

EMS • Eignungstest für das Medizinstudium in der Schweiz • 2005

*Bericht 11 über die Durchführung und Ergebnisse 2005
Trends und Schlussfolgerungen aus acht Jahren Testanwendung
Mit Studie "Berufswünsche der Kandidaten für Humanmedizin"*

avec un résumé en français

K.-D. Hänsgen und B. Spicher

Bericht 11 (2005)

Hänsgen, Klaus-Dieter; Spicher, Benjamin (2005):

EMS Eignungstest für das Medizinstudium in der Schweiz 2005;

Berichte des Zentrums für Testentwicklung, Nummer 11, im Auftrag der Schweizerischen Universitätskonferenz (SUK);

Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik, 2005.

Mit einem Beitrag „Berufswünsche der Kandidaten für Humanmedizin“
von B. Weickhardt, B. Spicher und K.-D. Hänsgen.

Unter Mitarbeit von Béatrice Weickhardt (Lektorat),

Patrick Bergem, Mauro Bernasconi, Sébastien Simonet und Maurizio Strazzeri.

Redaktion: Tanja Läng

Die Mitglieder des Beirates „Eignungstest“ der Schweizerischen Universitätskonferenz zum Zeitpunkt der Testabnahme 2005:

Dr. N. Ischi (Vorsitzender, Bern); Prof. Dr. P. Eggli (Bern); Prof. Dr. P. Groscurth (Zürich);

Dr. R. Hofer (Bern); Prof. Dr. M. Kleinmann (Zürich); Prof. Dr. J.-P. Montani (Freiburg);

Prof. Dr. K. Opwis (Basel); Prof. Dr. M. Perrez (Freiburg); Prof. Dr. J.-L. Reymond (Bern);

Prof. Dr. U. M. Spornitz (Basel); Prof. Dr. G. Trost (Bonn/Deutschland); K. Wechsler (Bern).

© ZTD Zentrum für Testentwicklung und Diagnostik 2005

am Departement für Psychologie der Universität Freiburg

Direktor: Prof. Dr. K.-D. Hänsgen

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

E-Mail: ztd@unifr.ch

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

Druck: Cric Print, Freiburg/Schweiz

Inhaltsverzeichnis

1	ZUSAMMENFASSUNG.....	5
1.1	Statistik	5
1.2	Organisation und Ablauf	5
1.3	Testgüte	6
1.4	Weiterentwicklung des EMS	6
1.5	Gleichbehandlung der Sprachgruppen	6
1.6	Differenzierung nach Alter	7
1.7	Differenzierung nach Geschlecht	7
1.8	Differenzierung nach Maturitätsquote und Kanton	8
1.9	Schlussfolgerungen.....	8
2	RESUME.....	9
2.1	Statistique	9
2.2	Organisation et déroulement.....	9
2.3	Qualité du test.....	10
2.4	Développement ultérieur de l'AMS	10
2.5	Egalité de traitement entre groupes linguistiques.....	10
2.6	Différences entre classes d'âge	11
2.7	Différences entre femmes et hommes	12
2.8	Différences entre cantons et taux de maturité	12
2.9	Conclusions	12
3	NEUES ZU EIGNUNGSTEST UND NUMERUS CLAUSUS	13
4	NUMERUS CLAUSUS (NC) UND MEDIZINSTUDIUM	16
5	ANMELDUNG ZU MEDIZINSTUDIUM UND TEST 2005.....	18
5.1	Anmeldungen, Testantritte und Zulassungen.....	18
5.2	Grösse der Testlokale	23
5.3	Testorte und Wunschuniversitäten	24
5.4	Wunschuniversitäten und Testorte nach Wohnkantonen	25
5.5	Testabsolvierung nach Alter und Geschlecht	29
5.6	Übernahme des Testergebnisses aus Vorjahren	32
5.7	Erneute Testteilnahme 2005	33
5.8	Sprachgruppen.....	35
5.9	Alter und Maturitätsjahr	37
6	BESCHREIBUNG DES VERWENDETEN EIGNUNGSTESTS	38
6.1	Aufbau des Tests	38
6.2	Berechnung der Werte	40
6.3	Mittlerer Rangplatz der Untertests	41

7	TESTANWENDUNG IN DER SCHWEIZ 2005	43
7.1	Verteilungsprüfung.....	43
7.2	Vergleich der Testfassungen 1998 bis 2005	46
7.3	Äquivalenz der Sprachversionen.....	48
7.3.1	<i>Sprachvergleich für die Untertests.....</i>	<i>50</i>
7.3.2	<i>Darstellung des Korrekturverfahrens</i>	<i>52</i>
7.4	Vergleichbarkeit der Testlokale	66
7.5	Vergleich der Geschlechter	67
7.6	Vergleich der Kantone	73
7.6.1	<i>Vergleich 2005 innerhalb der Deutschschweiz.....</i>	<i>73</i>
7.7	Vergleiche für Altersgruppen.....	78
7.8	Vergleiche nach Wunschuniversitäten.....	81
8	ERGEBNISSE ZUR TESTGÜTE.....	83
8.1	Zuverlässigkeit	83
8.2	Binnenstruktur.....	86
8.3	Ein neuer Untertest: Planen und Organisieren	92
8.4	Konzentriertes und sorgfältiges Arbeiten.....	95
8.5	Item-Trennschärfen	98
9	ZUSAMMENFASSUNG KANTONE 1998 BIS 2005.....	103
10	BEISPIELAUFGABEN FÜR DIE UNTERTESTS.....	110
10.1	Untertest: Muster zuordnen.....	110
10.2	Untertest: Medizinisch-naturwissenschaftliches Grundverständnis	110
10.3	Untertest: Schlauchfiguren.....	111
10.4	Untertest: Quantitative und formale Probleme.....	112
10.5	Untertest: Textverständnis	113
10.6	Untertest: Konzentriertes und sorgfältiges Arbeiten.....	114
10.7	Untertest: Figuren lernen.....	115
10.8	Untertest: Fakten lernen	115
10.9	Untertest: Planen und Organisieren	115
10.10	Untertest: Diagramme und Tabellen	117
11	LITERATUR.....	119
11.1	Originaltest zur Information und Vorbereitung	120
11.2	Frühere Berichte des ZTD.....	121
	ANHANG: BERUFSWÜNSCHE DER KANDIDATEN FÜR HUMANMEDIZIN.....	123

1 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2005 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Human-, Veterinär- und Zahnmedizin ab Wintersemester 2005/2006. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2004 herangezogen. **Die Ergebnisse sind nicht repräsentativ für Schlussfolgerungen hinsichtlich einer gesamthaften oder regionalen Evaluation von Bildungseinrichtungen und Bildungswegen.**

1.1 Statistik

Im Jahr 2005 ist erneut für **Humanmedizin (seit 1998), Veterinärmedizin (seit 1999) und Zahnmedizin (seit 2004)** ein Numerus Clausus (NC) für Personen notwendig, die sich an den Universitäten Basel, Bern, Freiburg oder Zürich angemeldet haben. Der EMS wurde in der Schweiz zum achten Mal durchgeführt.

Seit 5 Jahren steigen die Bewerberzahlen im Februar für die Universitäten mit NC um durchschnittlich 150 Personen pro Jahr, was immerhin jeweils der Kapazität der Universität Basel entspricht. Diese Entwicklung deckt sich mit internationalen Trends und wird nach Prognosen des Bundesamtes für Statistik weiter anhalten bzw. auf hohem Niveau bestehen bleiben. Selbst wenn man wegen des erhöhten Ärztebedarfs in Krankenhäusern eine Erhöhung der Ausbildungskapazitäten in realistischem Ausmass in Erwägung ziehen würde, erforderte der verbleibende extreme Bewerberüberhang weiterhin einen NC.¹

Im **Februar 2005** waren aufgrund der vorliegenden Anmeldungen die Kapazitäten der vier betroffenen Universitäten in **Humanmedizin zu 279% (Vorjahr 249%)**, in **Veterinärmedizin zu 212% (Vorjahr 182%)** und in **Zahnmedizin zu 134% (Vorjahr 146%)** ausgelastet. Der Rückgang in der Zahnmedizin kann auch darauf zurückzuführen sein, dass erfahrungsgemäss im zweiten Jahr der Notwendigkeit eines NC dissuasive Effekte wirken. Insgesamt 1598 Personen haben sich bis zur gesetzten Frist im **Mai 2005** bei der CRUS zum EMS-Test angemeldet, 60 Personen übernehmen ihr Testergebnis aus dem Vorjahr.

Erfolgreich absolviert haben den EMS **1501 Personen**. Eine Person hat den Test während des Vormittagsteils abgebrochen, 96 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen. Diese Rückzüge liegen im Bereich der Vorjahre.

1.2 Organisation und Ablauf

Der Test fand am 8.7.2005 gleichzeitig an acht Testorten in drei Sprachen (deutsch: Basel, Bern, Chur, St. Gallen, Zürich sowie erstmals Aarau; französisch: Freiburg; italienisch: Bellinzona) statt. Zu beachten ist, dass die Wahl des Testortes unabhängig vom zukünftigen Studienort möglich ist. Dank einer engagierten Mitwirkung aller Beteiligten kann die Testabnahme erneut als erfolgreich eingeschätzt werden. Am neuen Testort Aarau wurde der Test infolge einer guten Vorbereitung ebenso problemlos durchgeführt wie in den anderen Testorten.

¹ hierzu www.unifr.ch/ztd/ems/berichte/b10/FAQ_EMS_NC.pdf

1.3 Testgüte

Das überzeugendste Argument für die Verwendung des EMS als NC-Kriterium ist seine langjährige Stabilität aller wichtigen Gütekriterien auf hohem Niveau. Objektivität, alle Zuverlässigkeitswerte, die Faktorenstruktur, das Differenzierungsvermögen und die Qualität der Verteilungen liegen auch 2005 im Bereich der Vorjahre. Mit $r = 0.92$ liegt die Zuverlässigkeit des Testwertes nach der Testhalbierungsmethode sehr hoch. Auch die Konsistenz des Testprofils erreicht mit $\alpha = 0.82$ den zweithöchsten Wert seit 1998. Die Faktorenstruktur bleibt ebenfalls stabil. Es kann somit davon ausgegangen werden, dass auch die aktuelle Version des EMS eine hohe Prognosegüte für Studienerfolg aufweist.

