für die Untertests "Quantitative und formale Probleme", "Textverständnis" sowie "Diagramme und Tabellen" keinen signifikanten Einfluss des Faktors Sprache. In den anderen sprachabhängigen Untertests ("Medizinisch-naturwissenschaftliches Grundverständnis", "Fakten lernen") bestehen auch nach dem Sprachausgleich Differenzen, die nun aber mit grosser Sicherheit nicht auf adaptionsbedingte Ursachen zurückzuführen sind.

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.

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

7.4.2 Darstellung des Korrekturverfahrens

Die Identifikation und der Ausgleich testbedingter Unterschiede geschieht 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 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 16) zeigen sich in vier der fünf sprachabhängigen Untertests signifikante Differenzen. Die geringeren Unterschiede im Untertest "Diagramme und Tabellen" sind nicht signifikant, daher wird dort kein Sprachausgleich für die französischsprachigen Teilnehmer vorgenommen.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	886	9.95	3.093	.000	Ja
	f	59	7.80	2.981		
Quantitative und formale Probleme	d	886	11.14	3.697	.030	Ja
	f	59	10.05	4.015		
Textverständnis	d	886	9.22	3.411	.007	Ja
	f	59	7.98	3.149		
Fakten lernen	d	886	10.72	3.752	.000	Ja
	f	59	8.46	3.013		
Diagramme und Tabellen	d	886	10.81	3.145	.060	Nein
	f	59	10.02	3.093		

Tabelle 16: Signifikanzprüfung deutsche vs. französische Sprachform

Die italienische Form (Tabelle 17) weist in zwei sprachabhängigen Untertests signifikante Unterschiede auf: "Medizinisch-naturwissenschaftliches Grundverständnis" sowie "Diagramme und Tabellen".

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	886	9.95	3.093	.020	Ja
	i	28	8.57	2.860		
Quantitative und formale Probleme	d	886	11.14	3.697	.424	Nein
	i	28	10.57	3.948		
Textverständnis	d	886	9.22	3.411	.259	Nein
	i	28	8.50	3.180		
Fakten lernen	d	886	10.72	3.752	.100	Nein
	i	28	9.54	3.350		
Diagramme und Tabellen	d	886	10.52	3.07	.001	Ja
	i	28	8.43	3.41		

Tabelle 17: 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 Items für die 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 Itemschwierigkeiten z-standardisiert und anschliessend in " Δ -Werte" transformiert. Die Transformation erfolgt über die Formel

$$\Delta = 13 - 4z$$

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

Die aus den Delta-Werten abgeleitete Regressionsgerade $Y = AX + B$ beschreibt die Beziehung zwischen den interessierenden Sprachversionen. Eine graphische Darstellung der Delta-Werte (Delta-Plot) würde im Idealfall eine ellipsenförmige Anordnung von Punkten entlang der Diagonalen ergeben (Abbildung 24). Dies würde bedeuten, dass sowohl die Itemschwierigkeiten wie auch deren Reihenfolge in beiden Sprachversionen vergleichbar sind.

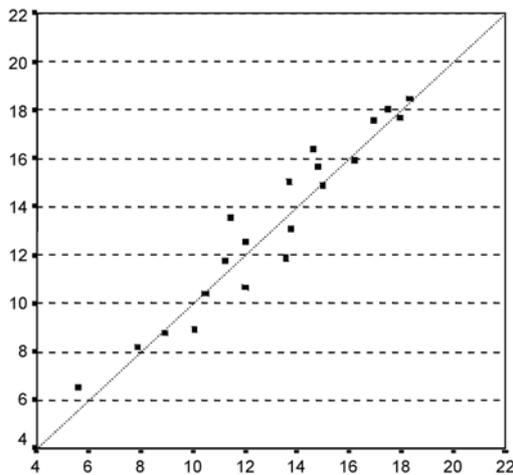


Abbildung 24: Beispiel für einen Delta-Plot

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 Achsen-schnittpunkt 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 unterbrochene 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 Achsen-schnittpunkt 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)

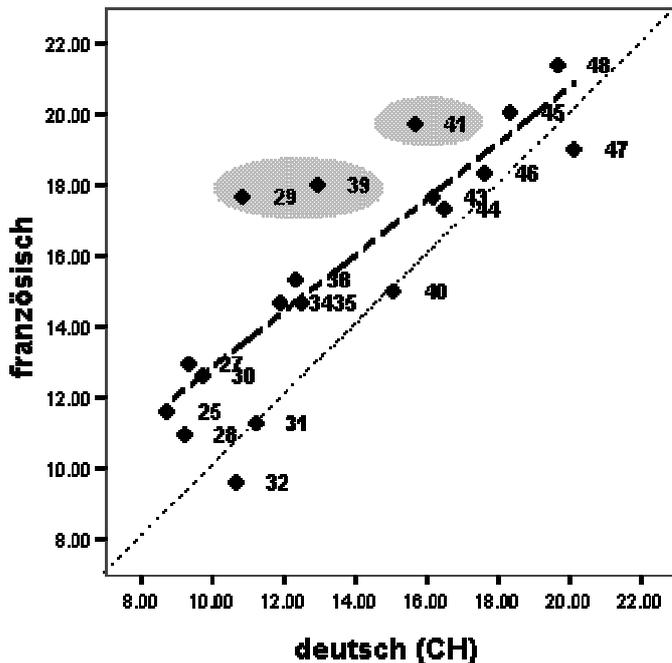


Abbildung 25 zeigt den Delta-Plot für den Untertest "MNGV" (Vergleich zwischen deutsch- und französischsprachiger Version). Die Items 29, 39 und 41 fallen hier als zu schwer auf. Die Items 32 und 47 andererseits wurden möglicherweise durch die Adaptation vereinfacht.

Abbildung 25: Delta-Plot für "MNGV" (deutsch-französisch)

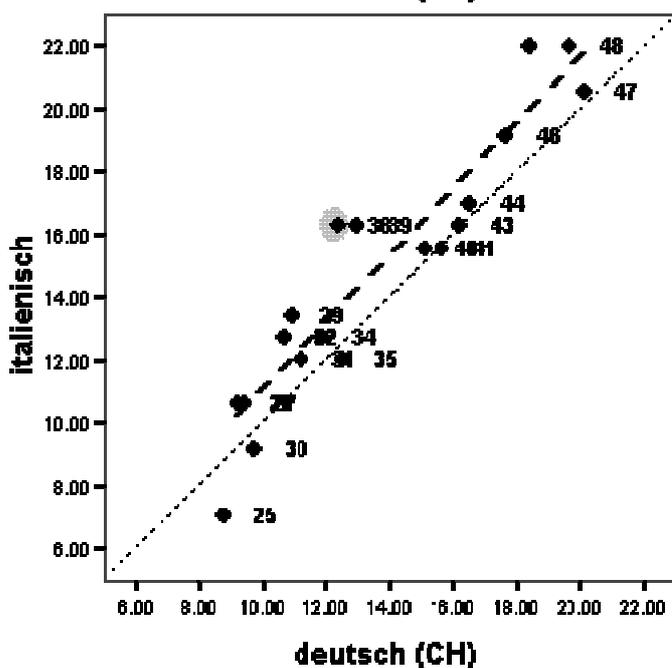


Abbildung 26 zeigt den entsprechenden Vergleich für die italienischsprachige Gruppe. Neben einer klaren Parallelverschiebung fallen hier die Items 36 und 38 (identische Werte) auf.

Abbildung 26: Delta-Plot für "MNGV" (deutsch-italienisch)

Die genauen Werte für beide Sprachvergleiche sind aus Tabelle 18 ersichtlich. Die ersten drei Spalten zeigen die Itemschwierigkeiten für die Sprachen, danach folgen die in Delta-Werte transformierten Angaben. DIF weisen jene Items auf, bei welchen die absolute Differenz der Deltawerte 1.5 oder mehr beträgt.

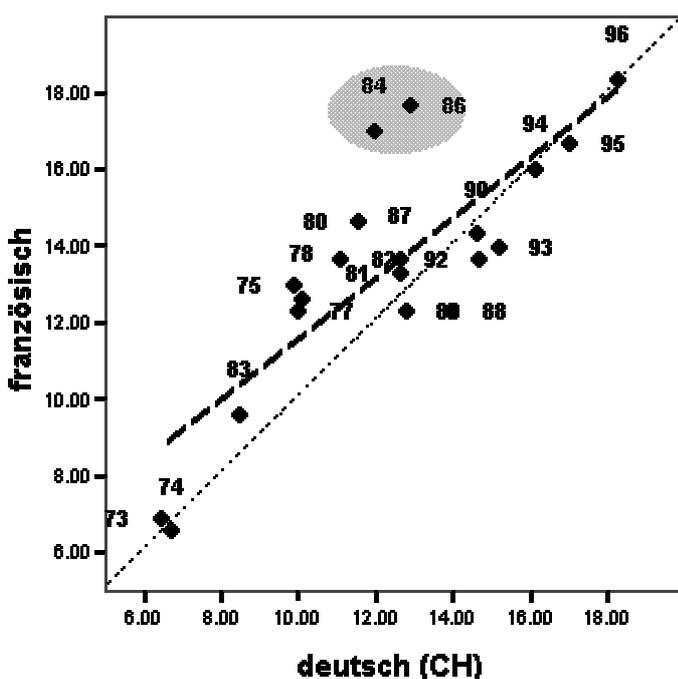
Da kein symmetrisches Verfahren gewählt wurde, sind nur negative Werte relevant. Der Bonus für die französische oder italienische Sprachgruppe berechnet sich aus der Differenz zwischen den ursprünglichen Itemschwierigkeiten.

Maximal können in diesem Untertest den französisch- oder italienischsprachigen Teilnehmern, welche die entsprechenden Items nicht gelöst haben, 0.82 (f) respektive 0.40 (i) Punkte gutgeschrieben werden.

Item	Schwierigkeit			Δ -Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
25	0.74	0.59	0.82	8.73	11.65	7.11	0.16	2.02		
27	0.71	0.53	0.64	9.35	13.00	10.66	-0.51	0.08		
28	0.72	0.63	0.64	9.22	10.98	10.66	0.99	-0.02		
29	0.63	0.29	0.50	10.86	17.71	13.50	-3.28	-0.73	0.35	
30	0.69	0.54	0.71	9.71	12.66	9.24	-0.02	1.31		
31	0.62	0.61	0.57	11.22	11.31	12.08	1.97	0.49		
32	0.64	0.69	0.54	10.65	9.63	12.79	2.94	-0.40		
34	0.58	0.44	0.54	11.87	14.68	12.79	-0.27	0.49		
35	0.55	0.44	0.57	12.52	14.68	12.08	0.13	1.45		
36	0.56	0.41	0.36	12.36	15.36	16.34	-0.50	-1.54		0.20
38	0.56	0.41	0.36	12.36	15.36	16.34	-0.50	-1.54		0.20
39	0.53	0.27	0.36	12.94	18.05	16.34	-2.25	-1.11	0.26	
40	0.42	0.42	0.39	15.07	15.02	15.63	1.45	0.94		
41	0.39	0.19	0.39	15.66	19.74	15.63	-1.89	1.37	0.21	
43	0.36	0.29	0.36	16.20	17.71	16.34	0.03	1.28		
44	0.35	0.31	0.32	16.51	17.38	17.05	0.49	1.04		
45	0.26	0.17	0.07	18.35	20.07	22.02	-0.48	-0.97		
46	0.29	0.25	0.21	17.63	18.39	19.18	0.39	0.42		
47	0.17	0.22	0.14	20.12	19.06	20.60	1.41	1.30		
48	0.19	0.10	0.07	19.65	21.42	22.02	-0.73	-0.01		

Tabelle 18: DIF-Analyse "Medizinisch-naturwissenschaftliches Grundverständnis"

7.4.2.4 Quantitative und formale Probleme



Die folgende Abbildung zeigt den Delta-Plot-Vergleich für den Unterstest "Quantitative und formale Probleme". Hier ist nur die französischsprachige Version vom Ausgleichsverfahren betroffen.

Die Regressionsgerade zeigt vor allem im unteren, "leichten" Bereich eine Abweichung von der Diagonalen. Darüber hinaus fallen die Items 84 und 86 auf, welche sich in der französischen Schwierigkeitshierarchie deutlich von der deutschsprachigen Version unterscheiden.

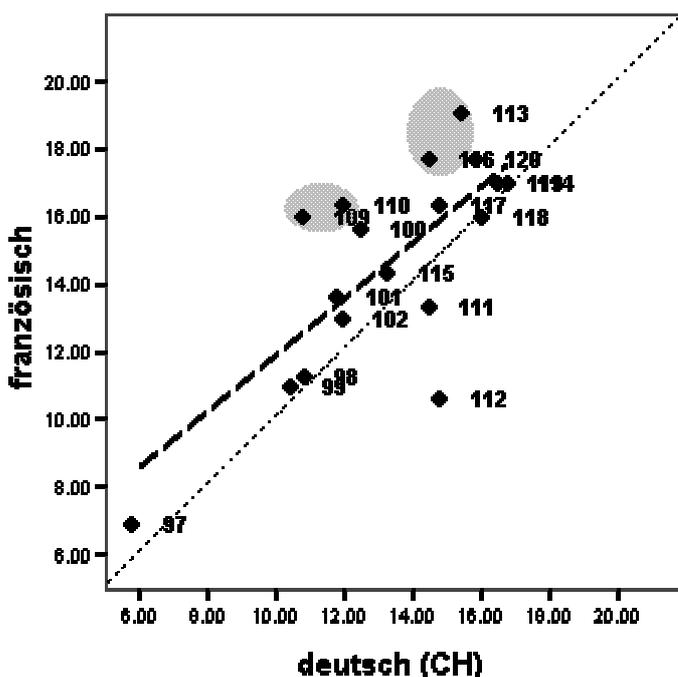
Abbildung 27: Delta-Plot für "Quantitative und formale Probleme" (deutsch-französisch)

Der maximal erreichbare Sprachbonus beträgt hier für die französischsprachigen Kandidaten 0.49 Punkte (Tabelle 19).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
73	0.86	0.83	0.86	6.44	6.93		1.44			
74	0.84	0.85	0.93	6.71	6.60		1.88			
75	0.68	0.53	0.61	9.89	13.00		-1.17			
77	0.68	0.56	0.61	10.00	12.32		-0.57			
78	0.67	0.54	0.68	10.12	12.66		-0.77			
80	0.60	0.44	0.61	11.57	14.68		-1.45			
81	0.54	0.51	0.79	12.67	13.33		0.29			
82	0.62	0.49	0.50	11.10	13.67		-0.95			
83	0.75	0.69	0.93	8.50	9.63		0.61			
84	0.58	0.32	0.36	11.98	17.04		-3.05		0.25	
86	0.53	0.29	0.32	12.94	17.71		-2.98		0.24	
87	0.54	0.49	0.43	12.65	13.67		0.01			
88	0.48	0.56	0.43	13.97	12.32		1.89			
89	0.53	0.56	0.57	12.81	12.32		1.16			
90	0.44	0.46	0.43	14.65	14.35		0.72			
92	0.44	0.49	0.32	14.71	13.67		1.29			
93	0.42	0.47	0.39	15.19	14.01		1.32			
94	0.37	0.37	0.14	16.17	16.03		0.34			
95	0.32	0.34	0.39	17.05	16.70		0.36			
96	0.35	0.32	0.33	15.95	16.62		-0.20			

Tabelle 19: DIF-Analyse "Quantitative und formale Probleme"

7.4.2.5 Textverständnis



Für den Untertest "Textverständnis" fallen im Vergleich der deutschsprachigen mit der französischen Version (Abbildung 28) vier Items bezüglich DIF auf. Maximal wären hier 0.80 Punkte gutzuschreiben, falls alle betreffenden Items nicht korrekt bearbeitet wurden.

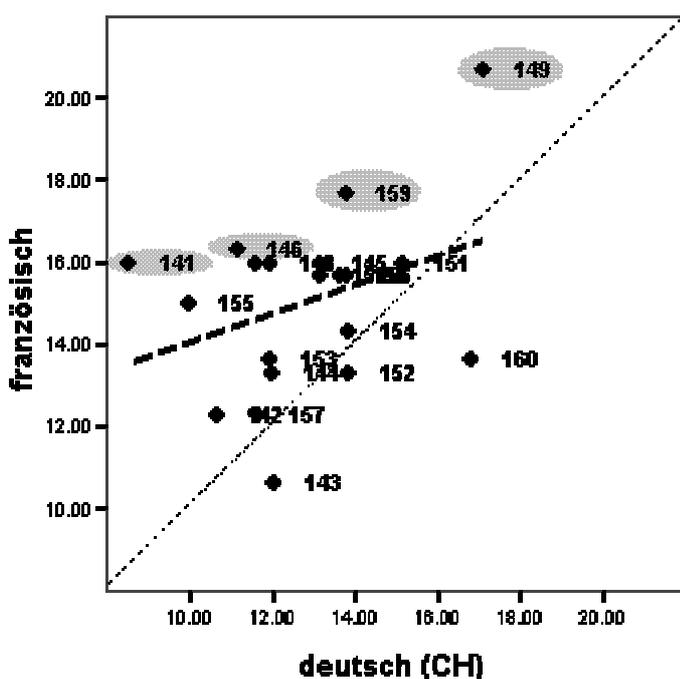
Item 112 erscheint in dieser Gruppe als deutlich leichter.

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

Item	Schwierigkeit			Δ -Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
97	0.89	0.83	0.82	5.79	6.93		1.16			
98	0.63	0.61	0.54	10.83	11.31		1.01			
99	0.65	0.63	0.57	10.45	10.98		1.03			
100	0.55	0.39	0.43	12.47	15.69		-1.31			
101	0.59	0.49	0.54	11.78	13.67		-0.20			
102	0.58	0.53	0.61	11.96	13.00		0.43			
109	0.64	0.37	0.61	10.81	16.03		-2.63		0.26	
110	0.58	0.36	0.75	11.98	16.37		-2.14		0.22	
111	0.45	0.51	0.36	14.49	13.33		1.79			
112	0.43	0.64	0.54	14.83	10.64		4.08			
113	0.40	0.22	0.29	15.45	19.06		-2.00		0.18	
114	0.33	0.32	0.14	16.80	17.04		0.42			
115	0.51	0.46	0.39	13.28	14.35		0.24			
116	0.45	0.29	0.46	14.49	17.71		-1.58		0.16	
117	0.43	0.36	0.50	14.83	16.37		-0.32			
118	0.37	0.37	0.25	16.06	16.03		0.72			
119	0.35	0.32	0.39	16.51	17.04		0.23			
120	0.38	0.29	0.32	15.84	17.71		-0.72			

Tabelle 20: DIF-Analyse "Textverständnis"

7.4.2.6 Fakten lernen



Im Gegensatz zu den Ergebnissen des Jahres 2001 besteht im Untertest "Fakten lernen" wieder eine signifikante Differenz zwischen den Leistungen der deutschsprachigen und jener der französischsprachigen Kandidaten.

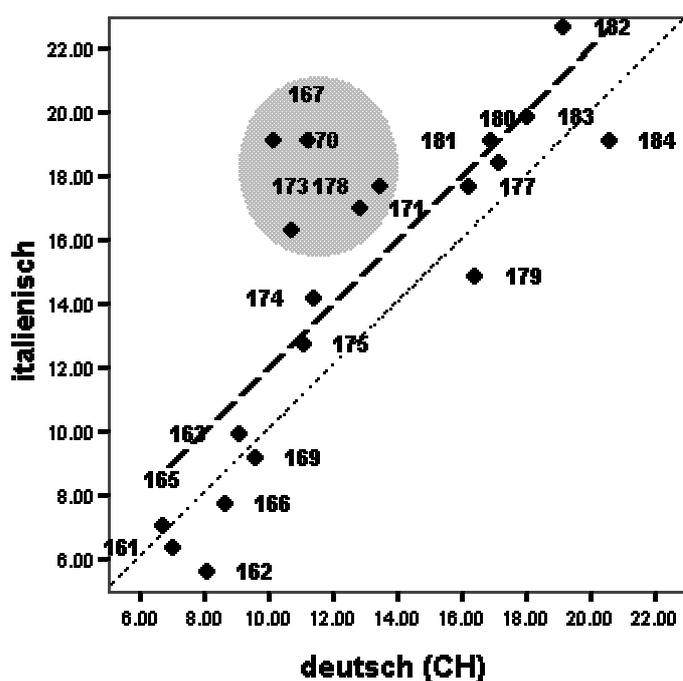
Abbildung 29 zeigt, dass die Adaptation hier möglicherweise die Itemkennwerte am stärksten beeinflusst hat. Besonders abweichend erscheinen die Items 141, 146, 149 und 159.

Abbildung 29: Delta-Plot für "Fakten lernen" (deutsch-französisch)

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
141	0.75	0.37	0.61	8.55	16.03	Kein Ausgleich	-2.35	Kein Ausgleich	0.38	Kein Ausgleich
142	0.64	0.56	0.50	10.65	12.32		1.85			
143	0.57	0.64	0.46	12.05	10.64		3.90			
144	0.58	0.51	0.50	12.00	13.33		1.34			
145	0.52	0.37	0.54	13.17	16.03		-0.81			
146	0.62	0.36	0.50	11.15	16.37		-1.80		0.26	
147	0.60	0.37	0.75	11.60	16.03		-1.33			
148	0.58	0.37	0.50	11.93	16.03		-1.22			
149	0.32	0.14	0.46	17.09	20.75		-3.96		0.18	
150	0.49	0.39	0.29	13.66	15.69		-0.33			
151	0.42	0.37	0.29	15.16	16.03		-0.15			
152	0.48	0.51	0.61	13.84	13.33		1.95			
153	0.58	0.49	0.29	11.96	13.67		1.01			
154	0.48	0.46	0.43	13.84	14.35		1.00			
155	0.68	0.42	0.71	9.98	15.02		-0.92			
156	0.49	0.39	0.54	13.77	15.69		-0.29			
157	0.59	0.56	0.46	11.66	12.32		2.18			
158	0.52	0.39	0.25	13.17	15.69		-0.50			
159	0.49	0.29	0.32	13.77	17.71		-2.20		0.20	
160	0.33	0.49	0.54	16.82	13.67		2.63			

Tabelle 21: DIF-Analyse "Fakten lernen"

7.4.2.7 Diagramme und Tabellen



Der Untertest "Diagramme und Tabellen" ist nur für die italienischsprachigen Kandidaten zu analysieren.

Einerseits ist hier wiederum eine deutliche Parallelverschiebung zu beobachten; im (deutschen) mittelschweren Bereich befindet sich allerdings eine Itemgruppe, welche möglicherweise durch die Adaptation erschwert wurde. Im Einzelnen sind dies die Items 167, 170, 171, 173 und 178.

Abbildung 30: Delta-Plot für "Diagramme und Tabellen" (deutsch-italienisch)

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
161	0.83	0.78	0.86	6.98		6.40		1.83		
162	0.77	0.69	0.89	8.05		5.70		3.10		
163	0.72	0.68	0.68	9.04		9.95		0.79		
165	0.84	0.83	0.82	6.68		7.11		1.13		
166	0.75	0.73	0.79	8.61		7.82		1.99		
167	0.62	0.51	0.21	11.22		19.18		-4.18		0.40
169	0.70	0.75	0.71	9.56		9.24		1.66		
170	0.67	0.66	0.21	10.12		19.18		-4.96		0.46
171	0.54	0.47	0.32	12.79		17.05		-1.57		0.21
173	0.64	0.61	0.36	10.70		16.34		-2.54		0.28
174	0.61	0.58	0.46	11.39		14.21		-0.55		
175	0.62	0.53	0.54	11.04		12.79		0.20		
177	0.36	0.32	0.29	16.22		17.76		0.36		
178	0.50	0.46	0.29	13.46		17.76		-1.59		0.22
179	0.35	0.24	0.43	16.40		14.92		2.49		
180	0.32	0.31	0.25	17.11		18.47		0.50		
181	0.33	0.31	0.21	16.85		19.18		-0.20		
182	0.22	0.14	0.04	19.16		22.73		-1.07		
183	0.27	0.27	0.18	18.01		19.89		0.13		
184	0.15	0.17	0.21	20.55		19.18		2.42		

Tabelle 22: DIF-Analyse "Diagramme und Tabellen"

7.4.2.8 Effekte der Korrektur

Die durch die Ausgleichsprozedur resultierenden Bonuspunkte können Tabelle 23 entnommen werden. Durchschnittlich werden den französischsprachigen Teilnehmern 2.21 Punkte, den italienischsprachigen Teilnehmern 1.41 Punkte gutgeschrieben. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der Bonus in der französischsprachigen Gruppe beträgt also 1 - 3, in der italienischsprachigen Gruppe 1 - 2 Punkte.

		n	Minimum	Maximum	m	s
		franz.	Med.-naturwiss. Grundverständnis	59	.00	.82
Quant. und form. Probleme	59		.00	.49	.34	.16
Textverständnis	59		.00	.82	.56	.22
Fakten lernen	59		.00	1.02	.70	.26
GESAMT	59		.76	2.97	2.21	.53
ital.	Med.-naturwiss. Grundverständnis	28	.00	.40	.26	.13
	Diagramme und Tabellen	28	.22	1.57	1.16	.34
	GESAMT	28	.62	1.77	1.41	.34

Tabelle 23: Mittelwerte und Standardabweichungen der Korrekturwerte für die Personen der französisch- und italienischsprachigen Gruppen

In Abbildung 31 und Abbildung 32 sind die Verteilungen der Bonuspunkte für die letzten vier Testdurchführungen dargestellt (1998 wurde noch kein Sprachausgleich vorgenommen). Die senkrechten Linien bezeichnen die Spannweite in jedem Jahr (Minimum bis Maximum), der jeweilige Mittelwert wird durch die kurze waagerechte Markierung gekennzeichnet.

Für die französischsprachigen Teilnehmer ist nach den letztjährigen geringen Bonuswerten wieder ein Anstieg auf das Niveau der Vorjahre zu verzeichnen. Auch die Werte für die italienischsprachigen Kandidaten liegen, nach einer Erhöhung im Jahre 2000, im üblichen Rahmen.

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.

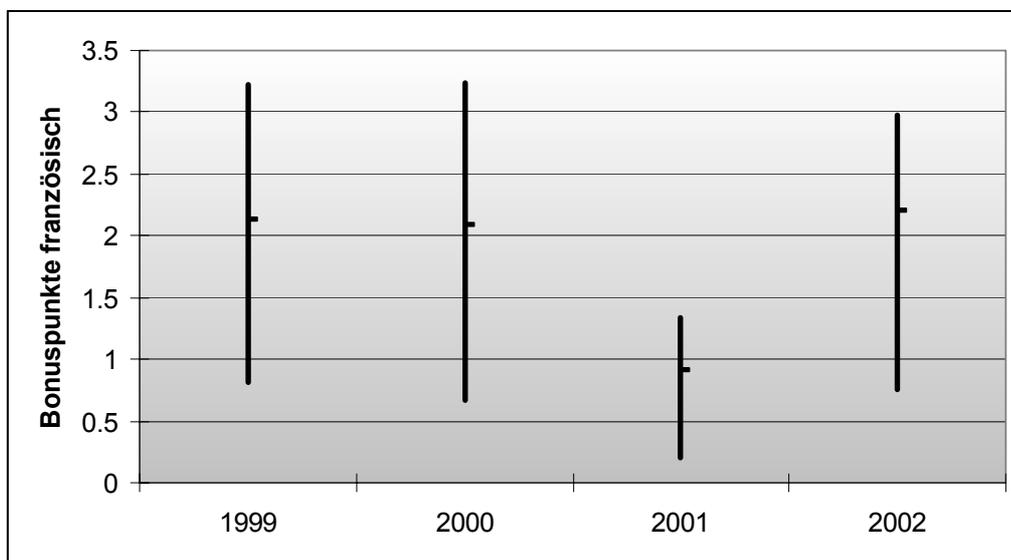


Abbildung 31: Auswirkungen der Korrektur über die Jahre für die französischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

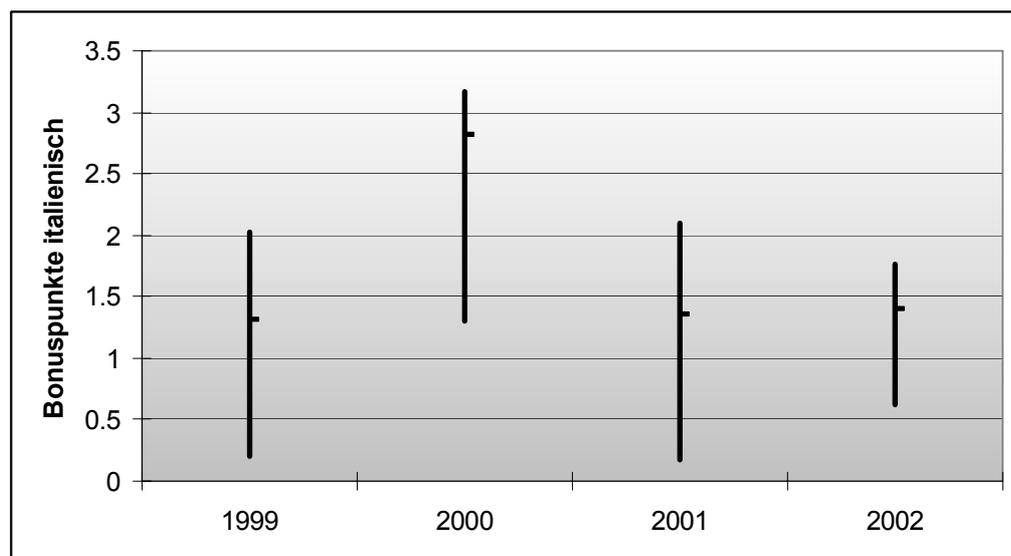


Abbildung 32: Auswirkungen der Korrektur über die Jahre für die italienischsprachigen Kandidaten (Minimum – Mittelwert – Maximum)

7.4.2.9 Evaluation: Gleichbehandlung der Sprachgruppen

Wirken sich die vorhandenen Unterschiede in den Punktwerten auf die Zulassung aus? Wie in den Vorjahren zeigen sich bezüglich der Zulassungsquoten zur Humanmedizin signifikante Unterschiede zwischen den Sprachgruppen. Bedingt durch die geringe Stichprobengröße und die hohe Zulassungsquote in der Veterinärmedizin lassen sich dort keine Differenzen nachweisen (Tabelle 24).

Disziplin	Sprache	Warteliste	zugelassen	Total
Humanmedizin	deutsch	104 14.2%	627 85.8%	731 100.0%
	französisch	6 31.6%	13 68.4%	19 100.0%
	italienisch	5 25.0%	15 75.0%	20 100.0%
	Total	115 14.9%	655 85.1%	770 100.0%
Veterinärmedizin	deutsch	10 6.5%	145 93.5%	155 100.0%
	französisch	6 15.0%	34 85.0%	40 100.0%
	italienisch	1 12.5%	7 87.5%	8 100.0%
	Total	17 8.4%	186 91.6%	203 100.0%

Tabelle 24: Zulassung nach Sprache des Tests

Die Signifikanzprüfung für die Zulassungsunterschiede kann Tabelle 25 entnommen werden. Zulassungsunterschiede sind knapp auf dem 5%-Niveau nachzuweisen.

Disziplin		Wert	df	Asymp. Sig. (2seitig)
Humanmedizin	Pearson Chiquadrat	6.026	2	.049
	Likelihood Ratio	4.990	2	.082
Veterinärmedizin	Pearson Chiquadrat	3.213	2	.201
	Likelihood Ratio	2.852	2	.240

Tabelle 25: Prüfung auf Verteilungsunterschiede für Sprachgruppen getrennt nach Disziplinen: signifikanter Unterschied für Humanmedizin, kein signifikanter Unterschied für Veterinärmedizin (aber geringe Stichprobengröße)

In den folgenden Abbildungen sind die Wartelisten-Quoten für Humanmedizin und Veterinärmedizin dargestellt. Hohe Werte bedeuten eine geringere Zulassungsquote. Nachdem 2001 zwischen deutschsprachigen einerseits und italienisch- und französischsprachigen Teilnehmern andererseits diesbezüglich deutliche Differenzen bestanden, wurden diese nun wieder etwas verringert (Abbildung 33).

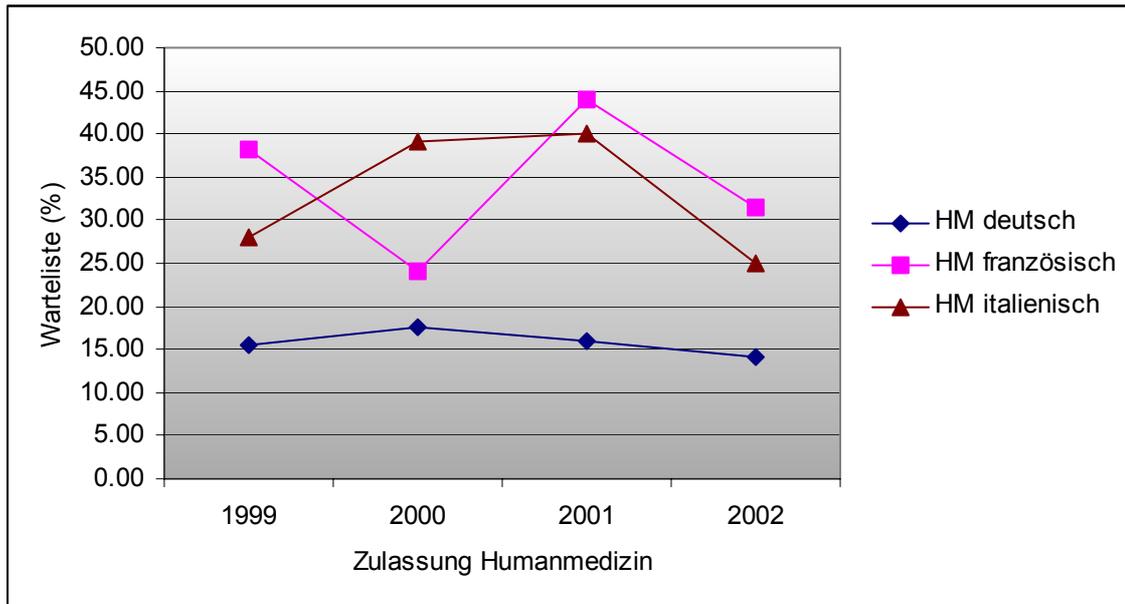


Abbildung 33: Zulassungsquoten Humanmedizin nach Sprache 1999 bis 2002

In der Veterinärmedizin wurde die Zulassungsquote wieder leicht gesenkt – in allen Sprachversionen stieg dort der Anteil der Personen auf der Warteliste (Abbildung 34). Damit wurden die eher zu "grosszügig" ausgefallenen Kapazitätserweiterungen des letzten Jahres (aufgrund der Doppelmaturität) korrigiert.

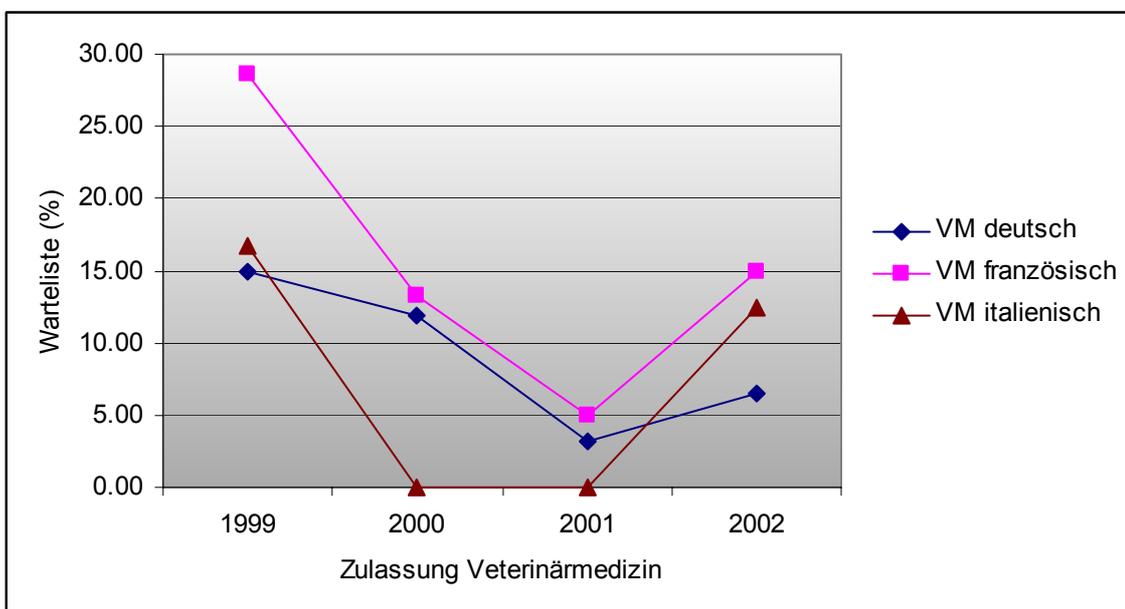


Abbildung 34: Zulassungsquoten Veterinärmedizin nach Sprache 1999 bis 2002

Es muss weiterhin sichergestellt sein, dass die unterschiedlichen Zulassungsquoten der Sprachgruppen nicht am Test selber 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 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 auch dieses Jahr für beide adaptierten Sprachversionen erneut 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. Durch die Korrektur werden diese Unterschiede weiter verringert, so dass eher von einer Überkompensation als von einer Benachteiligung der französisch- und italienischsprachigen Kandidaten ausgegangen werden kann.

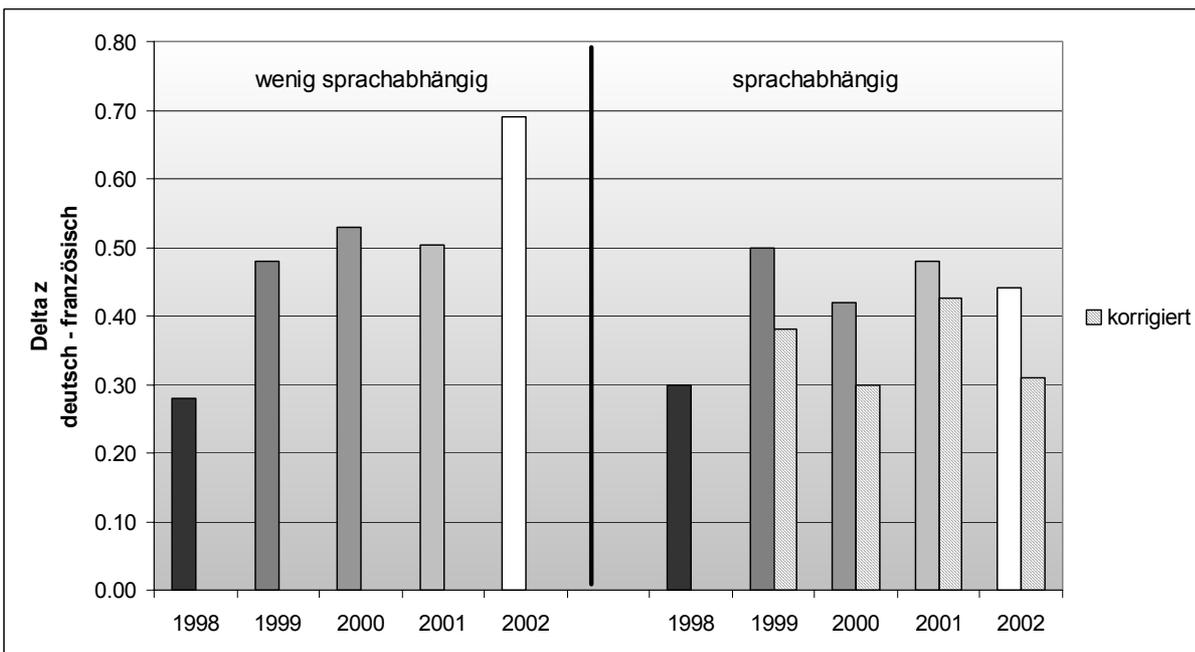


Abbildung 35: Abweichung (Delta z) für den Testwert zwischen deutscher und französischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

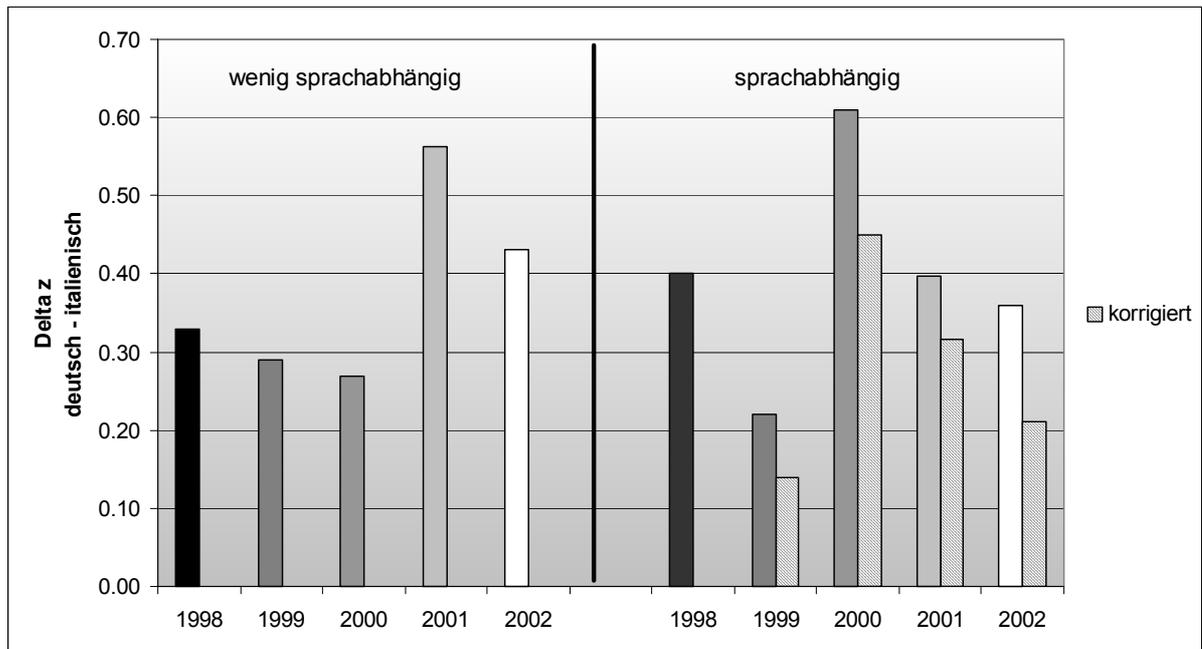


Abbildung 36: Abweichung (Delta z) für den Testwert zwischen deutscher und italienischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

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 – 2002

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	2001	773	14.03	3.59	.00
	2002	886	15.09	3.88	
französisch	2001	65	12.17	3.55	.03
	2002	59	13.66	3.79	
italienisch	2001	42	12.79	3.71	.38
	2002	28	13.64	4.25	

Tabelle 26: Verteilungskennwerte identischer Einstreuaufgaben 2001 und 2002

Abbildung 37 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. Erstmals ist dieses Jahr auch die Verbesserung der französischsprachigen Gruppe gegenüber dem Vorjahr signifikant. Diese Verbesserung hat sich wie bereits beschrieben auch in Form einer höheren Zulassungsquote ausgewirkt.