Wenn man die neue DIN-Norm 33430 „Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen“ zugrunde legt, sind alle wichtigen Anforderungen hinsichtlich der Qualität des Tests, der Schulung und Qualifikation des Personals sowie der Definiertheit und Transparenz der Abläufe erfüllt.

1.4 Weiterentwicklung des EMS

Der erstmals verwendete Untertest „**Planen und Organisieren**“ hat die in ihn gestellten Erwartungen voll erfüllt. Die Punktwerteverteilung differenziert zu den anderen Tests vergleichbar, die Zuverlässigkeitswerte des Testwertes werden durch diesen Untertest erhöht und er weist hohe Korrelationen zu den stärker für Studienerfolg prognoserelevanten Untertests auf, verstärkt deren Anteil. Männer und Frauen erreichen hier gleiche Ergebnisse, auch die Fairness zwischen den Sprachen ist gegeben. Aufgrund der grösseren Heterogenität der Aufgaben ist die Konsistenz erwartungsgemäss etwas geringer. Es wurden nur 19 Aufgaben gewertet, eine Aufgabe erfüllte die Gütekriterien nicht. Dies werten wir auch als Hinweis dafür, wie wichtig die Vorerprobung unter Ernstfallbedingungen im Rahmen von Einstreuaufgaben ist (was bei diesem Untertest nur in einem modifizierten Verfahren möglich war).

Beim Untertest „**Konzentriertes und sorgfältiges Arbeiten**“ wurde erneut die Art der zu erfüllenden Anforderung vorher nicht bekannt gegeben. Wie bereits 2004 konnte dadurch erreicht werden, dass sich exzessives Trainieren bei diesem Test praktisch nicht mehr auswirkt. Dies soll beibehalten werden.

Das Weglassen der nicht gewerteten Einstreuaufgaben hat trotz der Erweiterung dazu beigetragen, dass die Gesamttestzeit des EMS zu den Vorjahren vergleichbar blieb. Diese Änderung wurde vor allem deshalb positiv aufgenommen, weil keine Unklarheit mehr darüber besteht, welche Aufgaben gewertet werden und welche nicht. Für 2006 wird geprüft, ob durch weitere Strukturveränderungen die Pausengestaltung optimiert werden kann, weil der Vormittagsteil etwas länger geworden ist.

1.5 Gleichbehandlung der Sprachgruppen

Bereits der Übersetzung der Testaufgaben wird grosse Aufmerksamkeit gewidmet, um eine höchstmögliche **Äquivalenz** der drei Sprachfassungen zu erreichen. Besonders erwähnenswert ist, dass die Schlussredaktion von zweisprachigen Lehrern aus der Maturitätsstufe durchgeführt wird, die sich mit der in den Naturwissenschaften verwendeten Fachsprachen an den Schweizer Maturitätsschulen auskennen und speziell dahingehend geschult sind, dass es auf eine äquivalente Übersetzung der Schwierigkeit der deutschsprachigen Originalversion ankommt.

Seit 1999 wird das international eingeführte, so genannte DIF-Verfahren eingesetzt, welches mögliche testbedingte Unterschiede ausgleicht, andere Unterschiede aber in den Daten belässt. Diejenigen Aufgaben, die in einer Sprache deutlich von einem mittleren erwarteten Unterschied abweichen, werden ausgeglichen und führen zu einem Bonuswert. Die Zahl der Aufgaben, die überhaupt ausgeglichen werden mussten, hat über die Jahre allerdings deutlich abgenommen, was für die Optimierung der Übersetzungsprozesse spricht.

Die dennoch vorhandenen Unterschiede zwischen den Sprachgruppen werden hin und wieder fälschlich auf die Übersetzungsqualität oder den deutschen „kulturellen Background“ des Tests geschoben. Dies kann aufgrund der Daten aber eindeutig widerlegt werden:

Ab dem Jahr 2000 stellt man fest, dass die Unterschiede in den **wenig sprachabhängigen** Tests viel **deutlicher** sind als jene in den sprachabhängigen. Wenn man (nur) testbedingte Unterschiede aufgrund der Sprache erwarten würde, sollten sich Unterschiede auch nur (oder zumindest deutlicher) in den sprachabhängigen Tests zeigen, die weniger sprachabhängigen wären weniger unterschiedlich.

Nach der DIF-Korrektur sind die Ergebnisse in den sprachabhängigen Untertests sogar durchgängig und deutlich besser als in den wenig sprachabhängigen Untertests. Insofern kann man eventuell von einer geringfügigen Überkorrektur der Unterschiede ausgehen.

Es gibt in jeder Sprache auch die gleichen Möglichkeiten der Vorbereitung (Test Info, veröffentlichte Originalversion, Hinweise für einen Probelauf unter Echtzeitbedingungen). Die Freiburger Gymnasien haben beispielsweise einen Probelauf auf Französisch angeboten.

Die Leistungsfähigkeit der Sprachgruppen ist also definitiv nicht gleich – unabhängig von allen Spracheffekten. Sprachliche Faktoren bestimmen die Leistungsunterschiede zwischen den Sprachgruppen nachweislich, aber weniger stark als ein allgemeiner Leistungsunterschied. Die Fairness gegenüber den Sprachgruppen ist garantiert.

1.6 Differenzierung nach Alter

Der in den Vorjahren mehrfach bestätigte Befund, dass jüngere Studienbewerber höhere Testwerte erreichen, bestätigt sich nur tendenziell: Die mittlere Altersgruppe und die älteste Altersgruppe mit früher Maturität unterscheiden sich nur jeweils um 1.4 im Testwert von den Jüngeren. Im Vorjahr betragen diese Unterschiede beispielsweise 2.4 bzw. 2.8.

Alle Älteren mit später abgelegter Maturität erhalten einen um 3.3 geringeren Testwert als die jüngere Gruppe – im Vorjahr ergab sich noch ein Unterschied von 6.

Wir erinnern erneut daran, dass auch die Erfolgsquoten in den Vorprüfungen in etwa diesem Trend folgen und der Test damit nur die Prüfungsleistung korrekt vorhersagt.

1.7 Differenzierung nach Geschlecht

Wie der Evaluationsstudie zum EMS (Hänsgen & Spicher, 2002) zu entnehmen ist, ist die Erfolgswahrscheinlichkeit, die erste Vorprüfung zu bestehen, für Männer um 7% höher als für Frauen. Wenn der Test diese Erfolgswahrscheinlichkeit vorhersagen will, muss er auch diesen Unterschied feststellen. Bei dieser Frage ist weiterhin zu beachten, dass Frauen in der Bewerberkohorte deutlich überrepräsentiert sind, es keine vergleichbare Repräsentativität der Bewerbungen beider Gruppen für die Maturanden resp. die Altersgruppe gibt.

Der Testwertunterschied in Humanmedizin wie Zahnmedizin von 2 zugunsten der Männer liegt im Betrag im Bereich der Vorjahre – wurde wie 2004 in Humanmedizin allein aufgrund

der grösseren Bewerberzahlen erneut signifikant. In Veterinärmedizin ist der Testwert der Männer um 4 niedriger als derjenige der Frauen.

Insgesamt gibt es wiederum keine Hinweise, welche gegen eine Chancengleichheit sprechen – die Ergebnisse bestätigen diejenigen der Vorjahre.

1.8 Differenzierung nach Maturitätsquote und Kanton

Es handelt sich bei den Kandidaten um Bewerber für das Medizinstudium, die unterschiedlich repräsentativ für alle Maturanden der jeweiligen Kantone sind. In zweisprachigen Kantonen kommt dazu, dass Französischsprachige und Zweisprachler als Alternative in Genf, Lausanne oder Neuenburg ohne Numerus Clausus ein Studium aufnehmen können. Ebenso können „sozialpsychologische“ Phänomene in Kantonen unterschiedlich sein (auch im Vergleich der Jahrgänge innerhalb eines Kantons), indem sich z.B. einmal nur die besten Maturitätsabsolventen für ein Medizinstudium bewerben oder eben ein solcher „Gruppendruck“ nicht besteht.

Ein genereller Einfluss der Maturitätsquote auf den Testwert lässt sich wie in allen Vorjahren **nicht nachweisen** (3 Gruppen der Kantone nach Maturitätsquote zusammengefasst unterscheiden sich nicht signifikant). Die langjährigen Trends der Kantone mit einer ausreichenden jährlichen Bewerberzahl für statistische Analysen bleiben weiter stabil.

1.9 Schlussfolgerungen

Die achte Durchführung des EMS war wiederum erfolgreich. Die Testergebnisse sind gleich zuverlässig wie in den Vorjahren.

Mit dem EMS steht für den Numerus Clausus ein bewährtes Selektionsinstrument zur Verfügung, welches eine kapazitätsgenaue Zulassung aufgrund der Studieneignung ermöglicht und auch die Umleitungen nach einem objektiven Kriterium regelt.

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

2 Résumé

Le présent rapport expose les résultats obtenus lors du test réglant l'admission en médecine humaine, vétérinaire et dentaire au semestre d'hiver 2005/2006. Il fait également intervenir des données comparatives des éditions antérieures (1998-2004). Mais rappelons que ces résultats ne sont pas suffisamment représentatifs pour permettre de tirer des conclusions quant aux établissements de formation ou aux parcours éducationnels, que ce soit au plan suisse ou régional.

2.1 Statistique

En 2005, un numerus clausus a à nouveau été nécessaire pour les personnes souhaitant commencer des études en **médecine humaine (NC depuis 1998)**, en **médecine vétérinaire (NC depuis 1999)** et en **médecine dentaire (NC depuis 2004)** à l'Université de Bâle, Berne, Fribourg et Zurich. Cette année, le test d'aptitudes pour les études de médecine en Suisse (AMS) a été organisé pour la huitième fois.