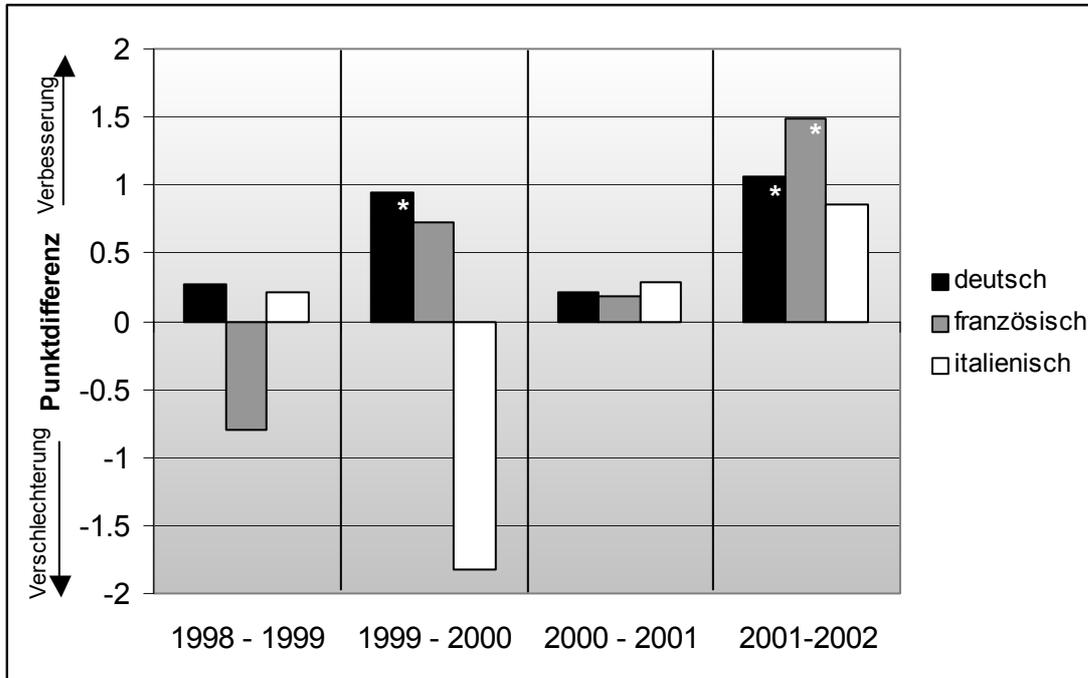


Abbildung 37: Verlauf der Leistungen in identischen Testaufgaben 1998 bis 2002 (Differenzverlauf); *= signifikante Differenz

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".

Auch eine Überprüfung der Daten zeigt keine Auffälligkeiten zwischen den Testlokalen für diesen Untertest (Tabelle 27). Einziger in den Teilen "Muster zuordnen", "Fakten lernen" sowie "Konzentriertes und sorgfältiges Arbeiten" treten statistisch signifikante Unterschiede zwischen den (deutschsprachigen) Testlokalen auf.

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testwert	<i>Zwischen Gruppen</i>	1602.93	11	145.72	1.51	.122
	<i>Innerhalb Gruppen</i>	84315.38	874	96.47		
	<i>Total</i>	85918.31	885			
Muster zuordnen	<i>Zwischen Gruppen</i>	156.42	11	14.22	1.83	.046
	<i>Innerhalb Gruppen</i>	6804.32	874	7.79		
	<i>Total</i>	6960.73	885			
Med.-naturwiss. Grundverständnis	<i>Zwischen Gruppen</i>	116.58	11	10.60	1.11	.350
	<i>Innerhalb Gruppen</i>	8347.61	874	9.55		
	<i>Total</i>	8464.19	885			
Schlauchfiguren	<i>Zwischen Gruppen</i>	150.23	11	13.66	1.11	.347
	<i>Innerhalb Gruppen</i>	10716.22	874	12.26		
	<i>Total</i>	10866.46	885			
Quant. und formale Probleme	<i>Zwischen Gruppen</i>	224.83	11	20.44	1.50	.124
	<i>Innerhalb Gruppen</i>	11871.81	874	13.58		
	<i>Total</i>	12096.65	885			
Textverständnis	<i>Zwischen Gruppen</i>	110.80	11	10.07	0.86	.575
	<i>Innerhalb Gruppen</i>	10184.40	874	11.65		
	<i>Total</i>	10295.20	885			
Figuren lernen	<i>Zwischen Gruppen</i>	65.92	11	5.99	0.41	.951
	<i>Innerhalb Gruppen</i>	12669.78	874	14.50		
	<i>Total</i>	12735.69	885			
Fakten lernen	<i>Zwischen Gruppen</i>	285.44	11	25.95	1.86	.041
	<i>Innerhalb Gruppen</i>	12172.45	874	13.93		
	<i>Total</i>	12457.89	885			
Diagramme und Tabellen	<i>Zwischen Gruppen</i>	69.61	11	6.33	0.64	.798
	<i>Innerhalb Gruppen</i>	8681.29	874	9.93		
	<i>Total</i>	8750.90	885			
Konzentr. u. sorgf. Arbeiten	<i>Zwischen Gruppen</i>	348.25	11	31.66	2.70	.002
	<i>Innerhalb Gruppen</i>	10229.87	874	11.70		
	<i>Total</i>	10578.12	885			

Tabelle 27: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den Testlokalen

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 sinkt absolut seit 1999 immer weiter ab (Abbildung 38). Bezieht man auch die Veterinärmedizin mit ein, so ist bereits seit 1998 ein Rückgang festzustellen. In Humanmedizin sind in diesem Jahr 65.2% Frauen, in der Veterinärmedizin sind es 81.8%.

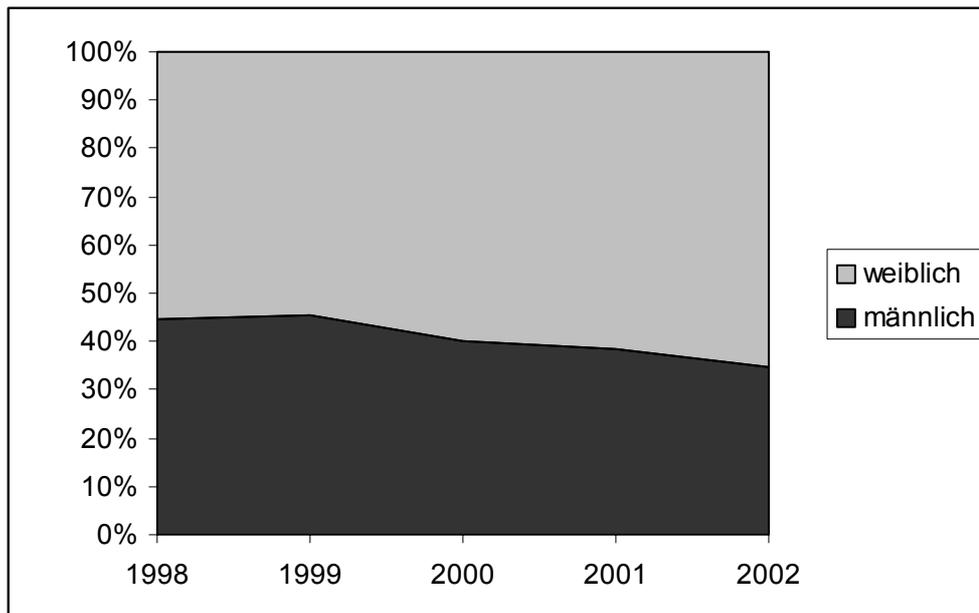


Abbildung 38: Bewerberkohorten (Humanmedizin) 1998 bis 2002 nach Geschlecht

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 wieder leicht über denjenigen der Kandidatinnen (Tabelle 28). Die kleinen Stichproben der französisch- und italienischsprachigen Gruppen lassen hier keine Aussagen zu.

Nachdem im vergangenen Jahr erstmals ein signifikanter Unterschied zwischen den Geschlechtern festgestellt wurde, ist die Differenz diesmal wie auch schon in den Vorjahren gering (Tabelle 29). Die Tendenz bleibt also bestehen, allerdings bewegt sich die Differenz in unbedeutenden Grössenordnungen. 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	102.10	92.17	100	97.18	93.2	-
	Stand.-abw.	9.62	12.49	8.85	8.60	7.36	-
	N	245	12	11	22	15	-
weiblich	Mittelwert	100.77	98.29	92.11	98	92.04	92.5
	Stand.-abw.	9.98	9.88	6.67	9.46	8.59	9.49
	N	486	7	9	133	25	8

Tabelle 28: 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	.652	.42	1.71	729	.08
	Veterinärmedizin	Varianzen homogen	.335	.56	-.38	153	.70
französisch	Humanmedizin	Varianzen homogen	.425	.52	-1.10	17	.28
	Veterinärmedizin	Varianzen homogen	1.185	.283	.43	38	.67
italienisch	Humanmedizin	Varianzen homogen	.631	.43	2.207	18	.04

Tabelle 29: Prüfung Varianzhomogenität und Mittelwert (t-Test) für Testwert zwischen Geschlechtern. Italienische Sprachgruppe Veterinärmedizin wg. Fallzahl nicht berechnet

Abbildung 39 zeigt den Vergleich zwischen männlichen Bewerbern und Kandidatinnen, aufgeteilt nach Disziplin, Abbildung 40 den Boxplot für die Testwerte nach Sprache und Geschlecht geschichtet. Die schiefen Verteilungen der französischen und italienischen Sprachgruppen kommen durch die geringen Teilnehmerzahlen zustande.

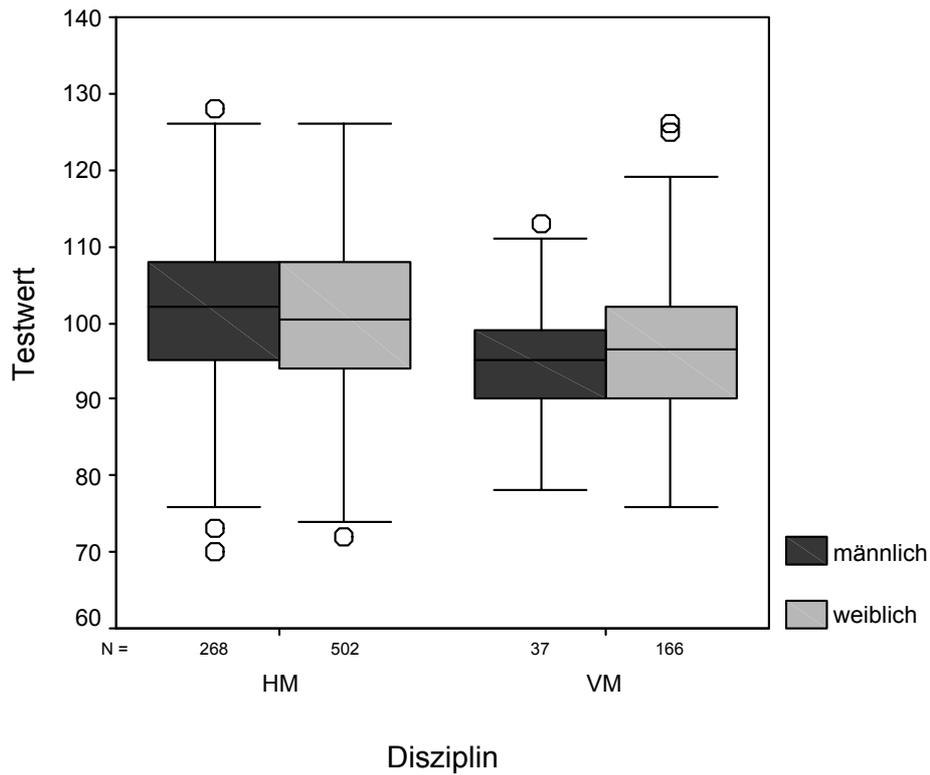


Abbildung 39: Boxplot für den Testwert nach Geschlecht und Disziplin (Humanmedizin und Veterinärmedizin)

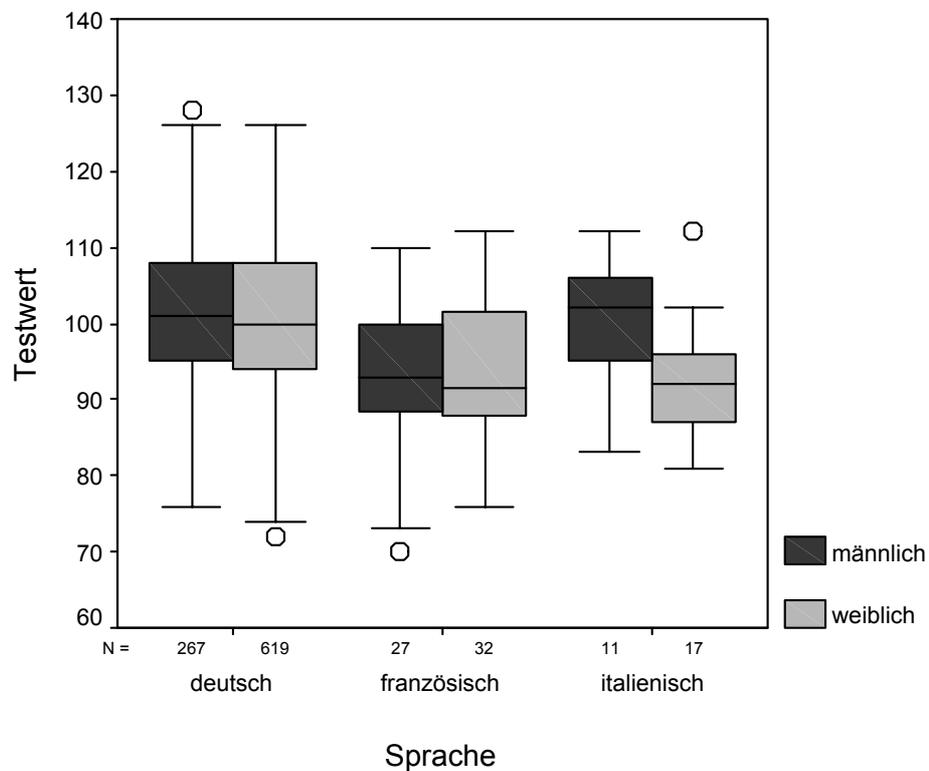


Abbildung 40: Boxplots für die Testwerte nach Geschlecht, getrennt für die drei Sprachgruppen

Wiederum zeigen sich Differenzen für die einzelnen Untertests zwischen den Geschlechtern. Ausser in den Untertests "Muster zuordnen", "Textverständnis" und "Figuren lernen" sind alle Differenzen signifikant. Die Unterschiede im Untertest "Muster zuordnen" waren als einzige in den bisherigen Durchführungen nie signifikant – in diesem Jahr erzielten hier die männlichen Teilnehmer sogar erstmals leicht bessere Resultate als die Kandidatinnen. In allen anderen Untertests bleibt bei wechselndem Ausmass die Richtung der Differenz über die Jahre konstant (Abbildung 42).

Frauen erzielen tendenziell bessere Werte in den Untertests "Figuren lernen", "Fakten lernen", "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". Im Untertest "Muster zuordnen" zeigen sich tendenziell keine Geschlechtsunterschiede.

Untertest	Geschlecht	m	s	Levene's Test Varianz		t-test Mittelwert			
					F	Sig.	t	df	Sig.
Muster zuordnen	männlich	11.38	2.71	Varianzhomogenität	1.343	.247	.363	884	.71
	weiblich	11.30	2.85						
Med.-naturwiss. Grundverst.	männlich	10.31	3.16	Varianzhomogenität	.591	.442	2.232	884	.02
	weiblich	9.80	3.05						
Schlauchfiguren	männlich	12.84	3.21	Keine Varianzhomogenität	4.670	.031	4.380	559.517	.00
	weiblich	11.78	3.58						
Quantitative und formale Probleme	männlich	12.39	3.83	Keine Varianzhomogenität	5.397	.020	6.559	466.603	.00
	weiblich	10.60	3.51						
Textverständnis	männlich	9.34	3.23	Varianzhomogenität	2.220	.137	0.700	884	.48
	weiblich	9.17	3.49						
Figuren lernen	männlich	15.57	3.85	Varianzhomogenität	.036	.850	-0.471	884	.64
	weiblich	15.70	3.77						
Fakten lernen	männlich	9.99	3.61	Varianzhomogenität	.479	.489	-3.842	884	.00
	weiblich	11.03	3.77						
Diagramme und Tabellen	männlich	11.53	3.01	Varianzhomogenität	1.598	.206	4.496	884	.00
	weiblich	10.50	3.16						
Konzent. u. sorgfältiges Arbeiten	männlich	13.06	3.55	Varianzhomogenität	.024	.876	-2.292	884	.02
	weiblich	13.63	3.41						

Tabelle 30: Mittelwerte (m) und Standardabweichungen (s) geschlechtsspezifisch und Ergebnisse der Prüfungen auf Varianzhomogenität und Mittelwertunterschiede

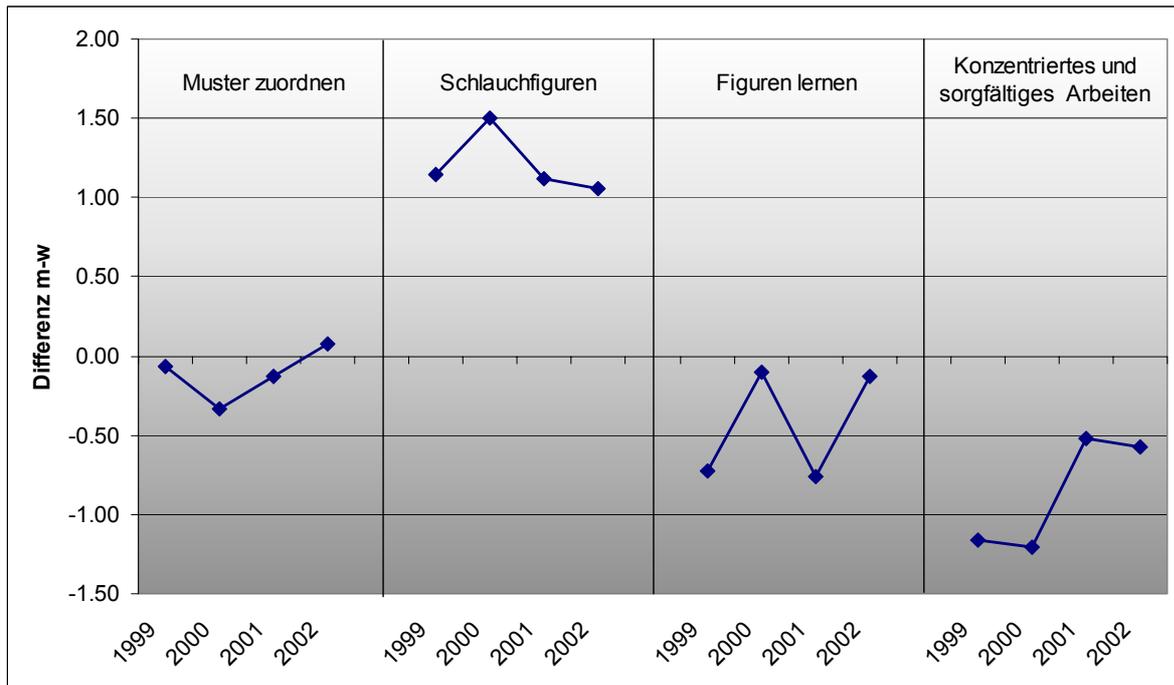


Abbildung 41: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (wenig sprachabhängige Untertests)

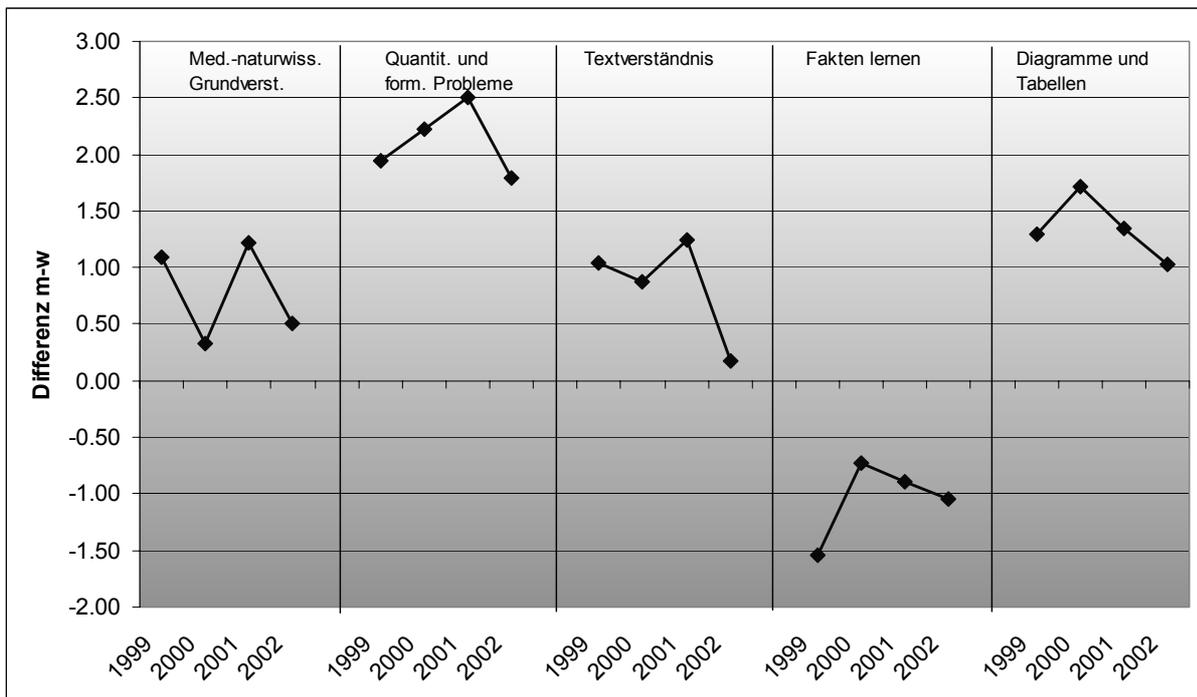


Abbildung 42: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (sprachabhängige Untertests)

Da bezüglich des Testwerts keine statistischen Geschlechtsunterschiede festgestellt wurden, weisen erwartungsgemäss auch die Zulassungsquoten keine signifikanten Differenzen auf (Tabelle 31 und Tabelle 32). Zwar liegen jene der männlichen Teilnehmer tendenziell etwas höher, jedoch nicht in einem statistisch bedeutenden Ausmass.

Disziplin	Geschlecht	Warteliste	zugelassen	Total
Humanmedizin	männlich	30 12.2%	215 87.8%	245 100.0%
	weiblich	74 15.2%	412 84.8%	486 100.0%
	Total	104 14.2%	627 85.8%	731 100.0%
Veterinärmedizin	männlich	2 9.1	20 90.9%	22 100.0%
	weiblich	8 6.0%	125 94.0%	133 100.0%
	Total	10 6.5%	145 93.5%	155 100.0%

Tabelle 31: Zulassungsquoten geschlechtsspezifisch

Disziplin		Wert	df	Asymp. Sig. (2-seitig)
Humanmedizin	Pearson Chiquadrat	1.187	1	.28
	Likelihood Ratio	1.212	1	.27
Veterinärmedizin	Pearson Chiquadrat	.296	1	.59
	Likelihood Ratio	.270	1	.60

Tabelle 32: Prüfung Unterschiede der Zulassung – geschlechtsspezifisch: keine signifikanten Unterschiede bei Humanmedizin und Veterinärmedizin

7.7 Vergleich der Kantone

Bereits in den letzten Jahren bestand ein grosses Interesse an Ergebnissen, die sich auf die Vergleiche zwischen den Kantonen beziehen. Die Ursache scheint auch ein gewisser Mangel an Vergleichsstudien im Bildungsbereich zu sein. 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 – dies würde aber Gegenstand einer eigenen Untersuchung sein müssen.

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 2002 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 1999¹ wurden die Kantone wie immer in die drei Gruppen geteilt, die hohe, mittlere und niedrige Maturitätsquoten repräsentieren (Tabelle 33).

hoch			mittel			niedrig		
Kanton	Maturitäts- quote	Anzahl	Kanton	Maturitäts- quote	Anzahl	Kanton	Maturitäts- quote	Anzahl
BS	21.2	39	SH	18.4	10	SG	12.8	59
BL	20.4	44	NW	16.5	6	SZ	12.7	15
ZH*	18.8	223	AG	15.9	60	AI	12.3	5
			GL	15.3	4	GR	12.3	22
			ZG	15	19	LU	11.5	75
			SO	13.9	43	TG	10.4	34
			AR	13.8	6	UR	10.2	5
			BE	13.1	169	OW	9.9	7
Summe		306	Summe		317	Summe		222

*= Doppeljahrgang

Tabelle 33: Gruppierung der Kantone nach der Maturitätsquote (Basis 1999)

Erstmals zeigt sich bezüglich des Testwerts ein Unterschied zwischen den nach Maturitätsquote sortierten Kantonen. Dieser ist allerdings von geringem Ausmass und ausserdem entgegen der Erwartung. Die Kantone mit hoher Maturitätsquote schneiden bezüglich des Testwerts insgesamt am besten ab (Tabelle 34). Die Analyse der Kantone weiter unten wird zeigen, dass es sich mit Zürich, Basel-Land und Basel-Stadt um die Kantone mit den höchsten, zweithöchsten und niedrigsten Ergebnissen handelt. Es gibt weiterhin 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.

¹ Zahlen aus: Bundesamt für Statistik (1999). Maturitäten 1998: Statistik der Schweiz Bd. 15, Tabelle 4, Seite 13.

	Maturitäts- quote	Mittel- wert	Stand.- abw.		Quadrat- summe	df	MQ	F	Sig.
Testwert	hoch	101.92	10.703	Zwischen Gruppen	739.502	2	369.751	3.736	.02
	mittel	99.78	9.578						
	niedrig	100.42	9.370						
Muster zuordnen	hoch	11.65	2.882	Zwischen Gruppen	50.458	2	25.229	3.173	.04
	mittel	11.25	2.785						
	niedrig	11.06	2.782						
Medizin.-natur- wiss. Grundver- ständnis	hoch	10.25	3.331	Zwischen Gruppen	38.940	2	19.470	2.012	.13
	mittel	9.76	2.987						
	niedrig	9.95	2.966						
Schlauchfigure n	hoch	12.50	3.644	Zwischen Gruppen	77.023	2	38.511	3.042	.04
	mittel	11.89	3.562						
	niedrig	11.86	3.430						
Quantitat. u.formale Probleme	hoch	11.28	3.904	Zwischen Gruppen	77.264	2	38.632	2.812	.06
	mittel	10.82	3.667						
	niedrig	11.56	3.475						
Textverständnis	hoch	9.35	3.530	Zwischen Gruppen	5.144	2	2.572	.221	.80
	mittel	9.22	3.305						
	niedrig	9.16	3.408						
Figuren lernen	hoch	15.99	3.885	Zwischen Gruppen	47.916	2	23.958	1.626	.19
	mittel	15.48	3.814						
	niedrig	15.52	3.810						
Fakten lernen	hoch	11.08	3.908	Zwischen Gruppen	71.125	2	35.563	2.535	.08
	mittel	10.41	3.812						
	niedrig	10.64	3.405						
Diagramme und Tabellen	hoch	10.79	3.357	Zwischen Gruppen	2.341	2	1.171	.118	.88
	mittel	10.83	2.950						
	niedrig	10.92	3.122						
Konzentr. u. sorgf. Arb.	hoch	13.99	3.704	Zwischen Gruppen	152.826	2	76.413	6.239	.00
	mittel	13.02	3.564						
	niedrig	13.32	3.093						

Tabelle 34: Mittelwerte und Standardabweichungen für drei Gruppen der Kantone nach der Maturitätsquote, keine signifikanten Unterschiede für Testwert und die Untertests

Für Vergleiche zwischen den Kantonen werden wiederum wie in den Vorjahren die gleichen, zahlenmässig stärksten Kantone berücksichtigt (Tabelle 35). Die Abbildung 43 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. Zum zweiten Mal in Folge weist Basel-Land, nach drei eher punkteschwachen ersten Jahren, das zweitbeste Resultat auf. Eine geringfügige Verbesserung (allerdings nicht rangmässig) gegenüber dem letzten Jahr kann auch in Basel-Stadt beobachtet werden. Der Kanton Aargau hat sich dieses Jahr verschlechtert. Alle anderen erreichten ungefähr das Niveau des Vorjahres (Abbildung 43).

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	99.75	11.10	9.42	11.80	11.10	9.50	14.43	11.33	10.98	13.0
	S	9.07	2.815	3.201	3.473	3.328	3.181	4.163	3.672	2.914	3.90
BE	M	99.73	11.47	9.78	12.08	10.62	9.09	15.65	10.23	10.79	12.9
	S	9.60	2.621	2.914	3.402	3.753	3.298	3.649	3.952	2.952	3.27
BL	M	102.00	12.02	10.16	12.73	11.41	8.73	16.20	11.32	10.98	13.5
	S	10.53	3.586	3.504	3.799	3.737	3.493	4.072	4.258	3.159	4.12
BS	M	96.05	10.44	9.33	11.69	9.69	8.38	15.05	9.85	9.41	11.6
	S	11.17	2.479	3.896	3.054	3.894	3.958	4.186	3.617	3.676	3.64
LU	M	100.27	10.57	10.04	12.32	11.24	9.29	15.64	10.71	10.84	13.0
	S	9.86	3.317	3.095	3.068	3.601	3.400	3.906	3.755	2.871	3.19
SG	M	101.85	11.66	10.14	12.54	12.41	8.88	15.20	10.71	11.63	13.6
	S	8.86	2.397	2.968	3.436	3.130	3.563	3.999	3.464	3.118	2.94
ZH	M	102.93	11.79	10.43	12.60	11.53	9.64	16.12	11.25	11.00	14.4
	S	10.37	2.750	3.176	3.704	3.890	3.428	3.788	3.862	3.294	3.46

Tabelle 35: Vergleich der Kantone für Testwert und die einzelnen Untertests

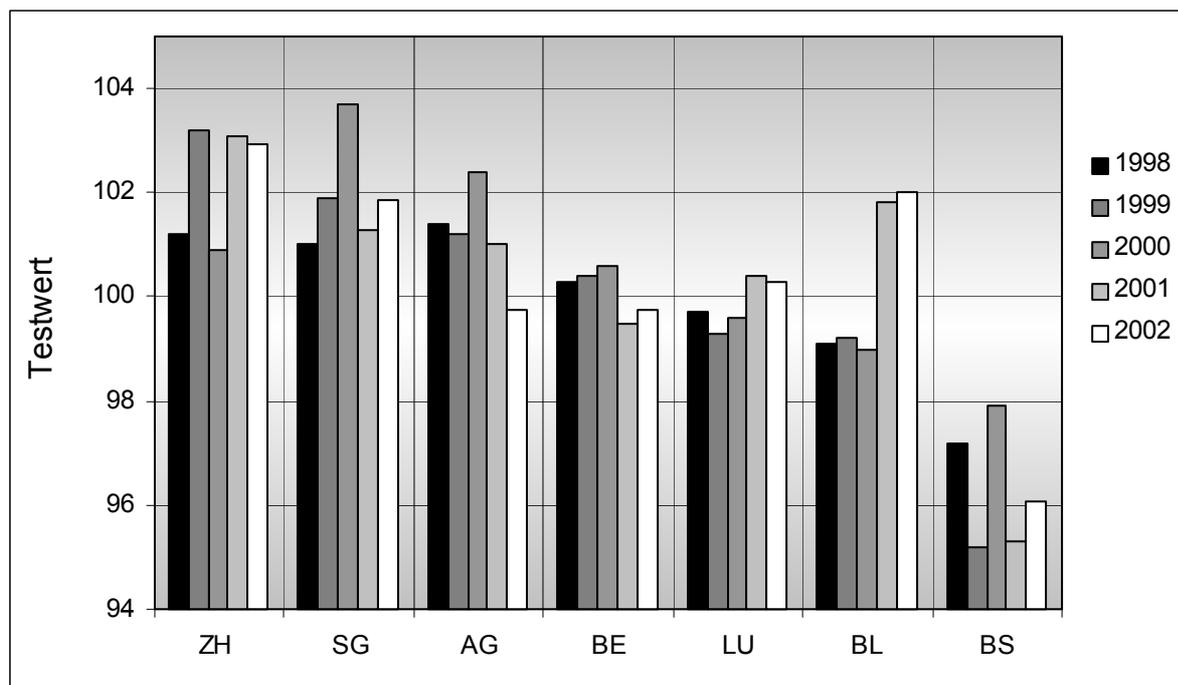


Abbildung 43: Testwert für Kantone im Vergleich

7.8 Vergleiche für Altersgruppen

Bisher wurde jedes Jahr nachgewiesen, dass Personen, die älter sind, etwas geringere Zulassungschancen haben. Dies war besonders der Fall, wenn die Personen die Maturitätsprüfung relativ spät abgelegt haben, also der Schulbesuch nicht zeitlich zusammenhängend erfolgte. Diese Personengruppe wird aus verschiedensten Gründen nicht frühestmöglich die Mittelschule besucht haben.

Für die Überprüfung der Alterseffekte im Jahr 2002 wurden die Gruppen auf der Basis der Tabelle 36 gebildet. Die Altersgruppen wurden um 1 Jahr gegenüber 2001 verschoben. Die älteste Gruppe wurde nahe dem Median des Maturitätsalters geteilt. Zusätzlich ist der Doppeljahrgang zu beachten, wodurch jüngere Bewerber (theoretisch bis Jahrgang 1986) berücksichtigt werden müssen.

Maturitätsjahr	Geburtsjahr			Total
	1954-1978	1979-1981	1982-1986	
1974	1			1
1983	1			1
1984	1			1
1986	2			2
1989	2			2
1990	1			1
1991	1			1
1992	1			1
1994	3			3
1995	1			1
1996	8			8
1997	5	1		6
1998	7	2		9
1999	5	18		23
2000	9	37	1	47
2001	6	114	86	206
2002	28	121	511	660
	82	293	598	973

Tabelle 36: Gruppenbildung für das Jahr der Maturitätsprüfung bezogen auf die Geburtsjahre

Erneut zeigen die älteren Jahrgänge mit später Maturität tiefere Testwerte als alle anderen Gruppen (Abbildung 44). Diese liegen innerhalb eines Bereichs von rund 1.5 Punkten (Tabelle 37). Die gleiche rangmässige Konstellation zeigt sich auch in allen Untertests, wobei die Differenzen bei "Fakten lernen" sowie "Konzentriertes und sorgfältiges Arbeiten" nicht signifikant sind.

	Geburtsjahr und Maturität	Mittelwert	Standard-abw.	Quadrat-summe	df	MQ	F	Sig.
Testwert	54-77, frühe Matur.	100.38	11.018	2920.313	3	973.438	10.04	.00
	54-77, späte Matur.	93.00	9.471					
	1978-80	99.39	10.266					
	1981-85	100.87	9.598					
Muster zuordnen	54-77, frühe Matur.	10.74	3.222	118.143	3	39.381	4.88	.00
	54-77, späte Matur.	9.85	2.989					
	1978-80	11.12	2.778					
	1981-85	11.39	2.836					
Med.-naturwiss. Grundverständnis	54-77, frühe Matur.	10.38	3.201	174.075	3	58.025	6.03	.00
	54-77, späte Matur.	8.13	2.606					
	1978-80	9.61	3.097					
	1981-85	9.97	3.136					
Schlauchfiguren	54-77, frühe Matur.	11.38	3.676	115.314	3	38.438	3.16	.02
	54-77, späte Matur.	10.58	3.512					
	1978-80	12.07	3.464					
	1981-85	12.09	3.489					
Quantitative und formale Probleme	54-77, frühe Matur.	11.91	3.888	260.655	3	86.885	6.35	.00
	54-77, späte Matur.	9.10	3.422					
	1978-80	10.80	3.751					
	1981-85	11.29	3.685					
Textverständnis	54-77, frühe Matur.	9.09	3.728	209.468	3	69.823	6.13	.00
	54-77, späte Matur.	7.60	3.044					
	1978-80	8.75	3.460					
	1981-85	9.43	3.335					
Figuren lernen	54-77, frühe Matur.	15.29	4.373	265.270	3	88.423	5.96	.00
	54-77, späte Matur.	13.35	3.750					
	1978-80	15.14	4.040					
	1981-85	15.67	3.734					
Fakten lernen	54-77, frühe Matur.	10.06	3.709	57.094	3	19.031	1.36	.25
	54-77, späte Matur.	9.58	3.758					
	1978-80	10.60	3.768					
	1981-85	10.63	3.724					
Diagramme und Tabellen	54-77, frühe Matur.	11.03	3.424	128.650	3	42.883	4.37	.00
	54-77, späte Matur.	9.17	2.644					
	1978-80	10.65	3.152					
	1981-85	10.84	3.145					
Konzentriertes und sorgfältiges Arbeiten	54-77, frühe Matur.	13.47	4.099	49.596	3	16.532	1.34	.26
	54-77, späte Matur.	12.27	3.648					
	1978-80	13.17	3.724					
	1981-85	13.30	3.354					

Tabelle 37: Testwert und Punktwerte für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung; Varianzanalytische Prüfung des Unterschiedes

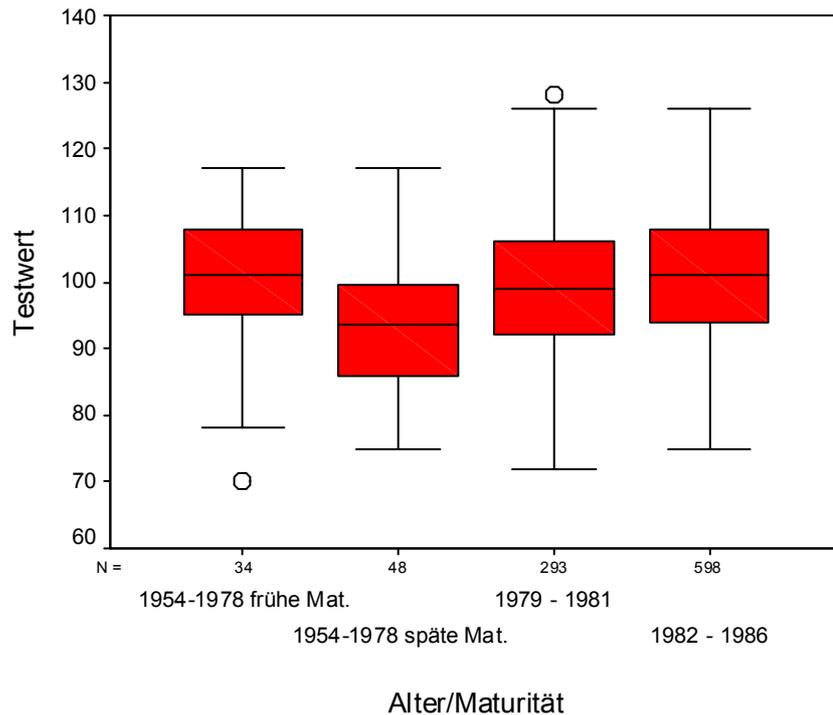


Abbildung 44: Boxplots für Testwert der Altersgruppen (Geburtsjahre) und Maturität (Mat.)

Eine Betrachtung der Zulassungsquoten (Tabelle 38) zeigt einen Wert von 89% für die Gruppe der jüngsten Teilnehmer, rund 85.5% für die zweitjüngste und die älteste Gruppe mit früher Maturität, und 60.4% für die älteren Jahrgänge mit später Maturität. Damit schneidet die ältere Gruppe mit später Maturität auch hier erwartungsgemäss schlechter ab als die anderen Kandidaten.

Geburtsjahr und Maturität	Warteliste	zugelassen	n
1954-1978, frühe Maturität	5 14.7%	29 85.3%	34
1954-1978, späte Maturität	19 39.6%	29 60.4%	48
1979 - 1981	42 14.3%	251 85.7%	293
1982 - 1986	66 11.0%	532 89.0%	598
Total	132 13.6%	841 86.4%	973

Tabelle 38: Zulassungsquoten für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung

Innerhalb der jüngsten Gruppe sind dieses Jahr keine Leistungsdifferenzen hinsichtlich des Testwerts oder der einzelnen Untertests zu finden.

7.9 Vergleiche nach Wunschuniversitäten

Die Zulassung zu den Universitäten erfolgt aufgrund des Testwertes (Tabelle 39).