Depuis cinq ans maintenant, le nombre (en février) des candidats au test d'aptitudes dans les universités avec NC a augmenté en moyenne de 150 personnes par an, ce qui correspond chaque année à la capacité de l'Université de Bâle. Cette évolution est conforme aux tendances internationales; selon les prévisions de l'Office fédéral de la statistique, elle va persister et le nombre de candidats à l'AMS va demeurer à un haut niveau. Même si l'on devait prendre en considération une augmentation des capacités de formation en fonction de l'accroissement du besoin de médecins dans les hôpitaux, il resterait encore un important surplus de candidats, si bien que le NC serait toujours nécessaire.

En **février 2005**, les pré-inscriptions dépassaient les capacités des quatre universités concernées, que ce soit en **médecine humaine à raison de 279% (année précédente 249%)**, en **médecine vétérinaire à raison de 212% (année précédente 182%)** ou en **médecine dentaire à raison de 134% (année précédente 146%)**. La diminution enregistrée en médecine dentaire peut s'expliquer par le fait que, comme l'expérience le montre, le NC exerce un effet dissuasif l'année suivant son introduction. En **mai 2005** – délai d'inscription – 1598 personnes s'étaient annoncées à la CRUS pour passer le test, 60 personnes reprenant leur résultat de l'année passée.

1501 personnes ont achevé l'AMS, 1 personne l'a interrompu en cours de matinée, 96 se sont désistées ou ne se sont pas présentées; ces valeurs correspondent à celles des années précédentes.

2.2 Organisation et déroulement

Le test, trilingue, s'est déroulé le 8 juillet 2005 au même moment en huit sites différents (français: Fribourg; allemand: Bâle, Berne, Coire, St-Gall, Zurich et pour la première fois Aarau; italien: Bellinzona). A noter qu'il est possible de choisir le lieu du test indépendamment du futur lieu d'études. Grâce à la collaboration active de tous les participants, l'édition 2005 peut elle aussi être considérée comme réussie. Le test a été organisé pour la première fois à Aarau;

il a été si bien préparé qu'il a pu s'y dérouler sans problème, comme dans les autres lieux du test.

2.3 Qualité du test

Depuis des années, tous les critères de qualité importants demeurent à un niveau élevé; cette stabilité constitue l'argument décisif en faveur de l'utilisation des résultats du test comme critère du NC. En 2005, l'objectivité, la fiabilité, la structure factorielle, la capacité de différenciation et la qualité de distribution ont des valeurs semblables à celles des années précédentes. Avec 0.92, la fiabilité selon la méthode "split-half" est très élevée. Quant à la consistance, elle atteint, avec 0.82, la deuxième valeur la plus haute depuis 1998. La structure factorielle reste elle aussi stable, si bien que l'on peut admettre que la version actuelle de l'AMS a également une capacité prédictive élevée pour le succès des études.

Si l'on considère la nouvelle norme DIN 33430 "Exigences en matière de méthode et d'application de tests d'aptitudes professionnelles", toutes les exigences importantes - à savoir celles portant sur la qualité du test, sur la formation et la qualification du personnel ainsi que sur la transparence du déroulement du test - sont remplies en l'occurrence.

2.4 Développement ultérieur de l'AMS

Le sous-test "Planifier et organiser", qui a été appliqué pour la première fois, a répondu à toutes les attentes placées en lui. La différenciation de la répartition des points est comparable à celle des autres tests, la fiabilité du test est augmentée grâce à ce sous-test et en raison des liens importants qui l'unissent aux autres sous-tests à capacité prédictive pour le succès des études, il renforce la valeur de ces derniers. Les hommes et les femmes parviennent aux mêmes résultats et l'équité linguistique est également réalisée. Vu que les items sont plus hétérogènes, la consistance est, comme il fallait s'y attendre, un peu moins importante. Seuls 19 items ont été évalués; l'un d'entre eux ne remplissait pas les critères de qualité. Cela montre également l'importance d'avoir au préalable testé ces items ailleurs, dans des conditions réelles (ce qui n'est possible pour ce sous-test que selon une procédure modifiée).

Cette année encore, la règle pour le marquage dans le sous-test "Travail avec soin et concentration" n'a pas été révélée à l'avance. Ainsi, comme en 2004, il n'a plus été possible qu'un surentraînement ait une influence notable sur les résultats. Ce principe doit être maintenu.

La suppression des questions d'essai non évaluées a eu pour conséquence que malgré l'introduction d'un nouveau sous-test, le temps nécessaire pour effectuer l'AMS est demeuré comparable à celui des années précédentes. Cette suppression a été accueillie favorablement, car elle élimine tout doute quant à la question de savoir quels items sont évalués et lesquels ne le sont pas. Pour 2006, des modifications structurelles sont à l'étude en vue d'améliorer l'organisation de la pause, la partie du test qui se déroule le matin étant devenue trop longue.

2.5 Egalité de traitement entre groupes linguistiques

Une grande attention est vouée à la traduction des questions du test, de manière à obtenir une **équivalence** maximale des trois versions linguistiques. Il sied notamment de relever dans ce contexte que la rédaction finale des questions relève de la compétence de deux professeurs de gymnase, chacun bilingue et connaissant la terminologie utilisée dans les gymnases suisses en matière de sciences naturelles; de plus, ces deux professeurs sont spécialement formés pour

qu'ils puissent formuler des versions (francophone et italophone) d'un niveau de difficulté équivalent à celui de la version originale (germanophone).

Depuis 1999, on applique une procédure introduite sur le plan international - dite procédure DIF -, qui permet de compenser d'éventuelles différences liées à la langue du test, alors que d'autres différences au niveau des données subsistent. Lorsque, dans une langue, certaines questions comportent des différences qui dépassent nettement une valeur moyenne, ces différences sont corrigées et un bonus est attribué. Au cours des dernières années, le nombre des différences qui ont dû être compensées a fortement diminué, ce qui témoigne de la qualité des traductions.

Lorsque des différences subsistent cependant entre les groupes linguistiques, elles sont parfois attribuées faussement à la qualité de la traduction ou au fait que le test a un "background culturel" allemand; les données dont nous disposons prouvent en effet le contraire:

Depuis l'année 2000, force est de constater que les différences sont beaucoup **plus nettes** dans les tests à **faible composante linguistique** que dans ceux qui ont une forte composante linguistique. Si l'on tablait (uniquement) sur des différences liées à la langue, les différences n'apparaîtraient elles aussi que (ou en tout cas elles apparaîtraient plus nettement) dans les tests à forte composante linguistique, alors que les tests à moindre composante linguistique comporteraient moins de différences.

Grâce à la procédure de correction DIF, les résultats des sous-tests à forte composante linguistique sont même nettement meilleurs que ceux des sous-tests à faible composante linguistique. Il est dès lors possible que les différences aient été quelque peu surcorrigées.

Les possibilités de se préparer au test sont les mêmes dans chaque langue (Test Info, publication des versions originales d'anciens tests, informations pour s'entraîner à exécuter le test dans sa durée originale). Les gymnases fribourgeois ont p.ex. préparé un test d'entraînement en français.

En fin de compte, l'efficacité des différents groupes linguistiques n'est donc pas la même – indépendamment des effets propres à chaque langue. Des facteurs linguistiques sont à l'origine des différences de prestations que l'on constate entre ces groupes; ces différences sont cependant moins grandes que celles engendrées par des différences de prestations sur un plan général. L'équité entre les groupes linguistiques est garantie.

2.6 Différences entre classes d'âge

Au cours des années précédentes, il a été constaté à maintes reprises que les candidats les plus jeunes obtiennent de meilleurs résultats. Cette année, on observe une tendance plus légère dans ce sens: le groupe des candidats d'âge moyen ainsi que celui des candidats plus âgés ayant passé tôt leur maturité ne se distinguent du groupe des plus jeunes candidats que par une différence de moyenne de 1.4. Auparavant, cette différence était p.ex. de 2.4 ou de 2.8.

Cette année, la moyenne du groupe des plus âgés ayant passé tardivement leur maturité est de 3.3 inférieure à celle du groupe des plus jeunes – cette différence était encore de 6 l'année passée.

A noter pour mémoire que le taux de réussite aux examens propédeutiques suit à peu près la même courbe; c'est donc bien ce que prédit le test.

2.7 Différences entre femmes et hommes

Il est notoire que la probabilité de réussite au premier examen est de 7% plus élevée chez les hommes que chez les femmes. Il faut que l'on retrouve cette différence au niveau du test, s'il est censé la prédire. Il convient en outre de relever dans ce contexte que le nombre de femmes candidates aux études de médecine est nettement supérieur à celui des hommes; cette différence de représentativité entre femmes et hommes ne se retrouve pas au niveau des candidats à la maturité ni dans le groupe plus âgé.

En médecine humaine et en médecine dentaire, on constate une différence de moyenne de test de 2 en faveur des hommes. Cette différence est comparable à celle des années précédentes; comme en 2004, elle est significative en médecine humaine, seulement en raison du nombre élevé de candidats. En médecine vétérinaire, la moyenne de test des hommes est inférieure de 4 à celle des femmes.

Dans l'ensemble, aucun élément ne va à l'encontre de l'égalité des chances; les résultats de cette année confirment ceux des années précédentes.

2.8 Différences entre cantons et taux de maturité

Il s'agit ici de données concernant des candidatures aux études de médecine dont la représentativité par rapport aux autres bacheliers du canton est variable. Dans les cantons bilingues, les francophones et les bilingues peuvent également choisir de commencer leurs études de médecine à Genève, Lausanne ou Neuchâtel sans être soumis à un *numerus clausus*. De même, certains phénomènes "socio-psychologiques" peuvent varier d'un canton à l'autre (voire d'une année à l'autre dans un même canton), au sens où une année, par exemple, seuls les meilleurs bacheliers décident d'étudier la médecine, alors qu'une autre fois, il n'y a pas un tel "effet de groupe".

Comme lors des années précédentes, il n'est **pas possible d'établir un lien** entre taux de maturité et résultats au test (les 3 groupes de cantons réunis en fonction du taux de maturité ne se différencient pas de manière significative). Les tendances mises en évidence pendant des années – dans les cantons ayant un nombre suffisant de candidats pour une analyse statistique – demeurent stables.

2.9 Conclusions

La huitième session de l'AMS s'est à nouveau déroulée avec succès. Les résultats obtenus sont tout aussi fiables que ceux des années précédentes.

L'AMS fournit, pour le *numerus clausus*, un instrument de sélection éprouvé afin de n'admettre aux études de médecine que les personnes les plus aptes en fonction des places disponibles et de procéder à des transferts au moyen d'un critère objectif.

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

3 Neues zu Eignungstest und Numerus Clausus

Die Nachfrage nach einem Medizinstudium steigt seit nunmehr 5 Jahren in der Schweiz permanent an, der Arztberuf wird offenbar immer attraktiver. Vor dem Hintergrund der wirtschaftlichen Entwicklung und ihrer Risiken in anderen Bereichen wird Arzt wahrscheinlich als eine relativ sichere Berufswahl erlebt. Der häufig thematisierte Bedarf an Krankenhausärzten sowie die Verbesserung der Arbeitsbedingungen hinsichtlich der Arbeitszeit bei Bereitschaftsdiensten sind Argumente, die den Arztberuf zusätzlich erstrebenswert machen.



Diese Entwicklung ist jedoch nicht spezifisch für die Schweiz. Der Austria Presse Agentur entnehmen wir die nebenstehende Abbildung. In der überwiegenden Zahl der Länder ist ein Numerus Clausus für Medizin notwendig. Unterschiede zwischen den Ländern ergeben sich vor allem daraus, ob weitere Fächer betroffen sind. Diese Sonderstellung des Medizinstudiums ergibt sich aus den besonderen Anforderungen: Es handelt sich um eine Berufsausbildung, erfordert teure Ressourcen und auch die Zahl der Patienten in den Universitätskliniken ist für eine ausreichende praktische Ausbildung nicht beliebig steigerbar.

In **Deutschland** ist die Nachfrage nach Studienplätzen für Humanmedizin in den letzten 3 Jahren ebenfalls von ca. 3000 auf ca. 5000 Personen angestiegen (siehe Abbildung 1). Dank eines NC kann man hier kapazitätsentsprechend zulassen – die Ablehnungsquote (4 von 5 Personen) ist allerdings deutlich höher als in der Schweiz. Bezüglich der Zulassungskriterien ist eine Entwicklung zurück zu spezifischen Eignungstests zu erkennen, die durch die veränderte Gesetzeslage gefördert wird, welche festlegt, dass die Hochschulen einen Teil ihrer Studierenden selbst auswählen können. Bisher ist es allerdings noch nicht wieder gelungen, eine praktikable Lösung analog des früheren TMS einzuführen, weil die zur Verfügung stehenden Mittel stark begrenzt und keine koordinierenden Strukturen vorhanden sind.

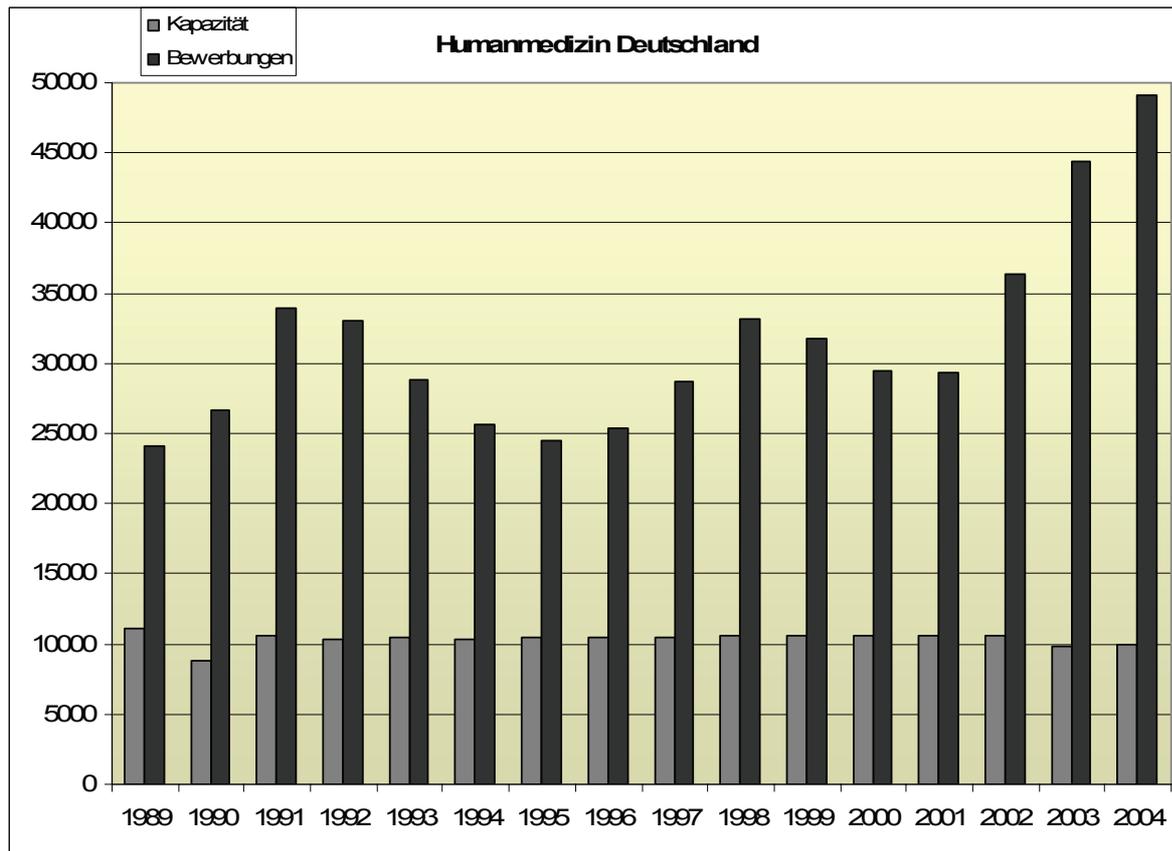


Abbildung 1: Bewerbungen und Kapazitäten Humanmedizin in Deutschland (zusammengefasst jeweils das WS und das SS des darauf folgenden Jahres – also 2004 = WS 04/05 und SS 05). Quelle: Zentralstelle für die Vergabe von Studienplätzen (ZVS) www.zvs.de.

In **Österreich** gab es vor 2005 noch keinen Numerus Clausus. Durch eine Zunahme der eigenen Bewerberzahlen und vor allem durch den Zustrom deutscher Studierwilliger, die aufgrund des Gleichbehandlungsgrundsatzes der europäischen Gesetzgebung nunmehr ebenfalls zuzulassen sind, werden auch dort die Kapazitäten so weit überschritten, dass Handlungsbedarf besteht. Man erwägt einen Eignungstest nach Schweizer Vorbild einzuführen.

Summa summarum ist die Schweiz durch den funktionierenden NC, der durch einen geeigneten Test als Zulassungskriterium realisiert werden kann, gut gerüstet für den Erhalt der Studienqualität im Fach Medizin bei steigender nationaler oder möglicher internationaler Nachfrage.

Seine Vorteile kann man wie folgt zusammenfassen:

- **Rechtsfähigkeit** der Ergebnisse durch Chancengleichheit und Willkürfreiheit (Forderung des Bundesgerichtes für einen Numerus Clausus), indem alle Abläufe standardisiert sind und eine laufende Evaluation hinsichtlich Alters-, Geschlechts-, Regional- und Sprachbesonderheiten erfolgt.
- **Prognose der Studieneignung** (als gesetzliche Grundlage des NC) ist evaluiert, Studiendauer und Noten werden nachweislich gut vorhergesagt.
- **Jährlich eine neue Testversion**, die vorher nicht bekannt ist (keine Vorteile für irgendwelche Insider, was die Rechtsfähigkeit einschränken würde), dadurch werden auch wiederholte Antritte problemlos möglich.
- Prinzipiell wird kein spezielles Fachwissen vorausgesetzt, das empfohlene **Informationskonzept** und die offiziell angebotenen Vorbereitungsmöglichkeiten sind nachweis-

lich genauso effektiv wie der Besuch professioneller Vorbereitungskurse (ist ausreichend evaluiert) – auch als Voraussetzung für die soziale Verträglichkeit.

Den Test kann man nur erfolgreich verwenden **und auf die Schweiz oder Deutschland Bezug nehmen**, wenn man das ganze System der Testanwendung beachtet. Es wurde in Deutschland bereits so angewendet, in der Schweiz weiterentwickelt und an die aktuellen Bedingungen laufend angepasst:

- **Empirisch ausreichend erprobte Aufgaben:** Mindestens als Einstreuaufgaben in einem Originaltest oder bereits unter Ernstfallbedingungen eingesetzte Aufgaben werden verwendet. Dies gewährleistet seit Jahren hohe und vergleichbare Testgütekriterien.
- **Standardisierung der Abnahmebedingungen und Auswertung:** Richtlinien für die Organisation der Testabnahme – gleiche Bedingungen an allen Orten inkl. präventiver Empfehlungen für besondere Vorkommnisse, Testleiter-Handbuch mit genauen Durchführungsvorschriften, erprobte und schnelle Auswertung mittels Belegleser, verständliches und kapazitätsgenaues System zur Studienplatzzuweisung als Kombination von Testwert und Rangplatz.
- **Ausreichende Information der Kandidaten:** Test Info als Broschüre ½ Jahr vorher an alle Kandidaten zur Vorbereitung, veröffentlichte Originalversionen zum Vertrautmachen mit dem Test, Handreichungen für Testleiter zur Durchführung von Probeläufen (an Gymnasien zentral angeboten), Info und Aufklärung im Internet (www.unifr.ch/ztd/ems).
- **Langfristige Garantie der Qualität:** Ausreichender Bestand an erprobten Aufgaben, der Wiederholungen – wenn überhaupt – nur in grossen Zeitabschnitten notwendig macht, Gewährleistung der Vergleichbarkeit der Testversionen von Jahr zu Jahr, Weiterentwicklung des Tests und Anpassung an neue Studienanforderungen („Planen und Organisieren“, weitere Massnahmen in Vorbereitung).

4 Numerus Clausus (NC) und Medizinstudium

Die Abbildung 2 verdeutlicht die Anmeldeverläufe 1991 bis 1997 (ohne NC, hellere Balken) sowie 1998 bis 2005 (mit NC, dunklere Balken). Die letzte Säule zeigt die Kapazitäten 2005 und das 20%-Überschreitungskriterium, ab welchem ein NC zur Anwendung kommt.