Wie auch in den vorangegangenen Jahren unterscheiden sich die Wunschuniversitäten nicht signifikant hinsichtlich ihrer Zulassungsquoten.

Disziplinen	Universität	Zulassung		Total
		Warteliste	zugelassen	
Humanmedizin	Bern	29	200	229
		12.7%	87.3%	100.0%
	Basel	28	149	177
		15.8%	84.2%	100.0%
	Freiburg	13	48	61
	21.3%	78.7%	100.0%	
Humanmedizin	Zürich	45	257	302
		14.9%	85.1%	100.0%
	Total	115	655*	770
		14.9%	85.1%	100.0%
	Veterinärmedizin	Bern	10	94
9.6%			90.4%	100.0%
Zürich		7	92	99
	7.1%	92.9%	100.0%	
Veterinärmedizin	Total	17	186	203
		8.4%	91.6%	100.0%

* 1 Anmeldung für Neuenburg

Tabelle 39: Zulassungsquoten für Disziplin und Wunschuniversität (erste Wahl)

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. Die Konsistenz des Untertests "Quantitative und formale Probleme" liegt höher als diejenige in Deutschland.

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

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

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

Tabelle 40: Zuverlässigkeit 1998 bis 2002; Vergleich der Skalenkonsistenzen 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 0.91 weiterhin in einem sehr hohen Bereich und in der aus Deutschland bekannten Grössenordnung.

Andererseits ist bei einigen Untertests dieses Jahr eine deutliche Verringerung der Reliabilität zu beobachten. Trotzdem würde ein Weglassen von einzelnen Untertests zu einer Verringerung der Gesamtkonsistenz des Testprofils führen (Tabelle 41). Diese fällt mit 0.80 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
Muster zuordnen	.82	.80	.80	.78	.79
Med.-nat. Grundverständnis	.79	.78	.78	.77	.77
Schlauchfiguren	.80	.79	.79	.78	.78
Quant. u. form. Probleme	.80	.79	.78	.78	.77
Textverständnis	.80	.79	.78	.77	.78
Figuren lernen	.81	.80	.80	.78	.78
Fakten lernen	.82	.81	.80	.79	.79
Diagramme und Tabellen	.80	.78	.78	.77	.77
Konz. u. sorgf. Arbeiten	.82	.81	.80	.79	.79
TESTPROFIL	.83	.81	.81	.80	.80

Tabelle 41: 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 liegen dieses Jahr sogar ausnahmslos über 0.57 (Tabelle 42). 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		0.25	0.35	0.24	0.23	0.30	0.31	0.28	0.42	0.57
Med.-naturw. Grundverst.	0.25		0.28	0.51	0.60	0.30	0.25	0.50	0.19	0.67
Schlauchfiguren	0.35	0.28		0.32	0.27	0.36	0.28	0.32	0.41	0.64
Quant. u. formale Probl.	0.24	0.51	0.32		0.54	0.26	0.20	0.56	0.19	0.68
Textverständnis	0.23	0.60	0.27	0.54		0.24	0.21	0.51	0.12	0.65
Figuren lernen	0.30	0.30	0.36	0.26	0.24		0.40	0.26	0.47	0.65
Fakten lernen	0.31	0.25	0.28	0.20	0.21	0.40		0.18	0.41	0.58
Diagramme und Tabellen	0.28	0.50	0.32	0.56	0.51	0.26	0.18		0.15	0.65
Konzent. u. sorgfält. Arbeiten	0.42	0.19	0.41	0.19	0.12	0.47	0.41	0.15		0.59
Punktwert	0.57	0.67	0.64	0.68	0.65	0.65	0.58	0.65	0.59	

Tabelle 42: Korrelationen zwischen Punktwerten und Untertests

Abbildung 45 zeigt, dass die Grössenordnung der Korrelationen mit der Gesamtpunktzahl für die einzelnen Untertests über die Jahre stabil bleibt.

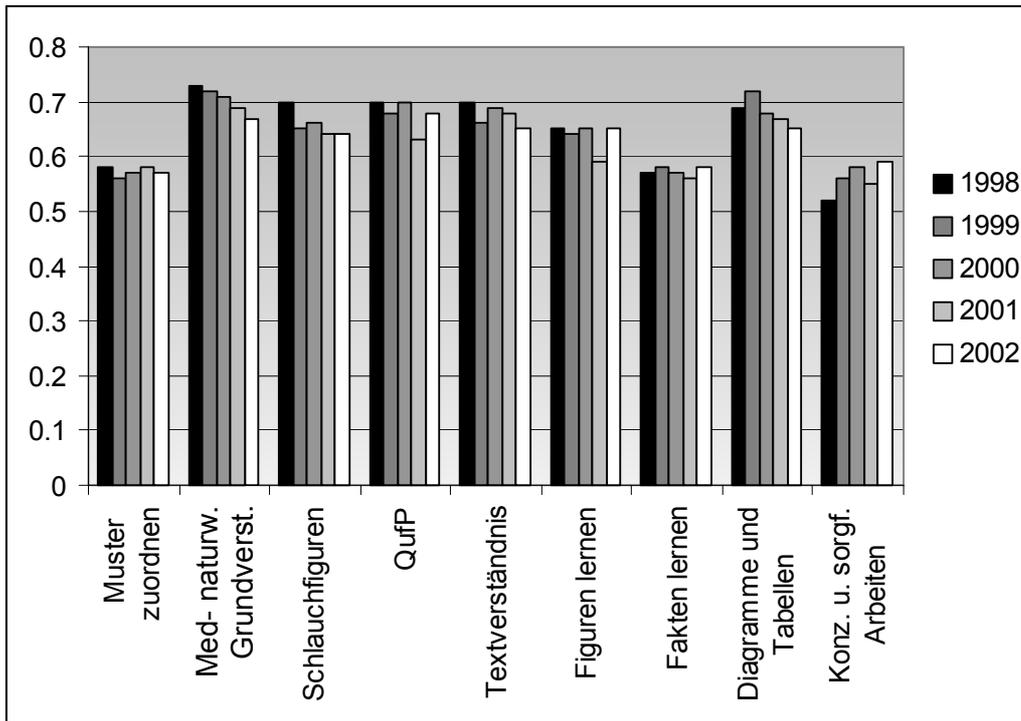


Abbildung 45: Korrelationen zwischen Punktwerten und Untertests nach Jahren

Die geringsten Zusammenhänge sind jeweils unter den Untertests "Muster zuordnen", "Fakten lernen" sowie "Konzentriertes und sorgfältiges Arbeiten" zu finden. Auch hier liegt der Wert jedoch durchwegs über 0.5. Die bereits bekannte Drei-Faktoren-Struktur bleibt auch in diesem Jahr stabil. Die Ladungsmuster stimmen weiterhin gut überein. Tabelle 43 zeigt die Vergleichbarkeit der unrotierten Lösungen über die letzten vier Jahre. Erneut laden alle Untertests signifikant im ersten unrotierten Faktor (Tabelle 45).

Unrotierte Lösung												
Faktor	Eigenwert				% Varianz				Kumuliert %			
	2002	2001	2000	1999	2002	2001	2000	1999	2002	2001	2000	1999
1	3.61	3.53	3.70	3.67	40.14	39.21	41.12	40.78	40.14	39.21	41.12	40.78
2	1.54	1.63	1.46	1.36	17.12	18.14	16.18	15.16	57.26	57.35	57.30	55.94
3	0.77	.85	.84	.91	8.53	9.39	9.31	10.16	65.79	66.74	66.61	66.10
4	0.68	.62	.66	.64	7.51	6.93	7.34	7.10	73.30	73.67	73.95	73.20
5	0.58	.57	.58	.59	6.41	6.35	6.40	6.50	79.71	80.03	80.35	79.70
6	0.53	.55	.53	.54	5.94	6.06	5.93	5.95	85.64	86.09	86.28	85.65
7	0.49	.50	.47	.53	5.40	5.53	5.23	5.86	91.05	91.61	91.51	91.50
8	0.42	.42	.39	.42	4.67	4.63	4.31	4.64	95.72	96.24	95.82	96.15
9	0.39	.34	.38	.35	4.28	3.76	4.18	3.85	100.00	100.00	100.00	100.00

Tabelle 43: Verlauf der unrotierten Lösung

In Tabelle 44 sind die rotierten Zwei- und Dreifaktorenlösungen dargestellt. Die Ladungen der einzelnen Untertests können Tabelle 46, beziehungsweise Tabelle 47 entnommen werden. Die Zweifaktorenlösung klärt 57% der Gesamtvarianz auf, die Dreifaktorenlösung knapp 66%.

Varimax-rotierte Lösungen											
Eigenwert				% Varianz				Kumuliert %			
2002	2001	2000	1999	2002	2001	2000	1999	2002	2001	2000	1999
Zwei-Faktorenlösung											
2.664	2.717	2.744	2.723	29.597	30.191	30.491	30.254	29.597	30.191	30.491	30.254
2.490	2.444	2.412	2.312	27.663	27.155	26.804	25.687	57.260	57.346	57.295	55.940
Drei-Faktorenlösung											
2.614	2.706	2.711	2.683	29.045	30.064	30.123	29.806	29.045	30.064	30.123	29.806
1.667	1.818	1.916	1.811	18.528	20.205	21.289	20.122	47.573	50.269	51.411	49.928
1.640	1.482	1.368	1.455	18.217	16.470	15.197	16.172	65.790	66.739	66.608	66.100

Tabelle 44: Varianzanteile der einzelnen Faktorenlösungen (rotierte Lösungen)

Im "Scree-Plot" (Abbildung 46) wird der Eigenwerteverlauf der Faktoren dargestellt. Der Verlauf für den EMS 2002 befindet sich links, rechts davon zum Vergleich der fast identische Verlauf aus dem Vorjahr. Die Ein-Faktorenlösung ist erneut deutlich hervorgehoben, die Zwei-Faktorenlösung und bedingt auch die Drei-Faktorenlösung unterscheiden sich weniger von den jeweils nachfolgenden. Ab der Drei-Faktorenlösung sind die Eigenwerte kleiner als 1 – der Varianzanteil des Faktors ist kleiner als der eines einzelnen Untertests.

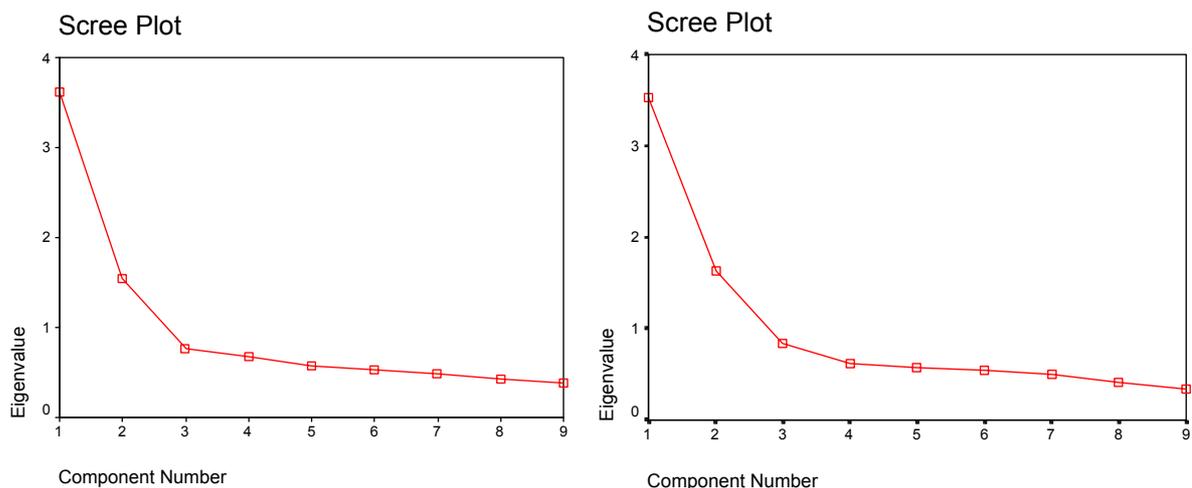


Abbildung 46: Scree-Plot (Eigenwerteverlauf) der Faktoren 1 bis 9 für die Untertests des EMS 2002 (links) und 2001 (rechts)

Die Abbildung 47 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.

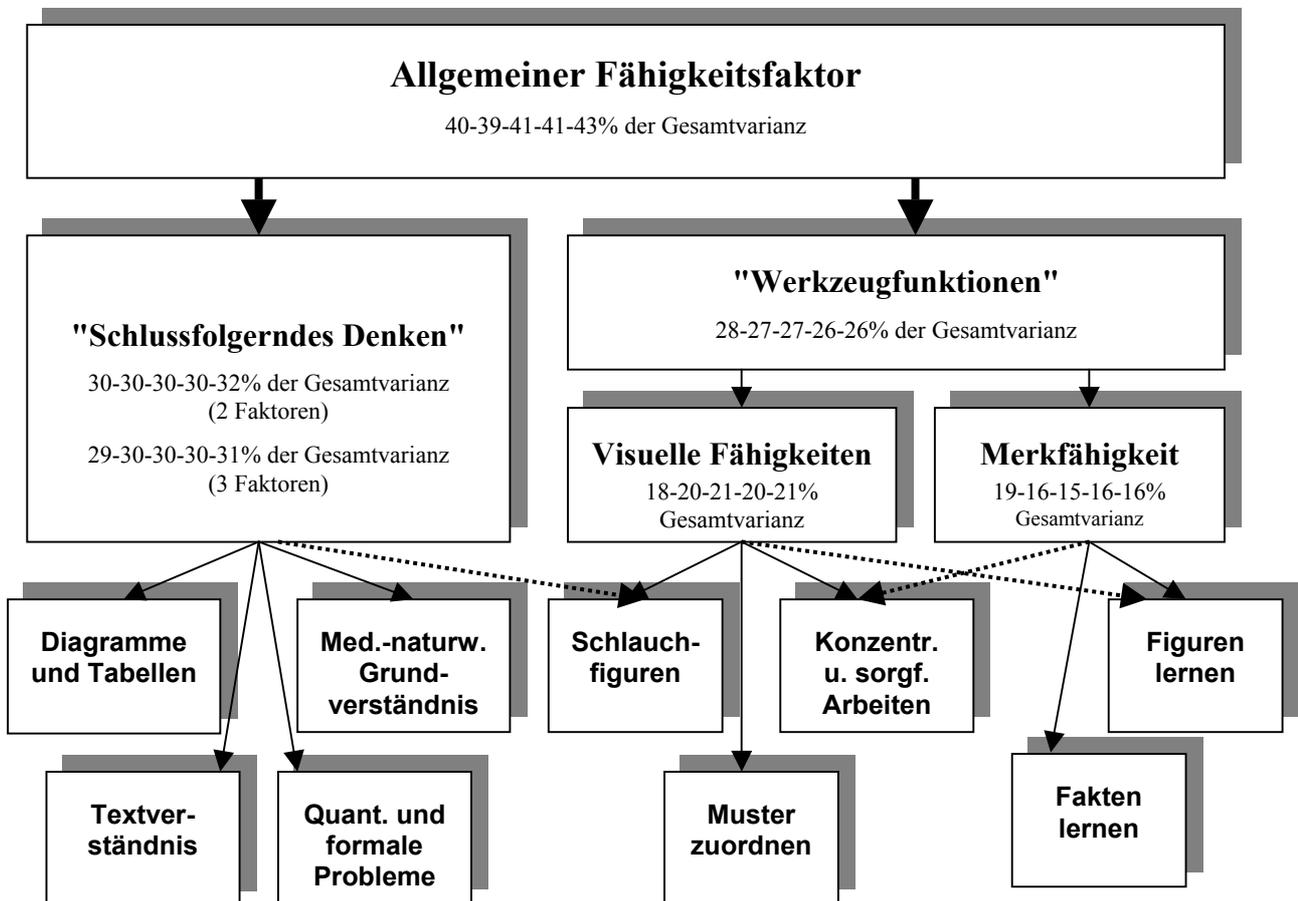


Abbildung 47: Struktur der Untertests des EMS, hierarchische Faktorenstruktur (2002 erster Wert, 2001 zweiter, 2000 dritter, 1999 vierter und 1998 fünfter Wert der Varianzaufklärung durch den Faktor)

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

Tabelle 45: Ladungen im ersten Faktor der unrotierten Lösung 1998 bis 2002 "Generalfaktor" bzw. allgemeiner Fähigkeitsfaktor)

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

Tabelle 46: Ladungen und Kommunalitäten der Zwei-Faktorenlösung, varimaxrotiert

In der Zwei-Faktorenlösung zeigen sich "Schlussfolgerndes Denken" und "Werkzeugfunktionen" als Faktoren. Der Untertest "Textverständnis" 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 ²		1	2	3
	2001	2002	2001	2002	2001	2002	2001	2002			
Diagramme und Tabellen	.77	.76	.13	.26	.17		.64	.65	.82		
Medizin.-naturwiss. Grundverständnis	.82	.78	.15		.12	.23	.70	.67	.81		.20
Quantitat. u. formale Probleme	.81	.78	.11	.20			.66	.65	.80	.18	
Textverständnis	.82	.82			.15	.13	.70	.70	.79		.20
Muster zuordnen	.16	.15	.80	.78	.10	.14	.68	.65		.81	
Konzentr. u. sorgfält. Arbeiten			.68	.59	.40	.58	.63	.68		.70	.40
Schlauchfiguren	.25	.25	.75	.70	.11	.18	.64	.59	.35	.71	
Fakten lernen	.14	.13	.11		.85	.83	.76	.71	.21		.87
Figuren lernen	.15	.20	.28	.25	.71	.72	.61	.62	.13	.47	.64

Tabelle 47: 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 "Textverständnis", "Muster zuordnen" und "Fakten lernen".

8.3 Item-Trennschärfen

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

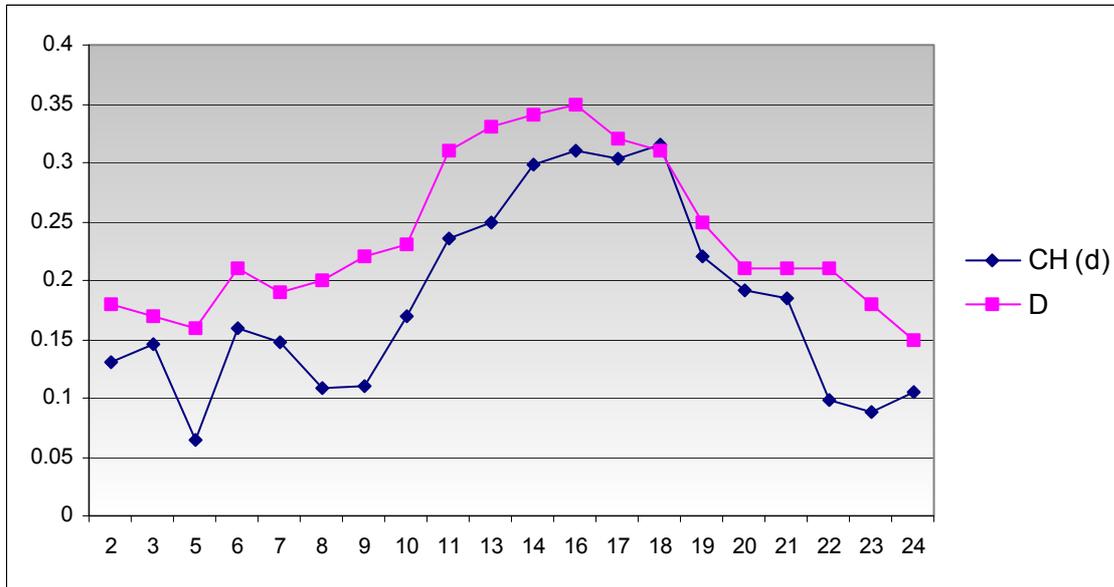


Abbildung 48: Trennschärfen für den Untertest "Muster zuordnen"

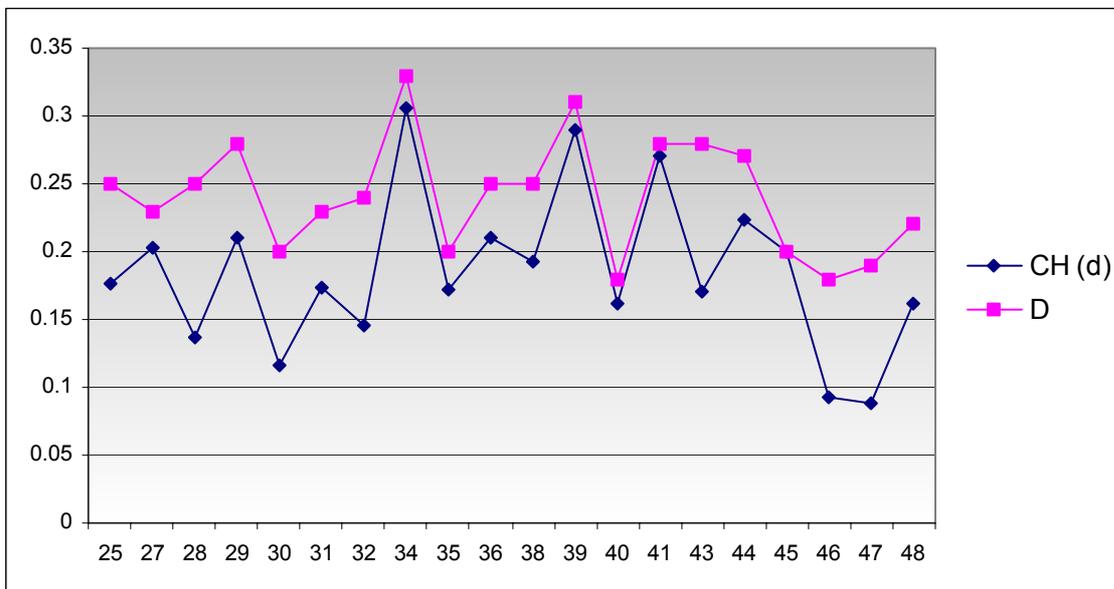


Abbildung 49: Trennschärfen für den Untertest "Medizinisch-naturwissenschaftliches Grundverständnis"