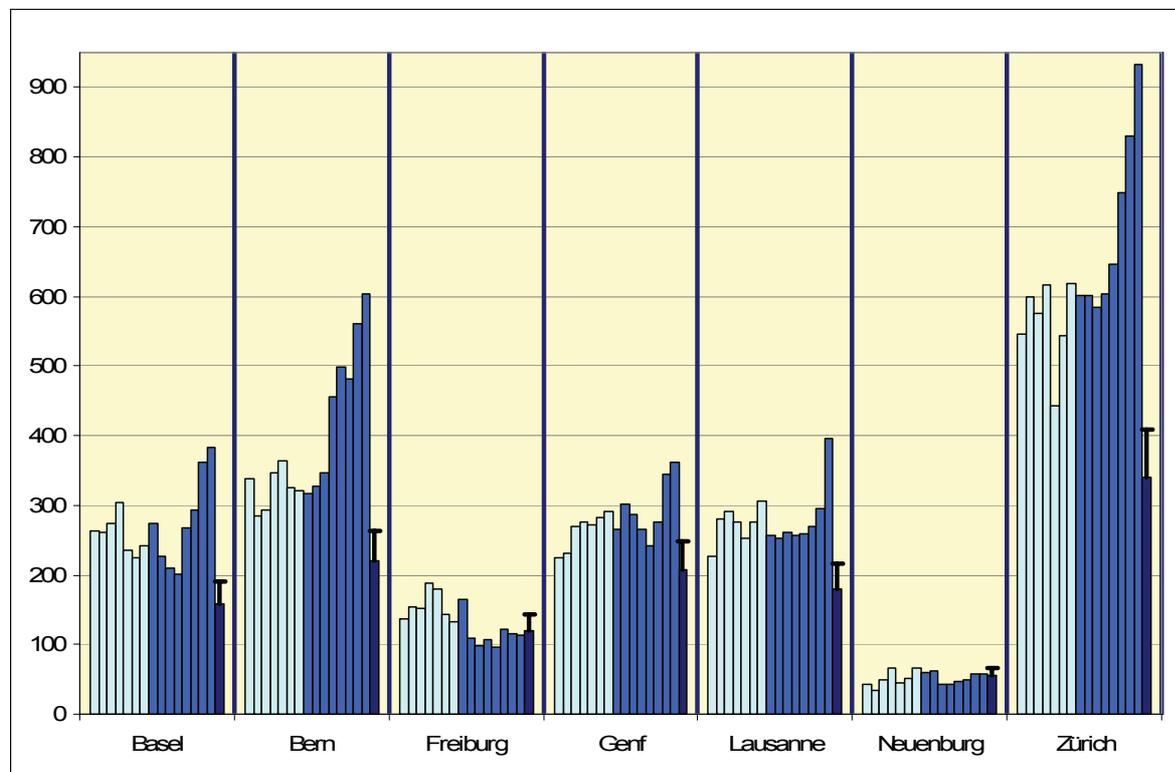


Abbildung 2: Anmeldungen zum Medizinstudium 1991 bis 2005, dunklere Säulen: Jahre mit NC; letzte Säulen: Kapazität 2005 und Markierung 20%-Überschreitungskriterium als Grenze für NC.

Die Anmeldezahlen in Basel, Bern und Zürich steigen in den letzten Jahren deutlich. Ein Zusammenhang mit den Jahrgängen der Doppelmaturität ist nicht mehr erkennbar, die Nachfrage nimmt unabhängig davon zu. Auch in der Westschweiz setzt sich die Zunahme der Nachfrage fort, sodass sich die Bedingungen im ersten Studienjahr (vor dem Zeitpunkt einer intrauniversitären Selektion) weiter verschlechtern. Es ist davon auszugehen, dass die Attraktivität des Medizinstudiums bzw. Arztberufes weiter zunimmt.

In der Abbildung 3, in welcher die Universitäten danach zusammengefasst sind, ob sie zum „NC-Pool“ gehören oder nicht, wird der Trend noch deutlicher.

Die Abbildung 4 zeigt, dass nicht mit einer Entspannung der Situation zu rechnen ist. Das Bundesamt für Statistik (BfS) veröffentlichte eine Prognose für die Studierendenzahlen der Schweiz. Eine wichtige Aussage ist, dass der nach den Jahrgängen mit Doppelmaturitäten vorhandene Nachfrageschub nicht mehr absinkt, das hohe Niveau bleibt erhalten.

Setzt man die Zahl der Anmeldungen bei der CRUS (bzw. früher der SHK) im Jahre 1998 zu 100%, ergibt sich zunächst für 1999 und 2000 ein relativer Rückgang der Anmeldezahlen (Abbildung 4). Dass es sich hier um einen Dissuasionseffekt durch die Neueinführung des EMS handelt, wird durch das Gleichbleiben bzw. leichte Ansteigen der Studieneintritte 1999 und 2000 nahe gelegt.

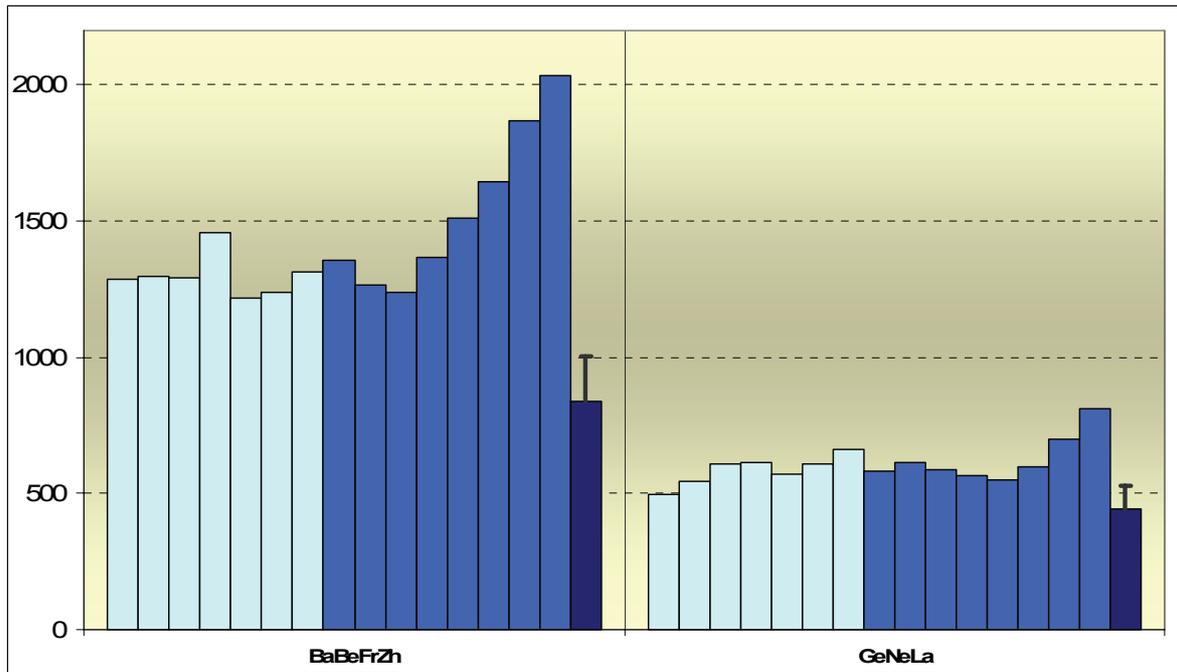


Abbildung 3: Anmeldungen zum Medizinstudium 1991 bis 2005 nach Gruppen (**B**asel, **B**ern, **F**reiburg, **Z**ürich vs. **G**enf, **N**euenburg, **L**ausanne). Dunklere Säulen: Jahre mit NC, letzte Säule Kapazität 2005 und Markierung 20%-Überschreitungskriterium als Grenze für NC.

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

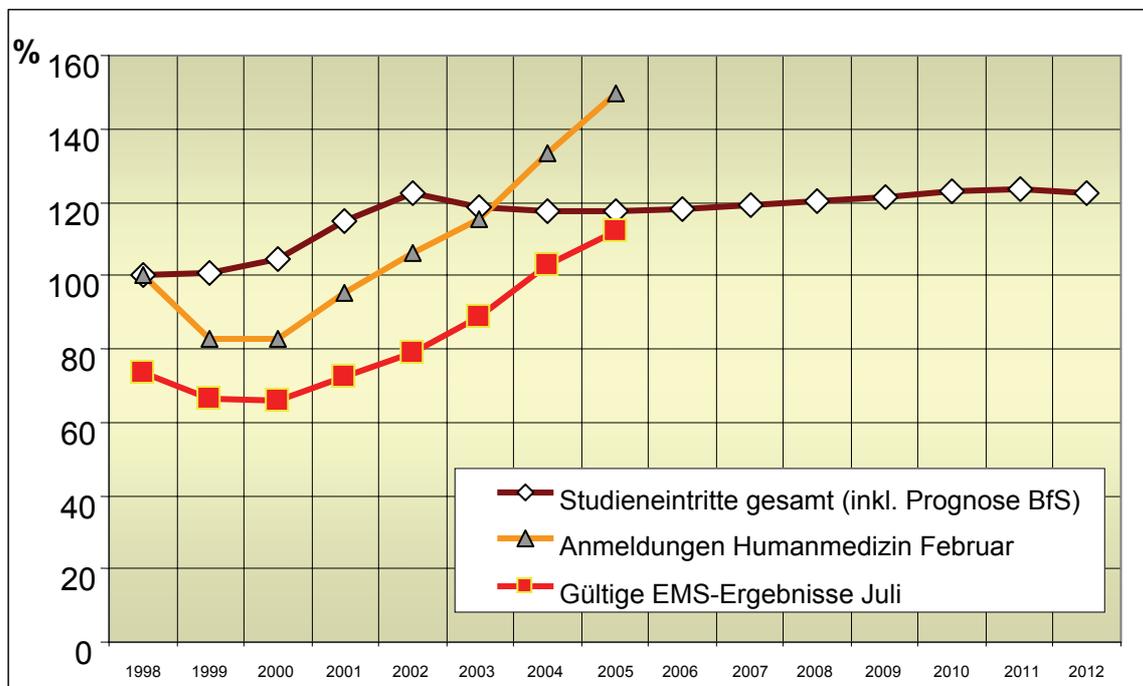


Abbildung 4: Studieneintritte an Universitäten und Hochschulen (indexiert, 1998 = 100%) im Vergleich zu Studienanmeldungen Februar und Bewerbungen mit gültigem EMS (indexiert, Anmeldungen 1998 = 100%) für Humanmedizin.

5 Anmeldung zu Medizinstudium und Test 2005

5.1 Anmeldungen, Testantritte und Zulassungen

Erstmals sind, trotz einer ebenfalls hohen Rückzugsquote gegenüber dem Stand der Anmeldungen im Februar, mehr als zwei Bewerbungen auf einen Studienplatz in Humanmedizin vorhanden.

	Humanmedizin							
	1998	1999	2000	2001	2002	2003	2004	2005
Kapazitäten (am NC beteiligte Univ.)	620	546	546	583	648	598	546	546
Anmeldungen bei CRUS im Februar	1020	845	844	974	1080	1177	1360	1525
... in % zu Kapazität	186	155	154	167	185	197	249	279
Anmeldungen EMS im Mai	797	717	678	752	827	917	1083	1182
+ Übernahme Testergebnis aus Vorjahr	0	5	33	27	40	47	38	41
= Total Bewerbungen Mai	797	722	711	779	867	964	1121	1223
... in % zu Kapazität	129	132	130	134	134	161	205	224
Erfolgreiche Absolvierung EMS	750	673	637	709	770	871	1026	1107*
Gesamtbewerbungen Juli EMS / Ergebnis Vorjahr	750	678	670	736	805	907	1051	1143
Rückzugsquote zwischen Februar und Juli	26.5	19.8	20.6	24.4	25.5	22.9	22.7	25.0
Prozent Gesamtbewerbungen, die Studienplatz erhalten	82.7	80.5	81.4	79.2	80.5	65.9	51.9	47.8
NC-Index Bewerbungen/Studienplatz	1.21	1.24	1.23	1.25	1.24	1.52	1.92	2.09

Tabelle 1: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; 1998 bis 2005 für Humanmedizin. (*inkl. 5 Testteilnehmer, die von Zahn- zu Humanmedizin in einem höheren Semester wechseln wollen und nicht als Bewerbungen zählen).

	Veterinärmedizin							Zahn- medizin	
	1999	2000	2001	2002	2003	2004	2005	2004	2005
Kapazitäten (am NC beteiligte Univ.)	150	150	165	185	170	150	150	142	142
Anmeldungen bei CRUS im Februar	231	213	217	254	271	273	318	208	190
... in % zu Kapazität	153	142	132	137	159	182	212	146	134
Anmeldungen EMS im Mai	194	176	179	206	225	207	257	163	159
+ Übernahme Testergebnis aus Vorjahr	0	5	5	3	11	13	8	1	11
= Total Bewerbungen Mai	194	181	184	209	236	220	265	164	170
... in % zu Kapazität	129	121	112	113	139	147	177	115	120
Erfolgreiche Absolvierung EMS	182	164	171	203	217	196	244	151	150
Gesamtbewerbungen Juli EMS / Ergebnis Vorjahr	182	169	176	205	228	209	252	152	161
Rückzugsquote zwischen Februar und Juli	21.2	20.7	18.9	19.3	15.9	23.4	20.8	26.9	15.3
Prozent Gesamtbewerbungen, die Studienplatz erhalten	82.4	88.7	93.7	90.2	74.5	71.8	59.5	93.4	88.1
NC-Index Bewerbungen/Studienplatz	1.21	1.13	1.06	1.11	1.34	1.39	1.68	1.07	1.13

Tabelle 2: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich; 1999 bis 2005 für Veterinärmedizin und 2004 bis 2005 für Zahnmedizin.

Im Fach Veterinärmedizin kam es zu einer Zunahme der Nachfrage, die Kapazitätsüberschreitung wird auch hier deutlicher (Tabelle 2).

Trotz einer gegenüber 2004 geringeren Zahl der Anmeldungen im Februar, ist die Nachfrage im Fach Zahnmedizin gestiegen. Ursache ist eine deutlich geringere Rückzugsquote zwischen dem Zeitpunkt der Anmeldung und der Testabsolvierung.

Die nachfolgende Tabelle 3 vergleicht Anmeldungen, Bewerbungen und Rückzüge nach Disziplinen und Universitäten. Die Abbildungen 5 bis 9 verdeutlichen die Anmeldesituationen nach Disziplinen. Summa summarum wird deutlich, dass die Verläufe zwischen Anmeldung im Februar, Testanmeldung und erfolgreicher Testabsolvierung, trotz der Zunahme der Bewerberzahlen, prozentual in etwa stabil bleiben. Nach wie vor kommt es aber in dieser Zeit zu Rückzügen in der Höhe von ca. 20%, die wir auch auf Dissuasionseffekte der „Bewährungssituation Test“ zurückführen, welcher sich nicht alle Interessenten stellen.

	Studienbeginn 2002			Studienbeginn 2003			Studienbeginn 2004			Studienbeginn 2005		
	Anmel- dungen am 15.2.	Pers. m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers. m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers.m. gültigem Testwert	Rück- züge in %	Anmel- dungen am 15.2.	Pers.m. gültigem Testwert	Rück- züge in %
Basel HM	239	181	24.3	256	155	23.8	299	247	17.4	350	267	22.3
Bern HM	305	245	19.7	294	234	20.4	357	277	22.4	396	299	24.5
Freiburg HM	87	66	24.1	113	84	25.7	107	89	16.8	105	82	21.9
Zürich HM	449	313	30.3	518	394	23.9	597	450	24.6	674	495	26.6
Total	1080	805	25.5	1181	907	23.2	1360	1063	21.8	1525	1143	24.7
Bern VM	132	106	19.7	133	108	18.8	137	111	19.0	152	121	20.4
Zürich VM	122	99	18.9	141	120	14.9	136	98	27.9	166	131	21.2
Total	254	205	19.3	274	228	16.8	273	209	23.4	318	252	20.8
Basel ZM	kein NC bzw. kein EMS						45	30	33.3	32	31	3.1
Bern ZM							65	52	20.0	56	47	16.1
Freiburg ZM							7	6	14.3	9	7	22.8
Zürich ZM							91	63	30.8	93	76	18.3
Total							208	151	27.4	190	161	15.3

Tabelle 3: Anmelungsverlauf für Universitäten, Human- (HM), Veterinär- (VM) und Zahnmedizin (ZM). Personen mit gültigem Testwert.

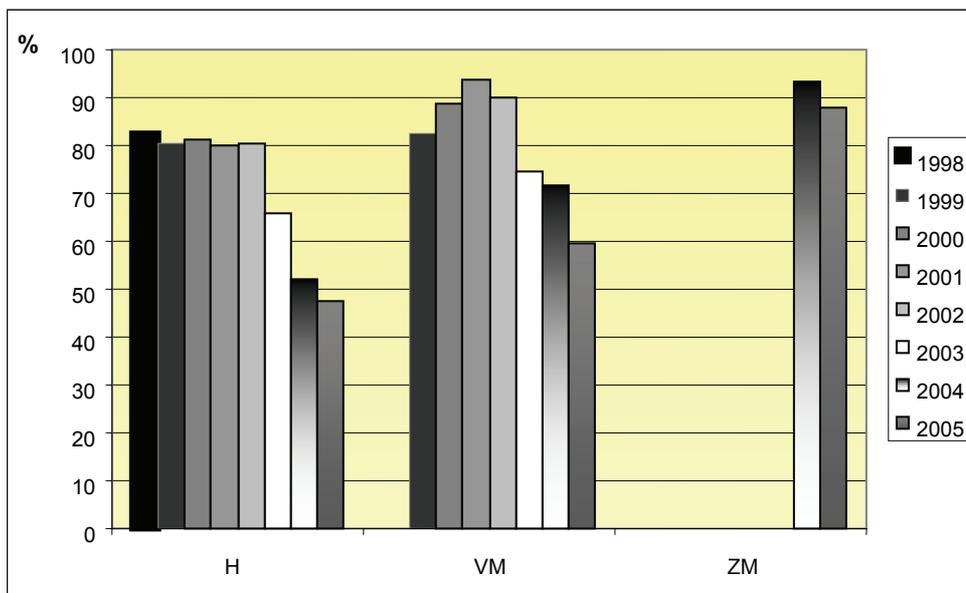


Abbildung 5: Prozent der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann – HM: Humanmedizin; VM: Veterinärmedizin; ZM: Zahnmedizin.