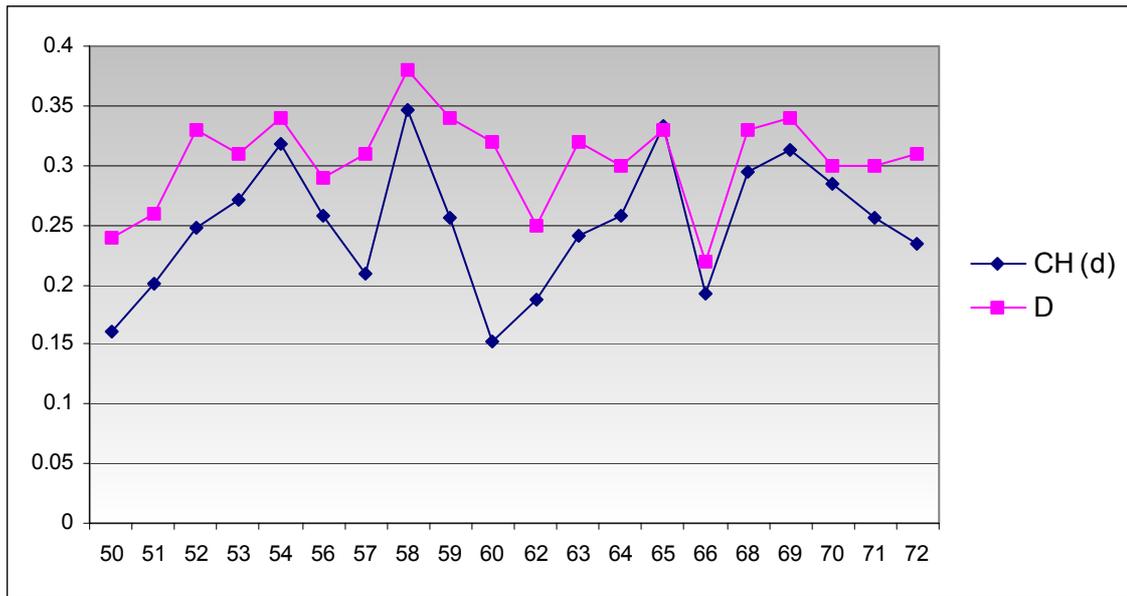


Abbildung 50: Trennschärpen für den Untertest "Schlauchfiguren"

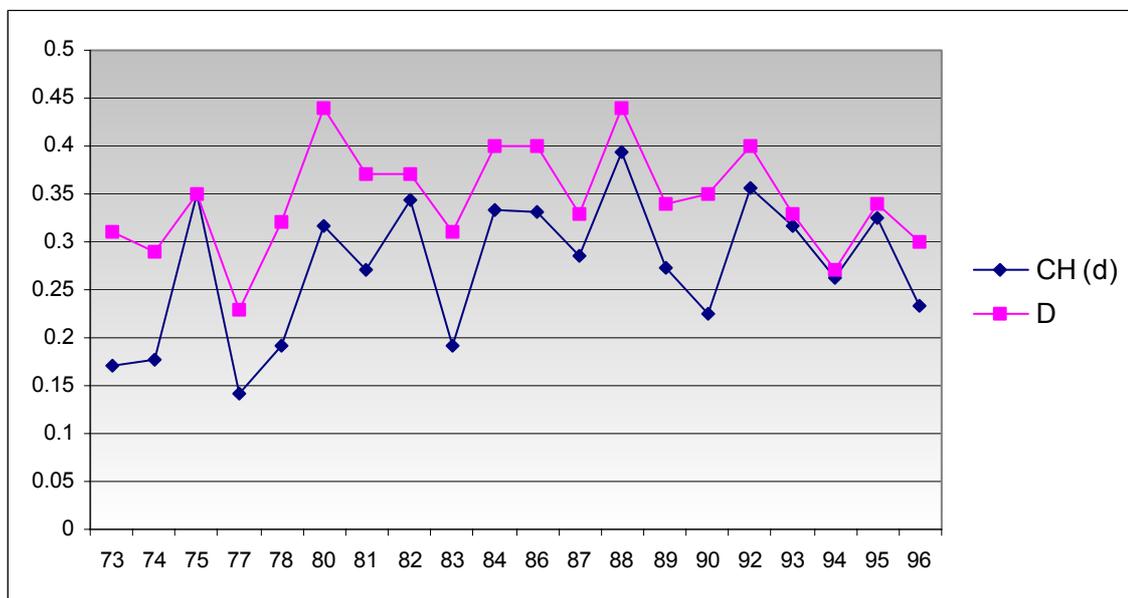


Abbildung 51: Trennschärpen für den Untertest "Quantitative und formale Probleme"

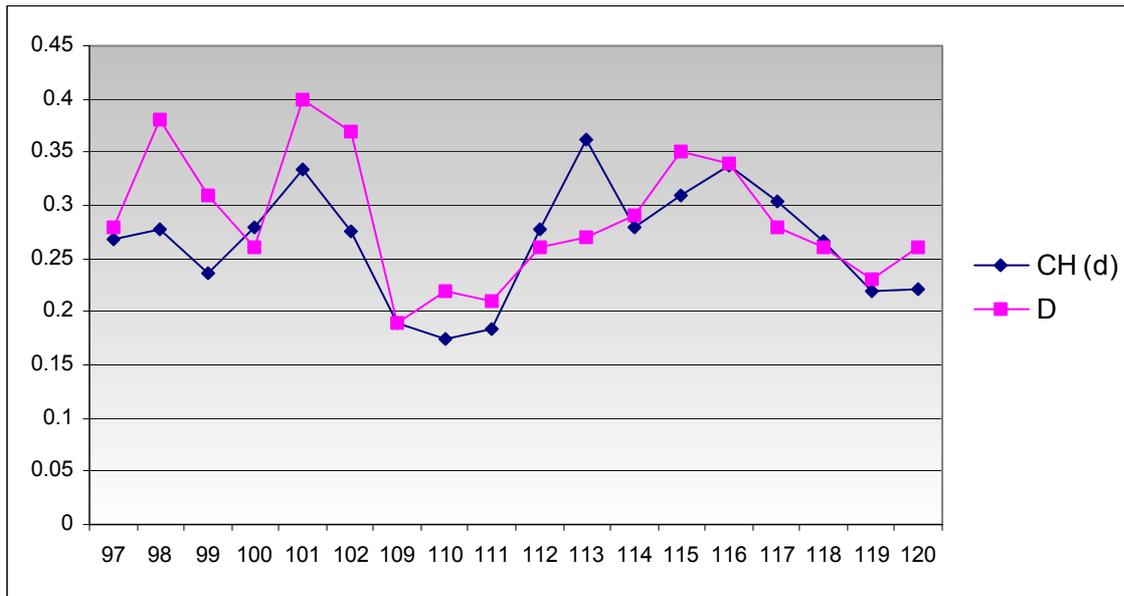


Abbildung 52: Trennschärfen für den Untertest "Textverständnis"

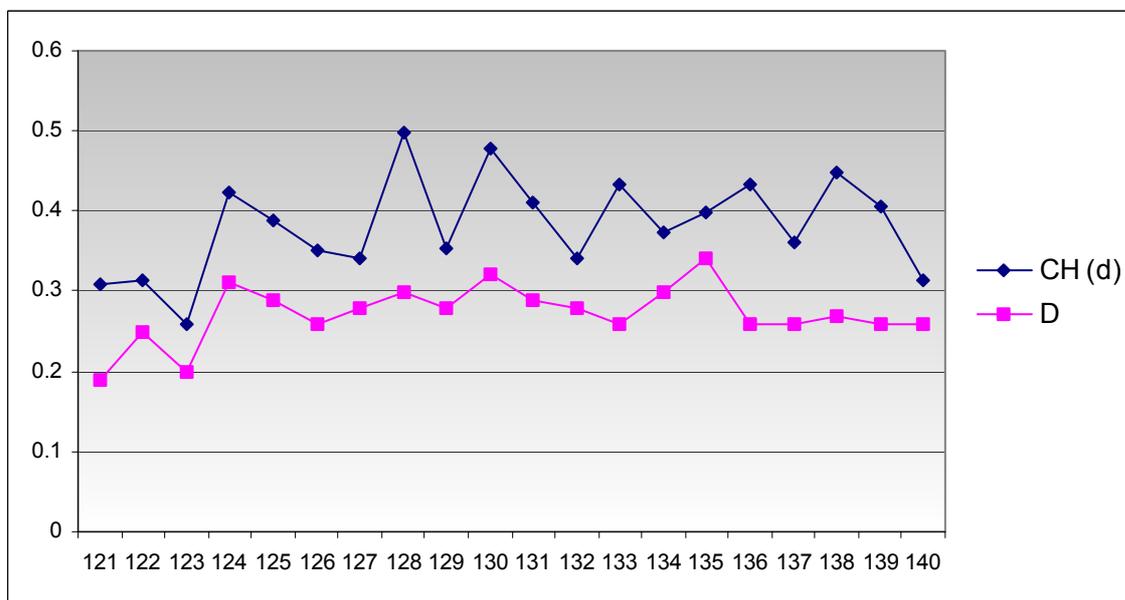


Abbildung 53: Trennschärfen für den Untertest "Figuren lernen"

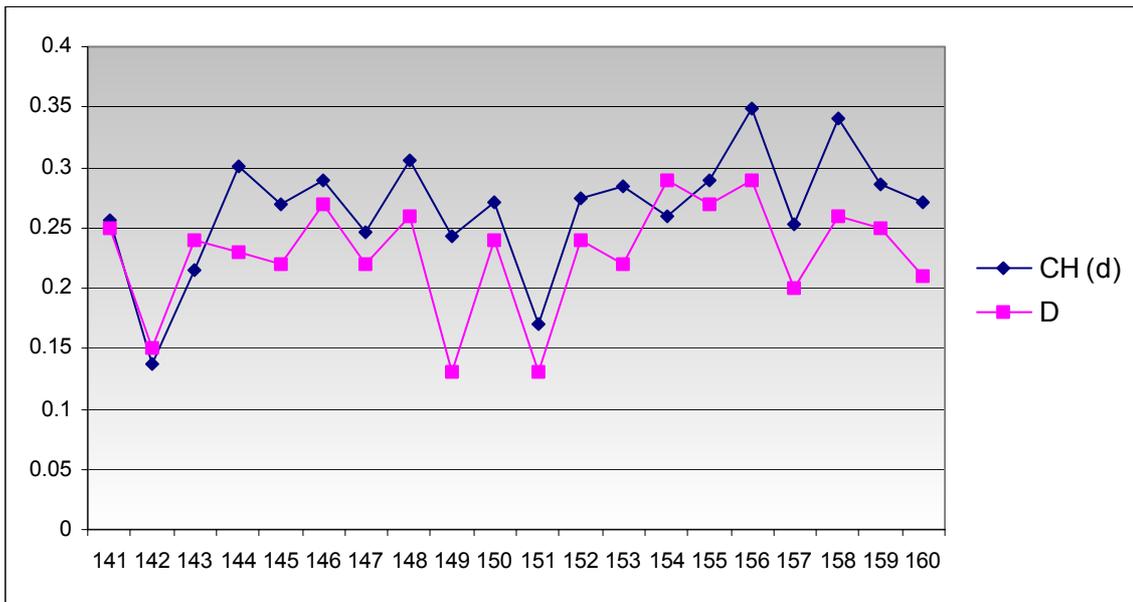


Abbildung 54: Trennschärfen für den Untertest "Fakten lernen"

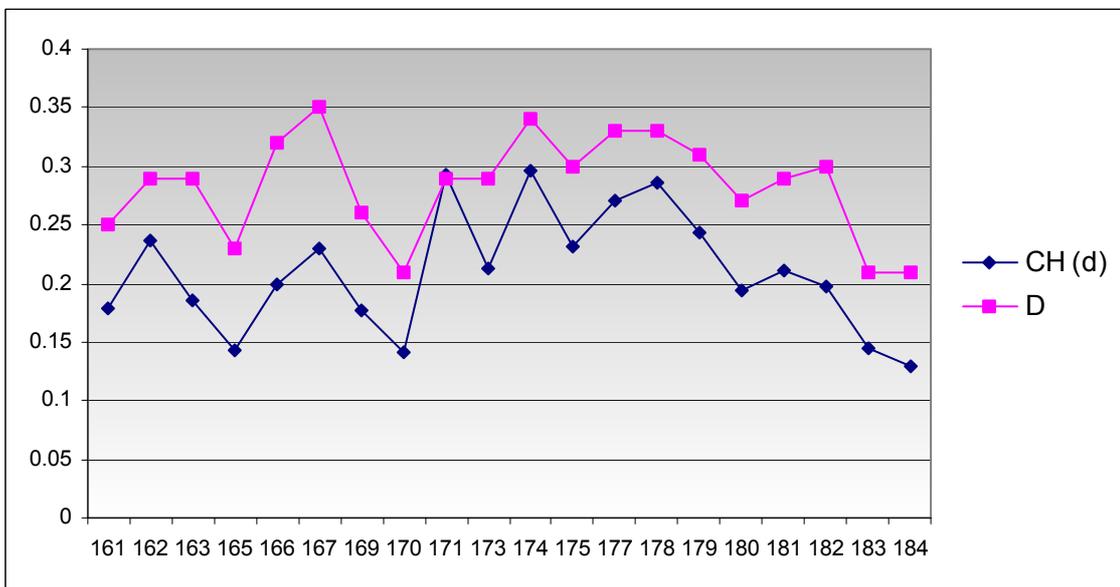


Abbildung 55: Trennschärfen für den Untertest "Diagramme und Tabellen"

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 48).

	Deutsch		Französisch		Italienisch	
	m	s	m	s	m	s
2002						
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
2001						
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
2000						
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
1999						
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
1998						
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 48: Kennwerte für den Untertest "Konzentriertes und sorgfältiges Arbeiten"

Ein signifikanter Zusammenhang besteht dieses Jahr zwischen der Anzahl richtig und fälschlicherweise markierter Zeichen. Mit 0.11 fällt dieser jedoch gering aus. 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 0.25 signifikant untereinander.

	Falsch	...übersehen	...fälschlich durchgestrichen
Richtig	.01 <i>.80</i>	-.03 <i>.31</i>	.11 <i>.00</i>
Falsch		.95 .00	.56 .00
...übersehen			.25 .00

Tabelle 49: Korrelationen der Parameter im Konzentrationstest; kursiv: Signifikanzniveau

Seit 1999 steigt die durchschnittliche Anzahl richtig bearbeiteter Zeichen in der deutschsprachigen Teilnehmergruppe stetig an (Abbildung 56). Die Leistung der italienischsprachigen Kandidaten bleibt diesbezüglich konstant. Die französischsprachigen Teilnehmer haben sich dieses Jahr erstmals nach zwei Durchführungen wieder leicht verschlechtert. Die beste Leistung wurde dort 1998 erreicht.

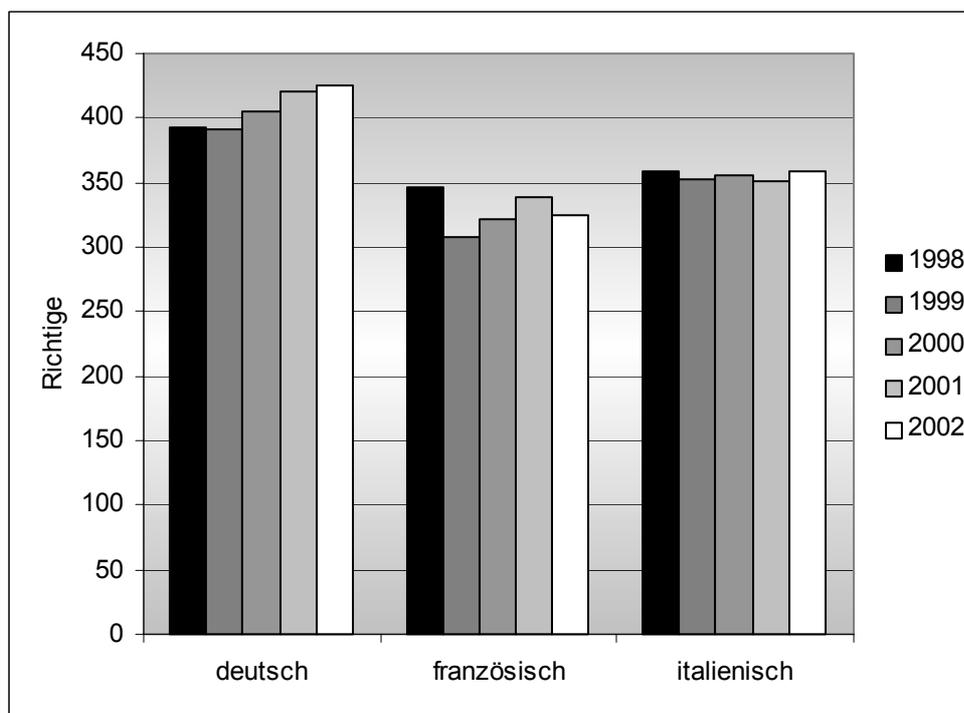


Abbildung 56: Anzahl richtig bearbeiteter Zeichen im Untertest "Konzentriertes und sorgfältiges Arbeiten", nach Jahren und Sprachgruppen

Betrachtet man die durchschnittliche Anzahl falsch bearbeiteter Zeichen, so zeigt sich für die deutschsprachige Gruppe seit 1998 ein Anstieg. Die französische Gruppe schneidet dieses Jahr diesbezüglich schlecht ab – 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 demjenigen des Vorjahres.

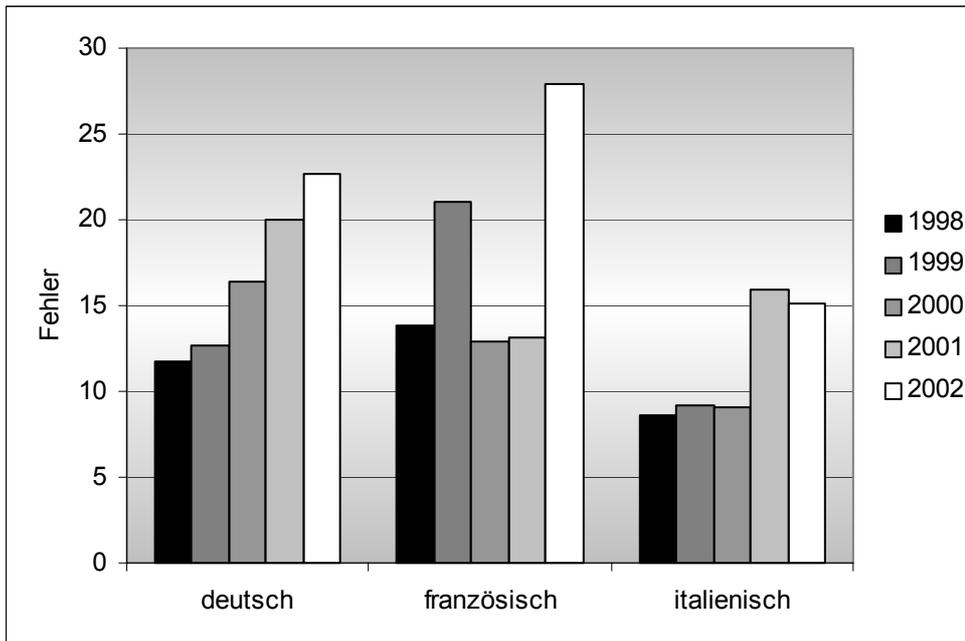


Abbildung 57: Anzahl falsch bearbeiteter Zeichen im Untertest "Konzentriertes und sorgfältiges Arbeiten", nach Jahren und Sprachgruppen

Abbildung 58 gibt den prozentualen Anteil falsch bearbeiteter Zeichen an der durchschnittlichen Anzahl korrekt bearbeiteter Zeichen für die Sprachgruppen wieder.

Die stetige Steigerung in der deutschsprachigen Gruppe geht mit einem ebenfalls stetig steigenden prozentualen Fehleranteil einher. Entgegen der Ergebnisse der Korrelationen zwischen der Anzahl richtig bearbeiteter Zeichen und der Fehlerzahl, scheinen längerfristig Anzeichen für eine Tendenz zu "schneller, aber unpräziser" zu bestehen.