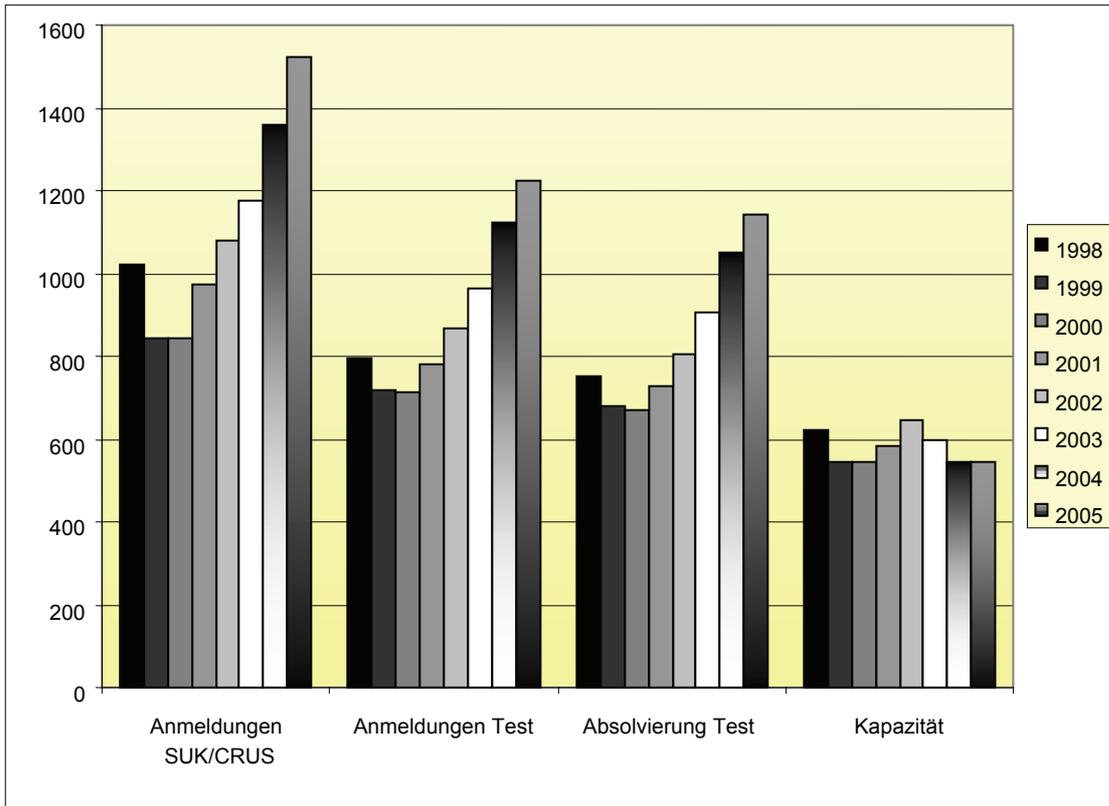


Abbildung 6: Anmeldeverlauf für Humanmedizin 1998 bis 2005 im Vergleich (mit NC).

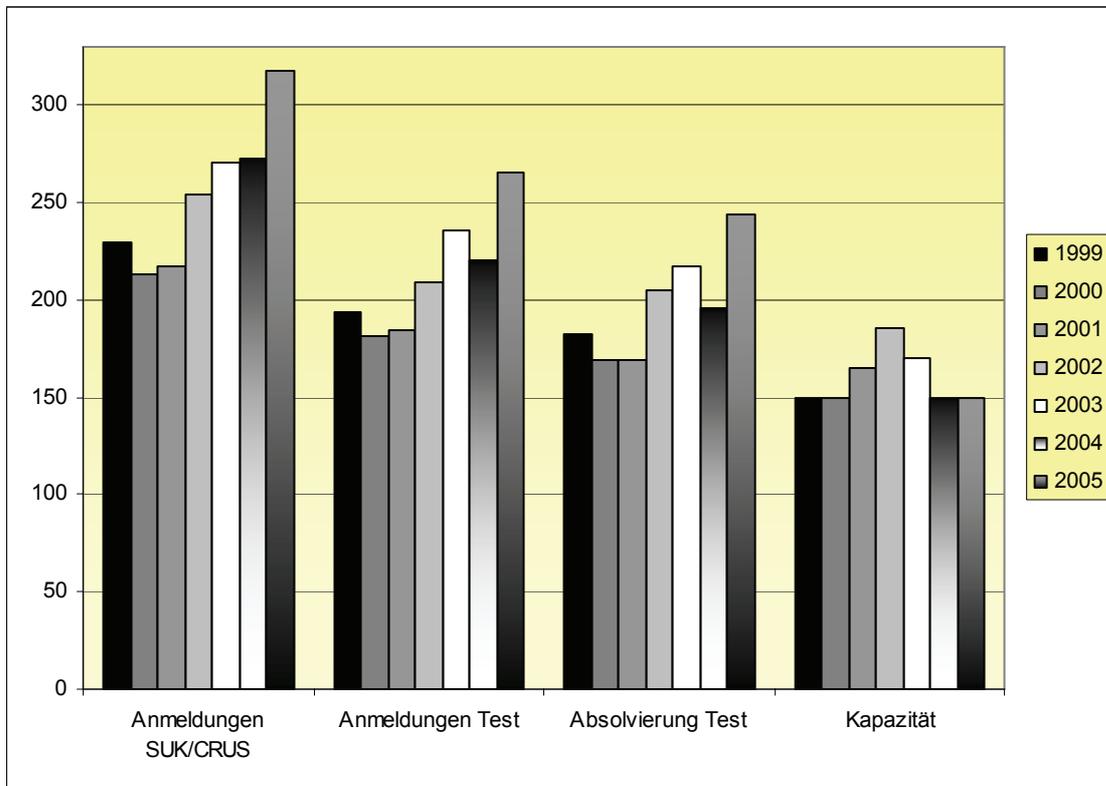


Abbildung 7: Anmeldeverlauf für Veterinärmedizin 1999 bis 2005 im Vergleich (mit NC).

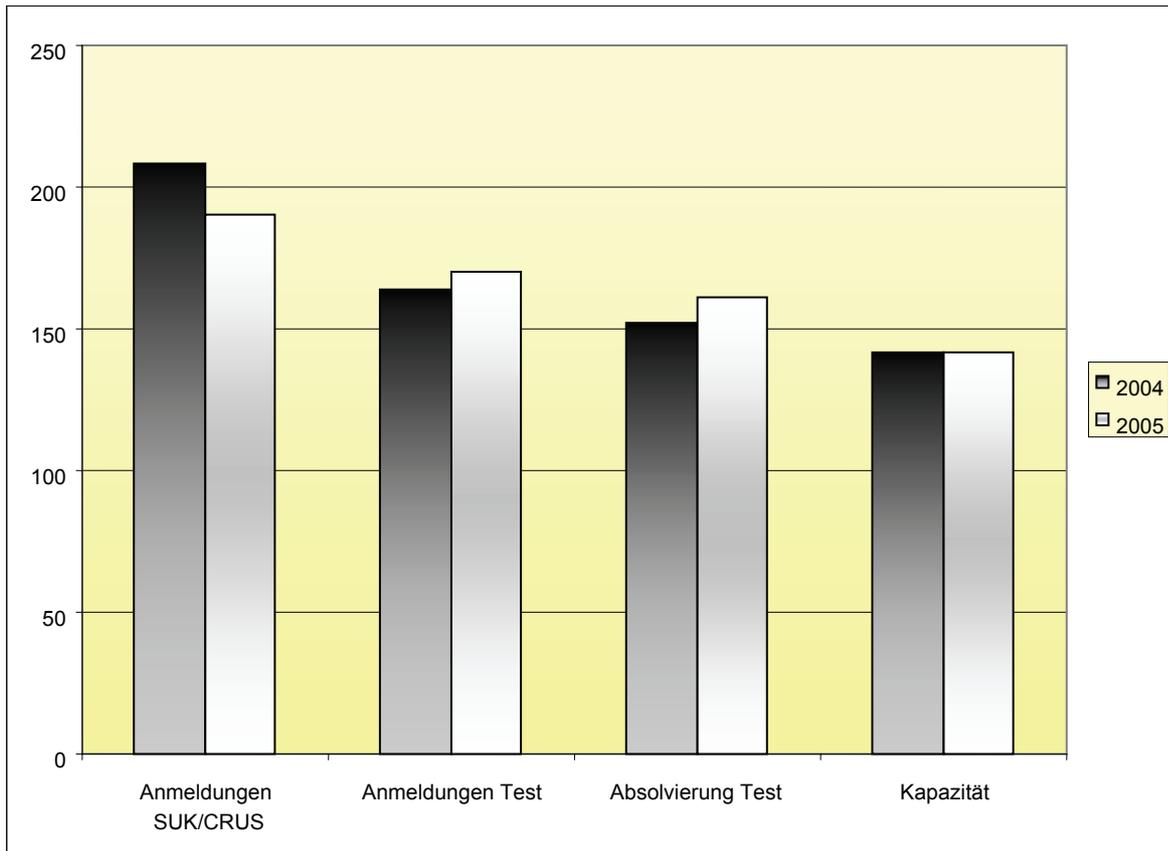


Abbildung 8: Anmeldeverlauf für Zahnmedizin 2004 und 2005 (mit NC).

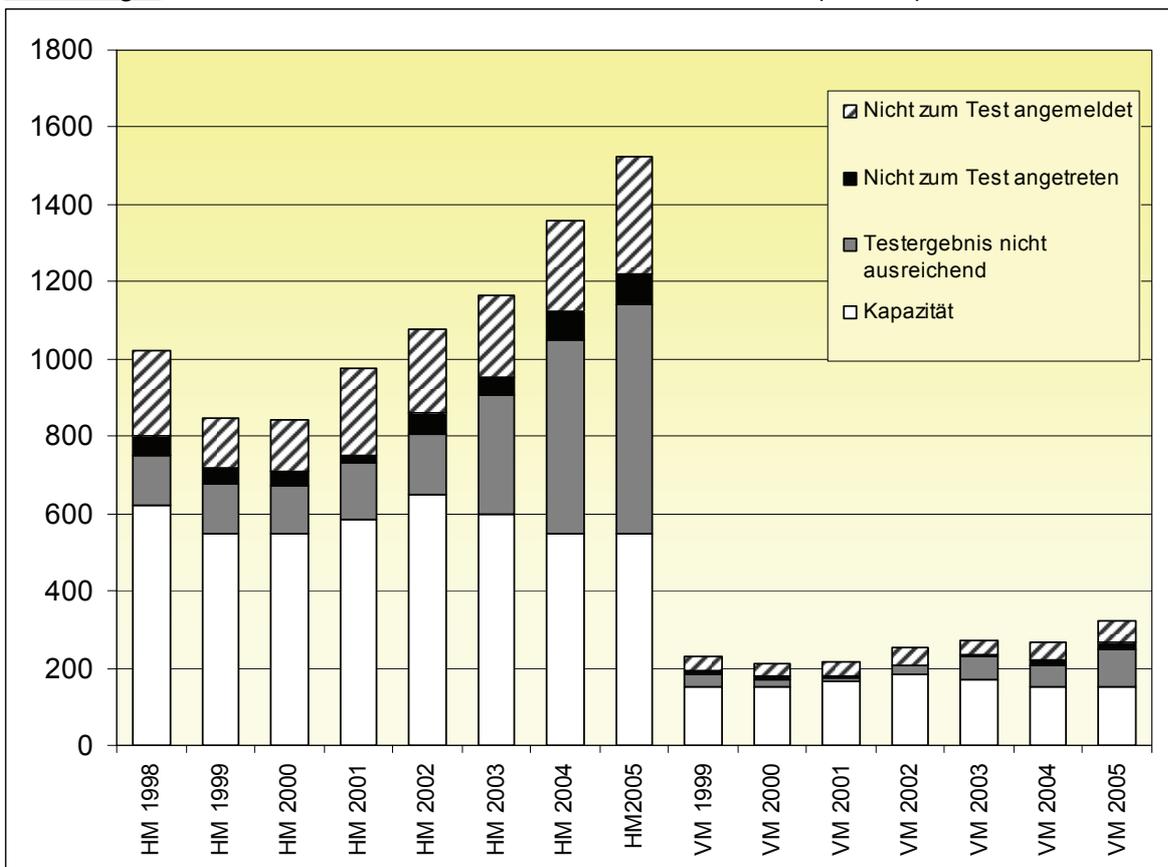


Abbildung 9: Kapazität, Anmeldungen und Testteilnahme für Human- (HM) und Veterinärmedizin (VM). Aufteilung der Voranmeldungen (Februar des jeweiligen Jahres).