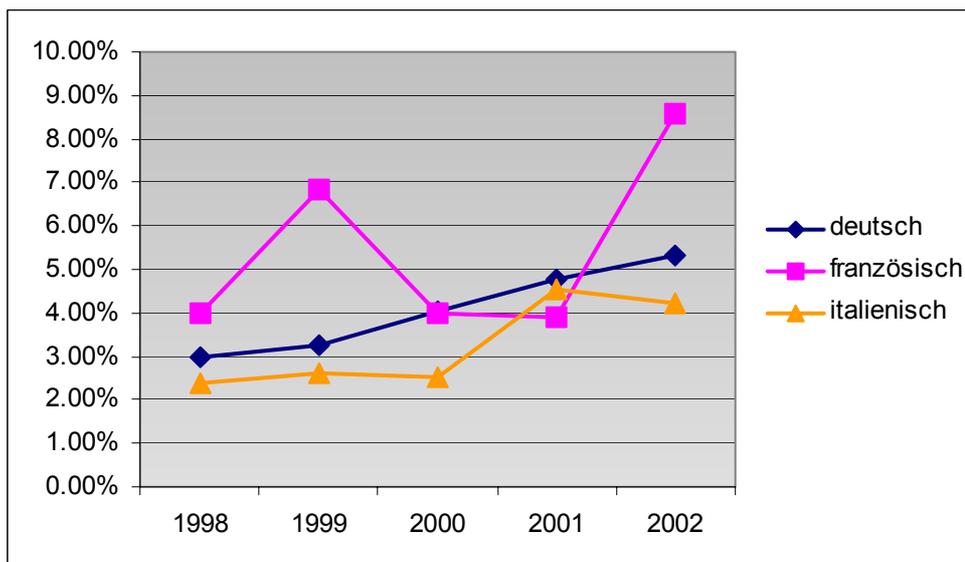


Abbildung 58: Prozentualer Anteil falsch bearbeiteter Zeichen

Wie aus Abbildung 59 ersichtlich, führt diese überproportionale Erhöhung der Fehlerzahlen dazu, dass die durchschnittlich erreichten Punktzahlen der deutschsprachigen Kandidaten trotz wachsender Anzahl korrekt bearbeiteter Zeichen nur schwach ansteigen, und zwar von 12.7 Punkten im Jahre 1998 auf 13.4 Punkte für das Jahr 2002.

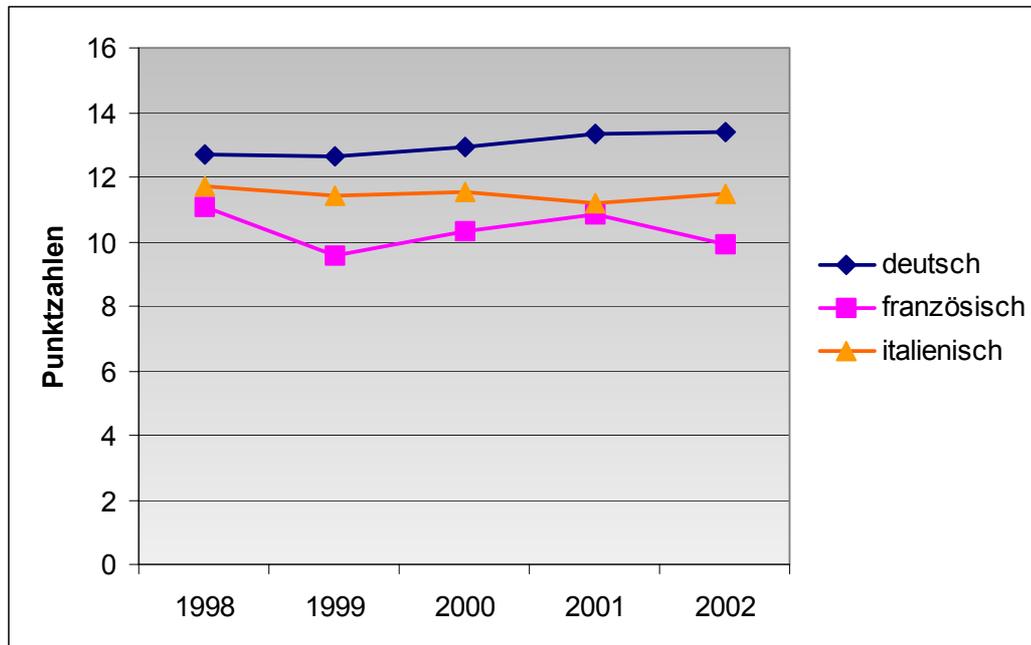


Abbildung 59: Entwicklung der Punktzahlen für "Konzentriertes und sorgfältiges Arbeiten"

9 Zusammenfassung Kantone 1998 bis 2002

Seit 1998 haben 4259 Personen am Test teilgenommen. In der Tabelle 50 sind alle Teilnehmer am EMS 1998 bis 2002 nach Kantonen aufgeführt. Zu beachten ist, dass 1998 nur Humanmedizin, 1999 bis 2001 Human- und Veterinärmedizin einem NC unterlagen.

Wohnkanton/ Wohnort	Disziplin		Geschlecht		Sprache			Total
	Human- medizin	Veterinär- medizin	männlich	weiblich	deutsch	franzö- sisch	italie- nisch	
AG	284	46	123	207	330			330
AI	8	4	4	8	12			12
AR	28	2	11	19	30			30
BE	563	121	247	437	670	14		684
BL	218	21	100	139	237	2		239
BS	169	15	61	123	183		1	184
FR	157	31	76	112	65	123		188
GE	3	39	12	30	1	41		42
GL	17	3	4	16	20			20
GR	105	20	50	75	119		6	125
JU	1	5	3	3		6		6
LU	210	40	95	155	250			250
NE	4	19	1	22	1	22		23
NW	18	2	4	16	20			20
OW	24		7	17	24			24
SG	257	30	112	175	287			287
SH	41	6	16	31	47			47
SO	117	25	62	80	141	1		142
SZ	79	8	32	55	87			87
TG	116	21	46	91	137			137
TI	123	28	73	78	2		149	151
UR	14	2	8	8	16			16
VD	8	60	27	41	3	65		68
VS	70	24	34	60	63	31		94
ZG	51	16	26	41	66	1		67
ZH	798	122	329	591	914	1	5	920
FL	17	3	9	11	20			20
Ausland	39	7	18	28	45	1		46
Total	3539	720	1590	2669	3790	308	161	4259

Tabelle 50: Gesamtzahl der Teilnehmer am EMS 1998 bis 2002. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die Testwerte sind durch die Standardisierung auf den jeweiligen Mittelwert und die Streuung des Geburtsjahres zwischen den Jahren direkt vergleichbar. Deshalb kann ein Vergleich über alle Jahre gezogen werden. 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	n	Mittelwert	Std. Abweichung	Minimum	Maximum
AG	330	101.2	8.9	74	127
AI	12	103.2	6.8	95	116
AR	30	100.8	9.5	88	127
BE	684	100.0	9.9	72	125
BL	239	100.2	9.6	74	125
BS	184	96.4	10.2	71	122
FR	188	94.5	8.9	70	121
GE	42	93.8	8.3	76	117
GL	20	103.2	10.1	89	130
GR	125	99.6	9.7	76	128
JU	6	90.3	4.1	86	96
LU	250	99.9	9.5	72	126
NE	23	92.2	9.9	76	116
NW	20	100.3	12.3	78	121
OW	24	100.4	7.1	92	126
SG	287	102.0	10.0	70	128
SH	47	100.2	9.3	82	120
SO	142	101.5	9.8	70	127
SZ	87	99.3	9.0	76	127
TG	137	101.1	9.5	77	122
TI	151	96.5	9.1	75	120
UR	16	101.8	8.5	87	120
VD	68	95.3	8.9	78	119
VS	94	97.8	9.9	70	120
ZG	67	100.7	9.4	78	120
ZH	920	102.2	10.4	70	130
FL	20	98.2	10.3	82	120
Ausland	46	96.2	9.4	78	113
Total	4259	100.0	10.0	70	130

Tabelle 51: Kennwerte für den Testwert von Kantonen. 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. Die mittleren Testwerte für beide Disziplinen unterscheiden sich, die Zusammensetzung der Bewerber ist nicht vergleichbar (grösserer Bewerberanteil für Veterinärmedizin aus der Westschweiz) und auch die Zulassungsquoten sind unterschiedlich. Bei der Interpretation sind die Stichprobengrössen der Tabelle 50 unbedingt zu beachten.

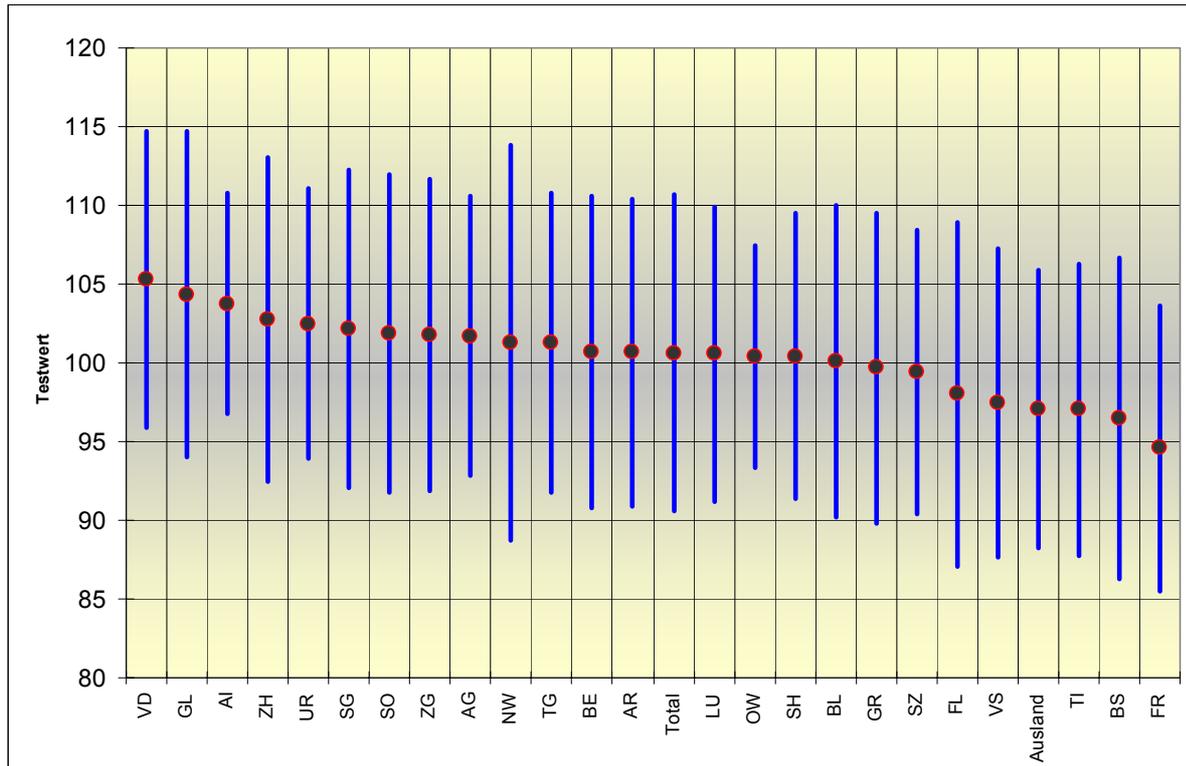


Abbildung 60: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone mit mehr als 5 Bewerbern 1998 bis 2002 zusammengefasst; **Humanmedizin**

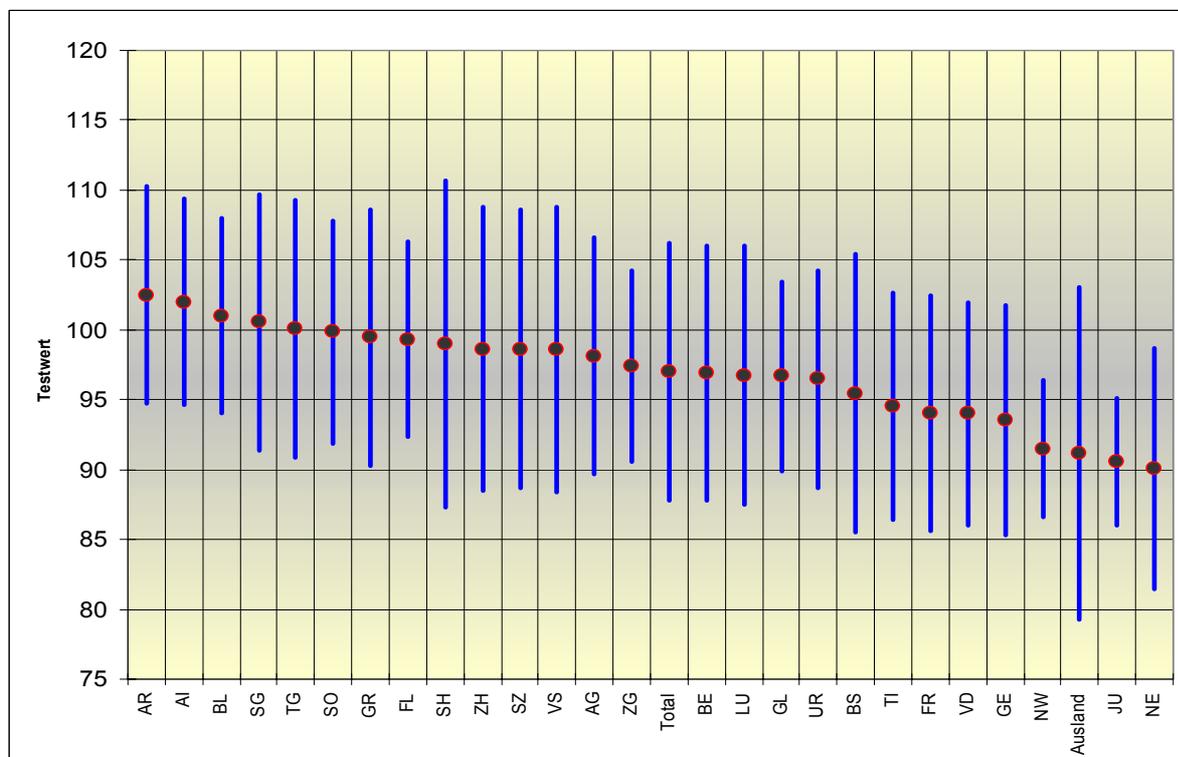


Abbildung 61: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1999 bis 2002 zusammengefasst; **Veterinärmedizin**

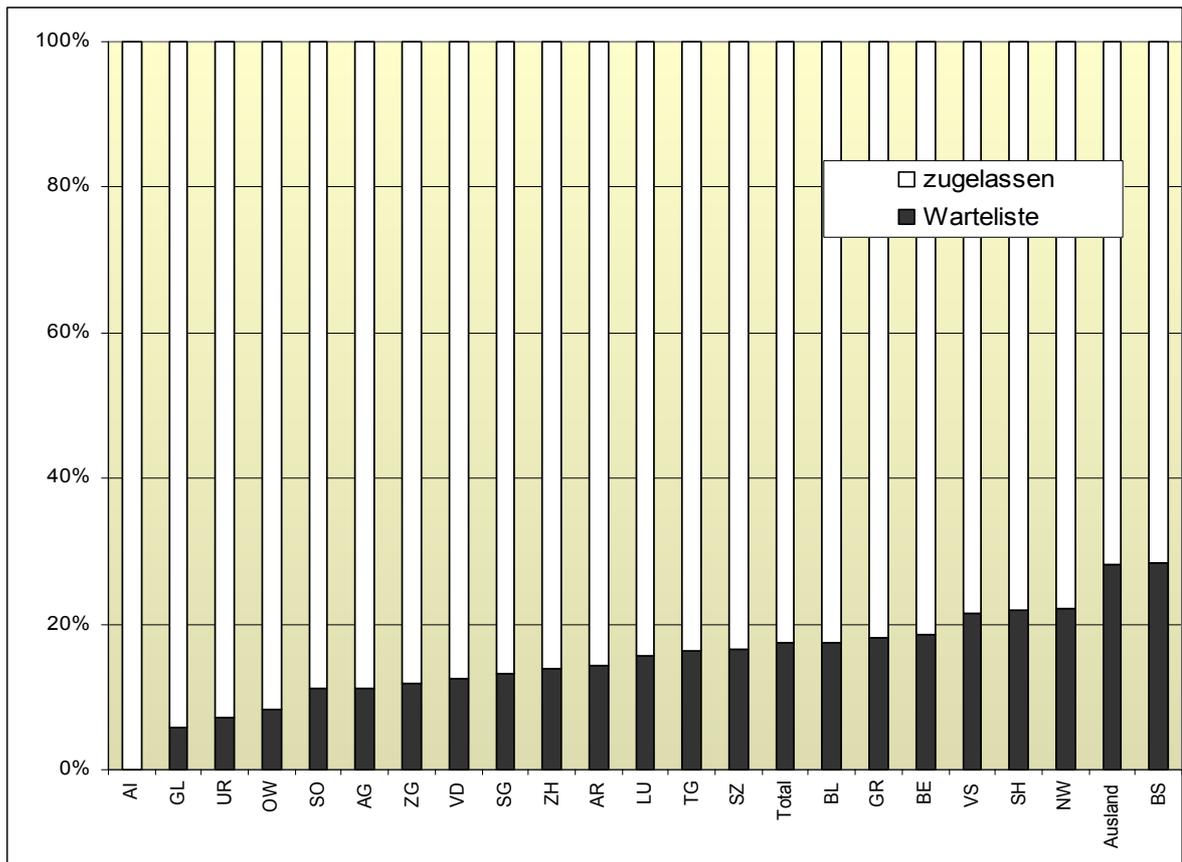


Abbildung 62: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1998 bis 2002; **Humanmedizin**

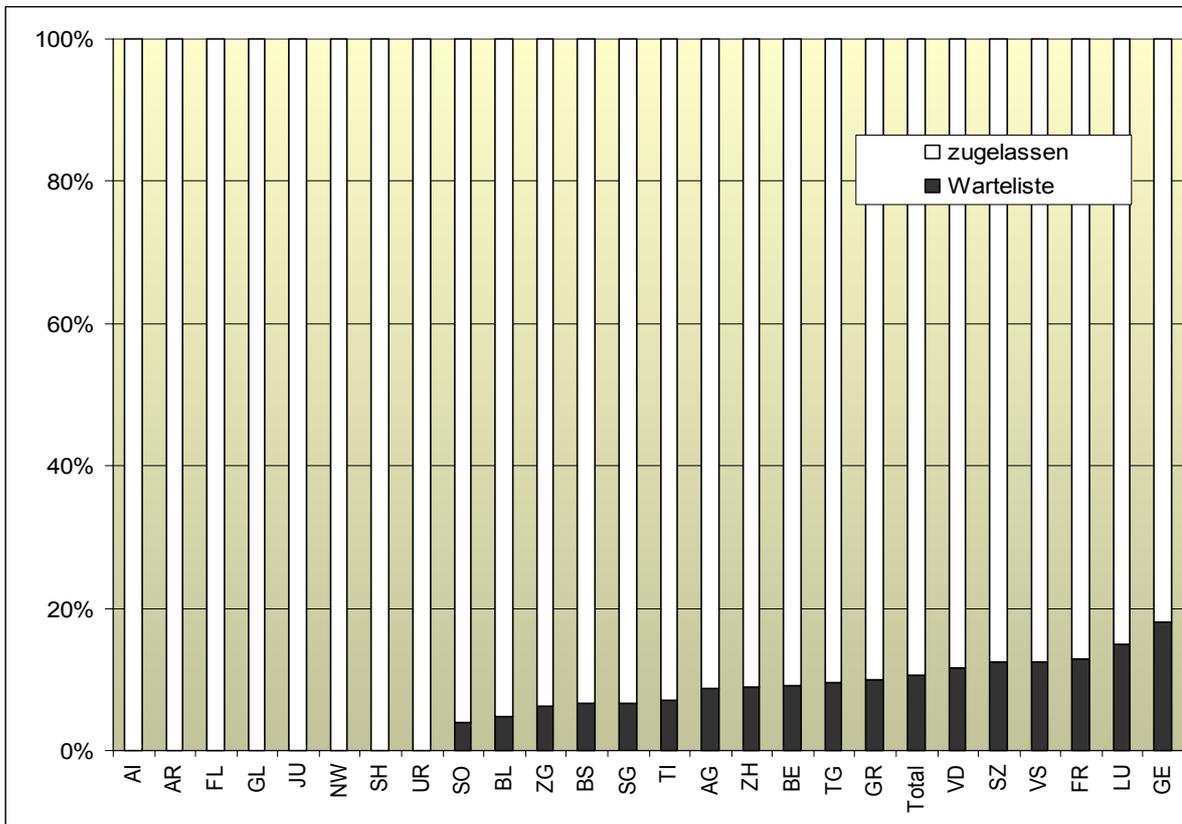


Abbildung 63: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1999 bis 2002; **Veterinärmedizin**

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
- 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.
- 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.
- 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, Université de Fribourg (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/welcome.htm>

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

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

11 Anhang zum Bericht 2002

11.1 Beispielaufgaben für die Untertests

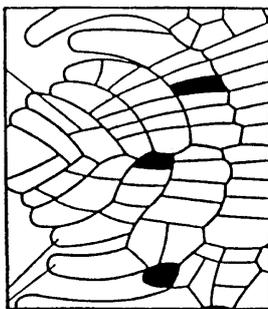
Nachfolgend wird pro Untertest eine Beispielaufgabe dargestellt. So können lediglich die Prinzipien der Aufgabenstruktur verdeutlicht werden – die Aufgaben unterscheiden sich innerhalb jedes Untertests bezüglich des Schwierigkeitsgrades und der Anforderung.

11.1.1 Untertest: Muster zuordnen

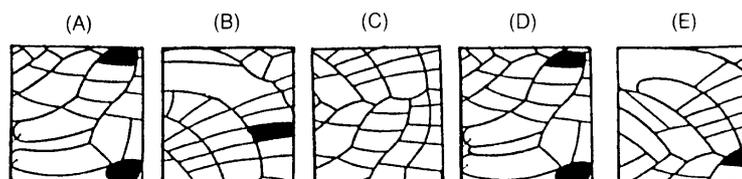
In diesem Untertest wird die Fähigkeit geprüft, Ausschnitte in einem komplexen Bild wiederzuerkennen. Dazu werden pro Aufgabe ein Muster und je fünf Musterausschnitte (A) bis (E) vorgegeben. Die Testteilnehmerin oder der Testteilnehmer soll herausfinden, welcher dieser fünf Musterausschnitte an irgendeiner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

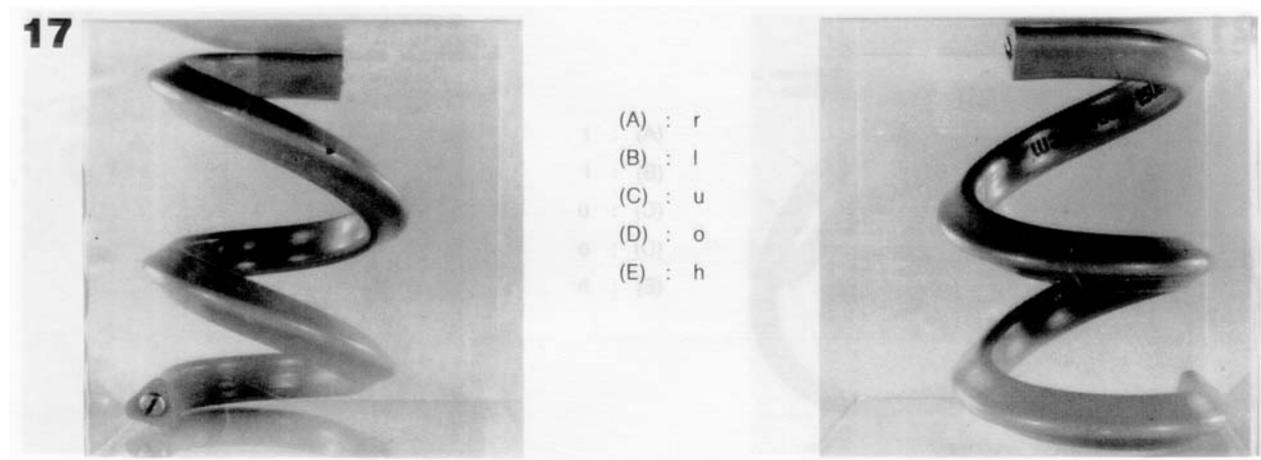
- I. Da bei der Mehrzahl der Linkshänder die Sprachregion in der linken Hemisphäre liegt, müssen sie also mit einer Sprachstörung rechnen, weshalb Aussage I falsch ist.
- II. Da es im Kindesalter noch offen ist, in welcher Hälfte des Gehirns die Sprachregion angelegt wird, besteht für ein Vorschulkind 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 Querstrichen versehen sind, die entweder beide unten, beide oben oder je einer unten und oben angebracht sind, markiert werden. Die Lösungsmenge ist ebenso wichtig wie die Fehlerfreiheit der Bearbeitung. Dieser Test ist trainierbar – im Test-Info wird darauf hingewiesen, diesen Untertest vor der Testabnahme mehrfach zu üben.

b b b

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

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

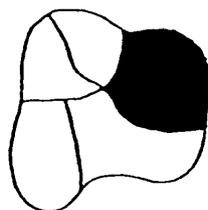
11.1.6 Untertest: Figuren lernen

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.

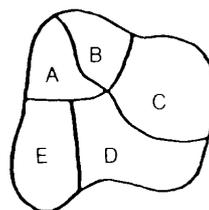
Der Untertest "Figuren lernen" prüft, wie gut man sich Einzelheiten von Gegenständen einprägen und merken kann.

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 Jodatomb eingebaut. So entstehen Monojodtyrosin-Reste (MIT), von denen ein Teil durch die Bindung je eines weiteren Jodatoms in Dijodtyrosin-Reste (DIT) umgewandelt wird. Durch die Verknüpfung von je zwei DIT-Resten entsteht schliesslich T_4 , während aus der Verbindung je eines MIT-Restes mit einem DIT-Rest T_3 hervorgeht. T_3 und T_4 werden dann in den Follikeln gespeichert und bei Bedarf über die Epithelzellen ins Blut freigesetzt.

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

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

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

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

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

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

11.1.9 Untertest: Diagramme und Tabellen

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

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

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

Welche Aussage lässt sich aus den gegebenen Informationen nicht ableiten?

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

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

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

Wenn der Test die Studieneignung erfassen will, ist seine Struktur natürlich abhängig von der Entwicklung dieser Studienanforderungen. In der Schweiz werden solche Veränderungen gegenwärtig in Reformkonzepten diskutiert und bereits umgesetzt. Sozialkommunikative Eignung, problemorientiertes Lernen oder mehr Verschränkung von Grundlagen und Anwendung werden hier genannt. Es werden auch zukünftig nicht alle der heute erfassten Anforderungen weniger wichtig für Studienerfolg werden – ohne intellektuelle Eignung ist die erfolgreiche Absolvierung eines Studiums kaum denkbar. Dennoch ist zu prüfen, ob sich andere Aspekte der Eignung gleichfalls erfassen lassen, um den Test ggf. zu aktualisieren.

Bei allen Aktualisierungen werden zwei Zielstellungen zu beachten sein:

- Der Aufwand für die Erfassung muss tragbar sein (z.B. wäre es unrealistisch, mit allen Bewerbern Eignungsgespräche durchzuführen).
- Die Trainierbarkeit der Eignungsabklärung muss gering bleiben – da andernfalls der Erfolg vom Trainingsaufwand und damit auch von finanziellen Möglichkeiten abhängig wird. Leider beruhen alle ernstzunehmenden Methoden der Eignungsfeststellung für den sozial-kommunikativen Bereich auf Anforderungen, die zumeist recht gut trainiert werden können.

11.2 Weitere Detailergebnisse

Gesamt	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen	1		1	2	7	7	28	38	83	100	130	133	121	121	82	55	34	14	11	3	2
Med.-nat. Grundv.	1	1	2	13	22	36	74	76	116	123	116	115	94	74	40	33	18	8	7	3	1
Schlauchfiguren			3	5	7	10	28	42	63	81	104	97	115	93	82	71	68	40	35	20	9
Qua.form Probl.			4	8	21	31	48	59	82	84	109	111	82	80	70	49	46	39	29	16	5
Textverständnis			11	17	44	76	81	110	114	98	102	83	68	49	46	40	14	13	7		
Figuren lernen			1	1	5	6	13	12	14	31	32	52	59	74	74	62	91	98	79	89	180
Fakten lernen		1	2	12	25	32	68	70	92	103	114	102	66	73	53	52	40	23	24	9	12
Diagr. und Tabellen			1	6	15	25	42	56	98	111	105	140	98	89	71	48	28	25	11	4	
Kon.sorgf. Arbeiten	2				1	2	13	25	42	67	80	90	97	107	104	83	65	58	49	64	24

Tabelle A 1: Häufigkeitsverteilung der Punkte Gesamt 2002

Deutsch	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen	1			1	5	5	21	35	76	86	115	124	112	116	78	51	31	14	10	3	2
Med.-nat. Grundv.	1	1	2	8	17	27	65	63	106	115	108	107	89	70	38	32	18	8	7	3	1
Schlauchfiguren			2	4	6	9	25	36	56	74	91	87	99	89	77	69	61	39	33	20	9
Qua.form Probl.			3	5	19	27	42	53	73	76	102	101	73	75	66	47	42	36	26	16	4
Textverständnis			8	14	41	70	71	98	97	90	93	80	64	44	46	39	12	12	7		
Figuren lernen					5	5	11	10	11	21	27	47	48	64	68	59	76	92	76	86	180
Fakten lernen		1	1	9	22	29	59	60	77	91	107	91	64	71	49	49	39	23	23	9	12
Diagr. und Tabellen			1	5	13	22	35	51	82	95	100	131	90	83	68	46	26	23	11	4	
Kon.sorgf. Arbeiten	1				1	1	9	19	35	55	71	76	86	98	100	80	62	56	49	63	24

Tabelle A 2: Häufigkeitsverteilung der Punkte deutsche Sprachgruppe 2002

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

Tabelle A 3: Häufigkeitsverteilung der Punkte französische Sprachgruppe 2002

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

Tabelle A 4: Häufigkeitsverteilung der Punkte italienische Sprachgruppe 2002

Human- medizin	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Muster zuordnen				2	5	4	21	30	62	73	94	110	100	100	69	49	26	11	10	3	1
Med.-nat. Grundv.	1	1	1	7	14	22	59	54	92	99	93	93	82	61	32	28	14	7	6	3	1
Schlauchfiguren			3	5	5	5	17	31	47	61	78	69	90	83	61	61	60	35	32	19	8
Qua.form Probl.			3	3	19	24	38	41	56	62	82	91	68	67	57	42	42	33	24	14	4
Textverständnis			6	10	35	60	65	83	85	82	76	68	58	42	40	33	10	12	5		
Figuren lernen			1	1	5	5	10	8	6	21	28	40	43	56	49	52	71	76	68	73	15 7
Fakten lernen		1		7	19	18	53	58	72	72	92	77	54	61	44	46	38	18	24	6	10
Diagr. und Tabellen			1	4	11	20	29	42	68	84	89	108	80	76	61	43	22	17	11	4	
Kon.sorgf. Arbeiten	2				1	1	8	16	28	48	59	66	74	87	84	69	57	50	43	54	23

Tabelle A 5: Häufigkeitsverteilung der Punkte Humanmedizin 2002

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		1		2	3	7	8	21	27	36	23	21	21	13	6	8	3	1		1
Med.-nat. Grundv.			1	6	8	14	15	22	24	24	23	22	12	13	8	5	4	1	1		
Schlauchfiguren					2	5	11	11	16	20	26	28	25	10	21	10	8	5	3	1	1
Qua.form Probl.			1	5	2	7	10	18	26	22	27	20	14	13	13	7	4	6	5	2	1
Textverständnis			5	7	9	16	16	27	29	16	26	15	10	7	6	7	4	1	2		
Figuren lernen						1	3	4	8	10	4	12	16	18	25	10	20	22	11	16	23
Fakten lernen			2	5	6	14	15	12	20	31	22	25	12	12	9	6	2	5		3	2
Diagr. und Tabellen				2	4	5	13	14	30	27	16	32	18	13	10	5	6	8			
Kon.sorgf. Arbeiten						1	5	9	14	19	21	24	23	20	20	14	8	8	6	10	1

Tabelle A 6: Häufigkeitsverteilung der Punkte Veterinärmedizin 2002

	Humanmedizin		Veterinärmedizin			Humanmedizin		Veterinärmedizin	
	Punktwert	Testwert	Punktwert	Testwert		Punktwert	Testwert	Punktwert	Testwert
36	1				101	17	34	5	8
48	1				102	15	28		9
50	1				103	15	10	7	4
52	1				104	16	27	6	1
53	1				105	15	39	6	5
55	2				106	19	23	2	5
56	1				107	17	27	4	3
57	1		2		108	11	34	5	3
58	1				109	10	20	4	6
59	1				110	15	27	1	3
60	1		2		111	12	23		3
61	1				112	21	21	4	3
62	2		3		113	18	16	1	1
63	1		1		114	12	10	2	3
64	1		1		115	11	10	3	
65	1		1		116	16	13	1	
66	3		1		117	11	6	2	
67	2		1		118	20	6	1	2
68	2		2		119	14	4	2	1
69	2		3		120	8	2	4	
70	3	1			121	12	5	2	
71	1				122	16	2	2	
72	4	1			123	11	3	1	
73	2	1			124	13	2	3	
74	4	2	3		125	10	3		1
75	5	2	1		126	8	3	2	1
76	5	2	1	2	127	13		1	
77	9	2	3		128	9	1		
78	5	2	4	2	129	7		1	
79	5	3	5	4	130	5		2	
80	9	2	7	2	131	5		1	
81	5	5	2	2	132	6			
82	12	4	3	5	133	4			
83	10	4	4		134	6			
84	9	6	3		135	7			
85	9	9	4	4	136	3			
86	12	14	2	4	137	3			
87	12	5	2	4	138	2		1	
88	10	14	2	12	139	4		1	
89	10	17	8	5	140	4		1	
90	12	19	4	7	141	1			
91	10	21	2	6	142	1			
92	13	22	3	4	143	3			
93	17	22	8	12	144	2			
94	8	23	5	5	145	1			
95	8	25	7	13	146	1			
96	15	23	4	11	148	3			
97	12	26	8	12	150	2			
98	14	31	4	7	151	2			
99	19	32	3	5	152	1		1	
100	12	31	4	13	153	2		1	
					154	1			
					157	1			

Tabelle A 7: Punkt- und Testwertverteilung nach Disziplinen 2002

	Punkt- wert	Test- wert	Mus- ter zu- ord- nen	Med.- naturw. Grundv.	Schlauch- figuren	Quantit. formale Proble- me	Text- verstän- dis	Figur- lernen	Fak- ten ler- nen	Dia- gramme und Tabellen	Konz. und sorgf. A.
Total											
Mittelwert	103.19	100.02	11.21	9.78	11.98	11.06	9.13	15.39	10.55	10.71	13.21
N	973	973	973	973	973	973	973	973	973	973	973
Standardabw.	19.383	9.985	2.858	3.126	3.501	3.730	3.400	3.882	3.740	3.151	3.513
Kurtosis	-.091	-.148	.143	-.030	-.376	-.507	-.456	-.179	-.361	-.279	-.440
Skewness	-.051	-.037	-.010	.162	-.012	.113	.300	-.689	.297	.075	-.042
Median	103.00	100.00	11.00	10.00	12.00	11.00	9.00	16.00	10.00	11.00	13.00
Deutsch											
Mittelwert	104.39	100.63	11.33	9.95	12.10	11.14	9.22	15.66	10.72	10.81	13.46
N	886	886	886	886	886	886	886	886	886	886	886
Standardabw.	19.097	9.853	2.805	3.093	3.504	3.697	3.411	3.793	3.752	3.145	3.457
Kurtosis	-.134	-.141	.118	.036	-.419	-.518	-.509	-.115	-.399	-.257	-.549
Skewness	-.035	-.035	.027	.166	-.017	.112	.279	-.744	.268	.047	-.051
Median	104.00	100.00	11.00	10.00	12.00	11.00	9.00	16.00	10.00	11.00	13.00
Französisch											
Mittelwert	89.58	93.10	9.73	7.80	10.69	10.05	7.98	12.27	8.46	10.02	10.32
N	59	59	59	59	59	59	59	59	59	59	59
Standardabw.	18.486	9.336	2.982	2.981	3.109	4.015	3.149	4.004	3.013	3.093	3.115
Kurtosis	.198	-.273	1.018	-.552	.534	-.446	.294	-.258	.371	-.413	.949
Skewness	-.384	-.226	.034	.410	-.386	.139	.215	-.350	.321	.479	-.325
Median	88.00	92.00	10.00	7.00	11.00	10.00	8.00	12.00	8.00	9.00	10.00
Italienisch											
Mittelwert	94.07	95.32	10.64	8.57	11.11	10.57	8.50	13.39	9.54	8.75	11.46
N	28	28	28	28	28	28	28	28	28	28	28
Standardabw.	17.295	8.957	3.391	2.860	3.635	3.948	3.180	2.936	3.350	2.703	2.848
Kurtosis	-.875	-.805	-.930	-.495	-.551	-.190	.674	-.750	.083	-.169	1.031
Skewness	.252	.270	-.246	-.116	.266	.442	1.128	-.446	.513	-.223	.662
Median	93.00	94.50	10.00	8.00	10.50	11.00	7.50	13.50	9.50	8.50	11.00
Humanmedizin											
Mittelwert	104.95	100.92	11.36	9.97	12.25	11.27	9.26	15.60	10.79	10.86	13.50
N	770	770	770	770	770	770	770	770	770	770	770
Standardabw.	19.338	9.961	2.812	3.080	3.533	3.742	3.389	3.883	3.743	3.142	3.499
Kurtosis	-.010	-.092	-.062	.078	-.341	-.549	-.534	.035	-.474	-.238	-.312
Skewness	-.148	-.129	.009	.172	-.087	.046	.281	-.791	.266	.026	-.135
Median	105.00	101.00	11.00	10.00	12.00	11.00	9.00	16.00	10.00	11.00	13.00
Veterinärmedizin											
Mittelwert	96.55	96.61	10.65	9.09	11.00	10.25	8.62	14.58	9.61	10.10	12.13
N	203	203	203	203	203	203	203	203	203	203	203
Standardabw.	18.114	9.340	2.964	3.211	3.197	3.579	3.405	3.780	3.589	3.115	3.358
Kurtosis	.205	.217	.778	-.349	-.292	-.077	-.096	-.662	.246	-.262	-.519
Skewness	.270	.268	-.011	.205	.188	.363	.392	-.364	.421	.265	.297
Median	96.00	96.00	10.00	9.00	11.00	10.00	8.00	15.00	9.00	10.00	12.00

Tabelle A 8: Verteilungskennwerte für Punktwert, Testwert und Untertests nach Sprachgruppen und Disziplinen 2002

Das Zentrum für Testentwicklung und Diagnostik im Internet:

www.unifr.ch/ztd

The screenshot shows a web browser window displaying the website <http://www.unifr.ch/ztd/ems/welcome.htm>. The browser's address bar shows the URL, and the page title is "Eignungstest für das Medizinstudium in der Schweiz". The website content includes:

- Navigation Menu (Left):**
 - Kandidatinnen & Kandidaten**
 - Termine & Ablauf
 - Wie vorbereiten?
 - Fragen & Antworten
 - CRUS (mehr Info)
 - Forschung zum EMS**
 - deutsch
 - français
 - Weitere Informationen der CRUS**
- CRUS (Center for Research in University Studies):**
 - Rektorenkonferenz der Schweizer Universitäten
 - Conférence des Recteurs des Universités Suisses
 - Conferenza dei Rettori delle Università Svizzere
- Contact:** E-MAIL : ZTD@unifr.ch
- Main Content Area:**
 - Information des ZTD für Teilnehmende am EMS**
 - Eignungstest 2002 für das Medizinstudium**
 - Am 5.7. fand der Eignungstest in 7 Testorten und 3 Sprachen mit Erfolg statt.**
 - Text: Gegenwärtig werden die Ergebnisse ausgewertet - Hier finden Sie einen ersten statistischen Bericht auf deutsch [Statistischer Bericht 2002](#). Im Spätherbst werden die Ergebnisse im Bericht 8 des ZTD publiziert.
 - Text: Actuellement les questionnaires sont en cours de dépouillement - Ici vous trouvez le rapport en français [Rapport statistique 2002](#). Un rapport complet du CTD qui comprendra également les détails des résultats du test sera disponible en automne 2002.
 - Weitere Informationen des ZTD**
 - FAQ zum EMS:** Antworten auf häufig gestellte Fragen zum EMS. Hier finden Sie Antworten auf Fragen, die uns bezüglich Numerus Clausus und Test häufig gestellt werden.
 - ORIGINAL-VERSIONEN:** Wie bereitet man sich auf dem EMS vor? Comment se préparer pour le test AMS? Come prepararsi per il test AMS?
 - TRAINING des EMS NOTWENDIG...** ... oder Geschäft mit der Angst??? - Geringe Trainierbarkeit ist ein Vorzug des EMS.
 - FÜR INTERESSIERTE AN FORSCHUNGSERGEBNISSEN**
 - Text: Der Bericht Nummer 7 über den Eignungstest 2001 ist fertiggestellt. Er enthält zugleich die Ergebnisse zur Vorhersage des Prüfungserfolges bei der ersten und zweiten ärztlichen Vorprüfung für bereits zwei Test-Jahrgänge. Die Ergebnisse bestätigen, dass der Test tatsächlich die Eignung erfasst.