5.2 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme in 18 Testlokalen verteilt auf 8 Testorte. Aarau wurde 2005 erstmals zusätzlich angeboten und die Testabnahmestelle hat ihre Bewährungsprobe hervorragend bestanden – nicht zuletzt durch eine intensive Vorbereitung, die bereits im vergangenen Jahr begann.

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

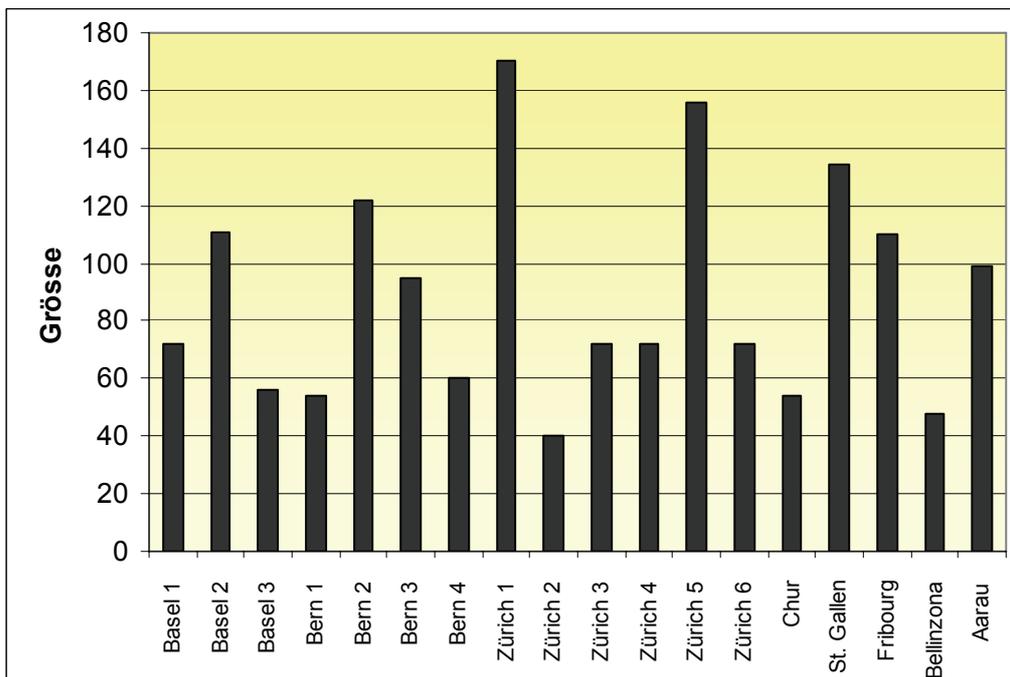


Abbildung 10: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2005.

5.3 Testorte und Wunschuniversitäten

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

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

Testort	Wunschuniversität				
	Bern	Basel	Freiburg	Zürich	Gesamt
AARAU	28 6.2%	38 12.9%	0 -	30 4.5%	96 6.4%
BASEL	13 2.9%	175 59.5%	0 -	21 3.1%	209 13.9%
BELLINZONA	9 2.0%	4 1.4%	2 2.4%	31 4.6%	46 3.1%
BERN	268 58.9%	13 4.4%	21 25.3%	8 1.2%	310 20.7%
CHUR	15 3.3%	9 3.1%	1 1.2%	25 3.7%	50 3.3%
FREIBURG	57 12.5%	2 0.7%	42 50.6%	11 1.6%	112 7.5%
ST. GALLEN	26 5.7%	20 6.8%	8 9.6%	76 11.4%	130 8.7%
ZÜRICH	39 8.6%	33 11.2%	9 10.8%	467 69.8%	548 36.5%
Gesamt	455	294	83	669	1501

Tabelle 4: Testorte und Wunschuniversität (erste Wahl) 2005.

5.4 Wunschuniversitäten und Testorte nach Wohnkantonen

Nachfolgend sind die Testorte im Verhältnis zum Wohnkanton für die einzelnen Disziplinen tabelliert. Angegeben sind auch die Vergleichswerte seit 1998 bzw. für die Jahre mit NC.

Wohnkanton/ Wohnort	Voranmeldung HM 2005					Total						
	BE	BS	FR	ZH	Total	2004	2003	2002	2001	2000	1999	1998
AG	6	39	0	36	81	93	54	52	42	62	62	66
AI	1	0	0	0	1	4	0	2	1	3	0	2
AR	2	1	0	1	4	8	6	6	8	4	6	4
BE	163	2	8	1	174	179	135	139	148	88	87	101
BL	0	63	0	1	64	54	44	37	41	38	52	50
BS	3	55	0	3	61	48	41	35	31	26	37	40
FR	5	0	45	4	54	42	44	28	24	33	39	34
GE	0	0	0	0	0	0	1	1	1	0	0	1
GL	0	1	0	4	5	1	6	3	3	3	4	4
GR	11	8	0	15	34	41	46	18	30	16	22	19
JU	0	0	0	0	0	3	1	0	0	0	1	0
LU	33	17	1	14	65	53	52	62	42	29	34	43
NE	2	0	0	1	3	1	1	1	0	0	1	2
NW	0	4	0	0	4	5	9	4	3	2	3	6
OW	0	3	0	2	5	5	2	7	3	9	2	3
SG	12	9	7	53	81	55	48	50	41	51	51	64
SH	2	4	0	5	11	6	10	10	15	5	4	7
SO	15	24	0	3	42	42	38	35	21	13	22	26
SZ	1	1	1	16	19	18	11	11	15	16	16	21
TG	7	5	1	25	38	24	22	28	30	20	16	22
TI	5	2	2	17	26	34	30	18	31	21	23	30
UR	0	2	0	1	3	1	3	4	4	2	1	3
VD	0	1	1	1	3	1	2	1	2	2	0	3
VS	8	0	1	1	10	7	5	13	12	11	12	22
ZG	5	1	0	14	20	17	12	16	19	6	4	6
ZH	7	12	7	244	270	263	234	178	130	169	159	165
FL	2	4	0	3	9	5	3	3	5	4	2	3
Übriges Ausland	1	7	2	10	20	16	11	8	7	4	13	3
Total	291	265	76	475	1107	1026	871	770	709	637	673	750

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

Wohnkanton/ Wohnort	Vor Anmeldung VM 2005			Total					
	Bern	Zürich	Total	2004	2003	2002	2001	2000	1999
AG	4	21	25	11	13	8	7	9	22
AI	0	0	0	1	0	3	0	1	0
AR	0	2	2	2	2	0	2	0	0
BE	23	1	24	29	32	30	39	26	26
BL	5	5	10	12	8	7	4	3	7
BS	4	5	9	4	2	4	7	2	2
FR	8	1	9	6	8	9	3	11	9
GE	6	0	6	2	4	6	9	11	13
GL	0	4	4	0	1	1	0	2	0
GR	1	5	6	10	13	4	6	3	7
JU	0	0	0	4	4	2	1	1	1
LU	4	2	6	11	7	13	5	12	10
NE	10	0	10	4	7	6	3	6	4
NW	1	1	2	1	2	2	0	0	0
SG	4	15	19	14	13	9	7	4	10
SH	1	0	1	1	2	0	3	3	0
SO	5	2	7	7	7	8	4	7	6
SZ	1	1	2	4	2	4	0	2	2
TG	2	6	8	5	11	6	8	3	4
TI	4	6	10	6	9	9	7	6	6
UR	0	1	1	2	1	1	1	0	0
VD	28	3	31	15	15	13	21	15	11
VS	5	0	5	7	4	7	7	7	3
ZG	0	3	3	4	6	3	4	3	6
ZH	0	39	39	29	42	46	19	25	32
FL	0	0	0	0	0	1	2	0	0
Übriges Aus- land	3	2	5	1	2	1	2	2	1
Total	119	125	244	196	217	203	171	164	182

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

Wohnkanton Wohnort	Vor Anmeldung Zahnmedizin (ZM) 2005					2004
	Bern	Basel	Freiburg	Zürich	Total	
AG	2	3	0	5	10	14
AI	0	0	0	1	1	0
AR	2	0	0	0	2	4
BE	25	0	2	0	27	31
BL	0	10	1	0	11	10
BS	1	5	0	0	6	4
FR	0	0	4	0	4	1
GL	0	0	0	3	3	0
GR	0	1	0	2	3	8
LU	2	2	0	1	5	2
NW	1	0	0	0	1	1
OW	0	0	0	0	0	2
SG	3	0	0	9	12	4
SH	0	0	0	5	5	3
SO	1	1	0	0	2	9
SZ	0	1	0	0	1	2
TG	1	1	0	7	9	6
TI	1	1	0	1	3	8
UR	1	1	0	0	2	0
VS	4	0	0	0	4	4
ZG	0	0	0	1	1	1
ZH	1	0	0	32	33	33
FL	0	1	0	1	2	2
Übriges Aus- land	0	2	0	1	3	2
Total	45	29	7	69	150	151

Tabelle 7: Gewünschte Studienorte Zahnmedizin nach Wohnkantonen und Vergleich zu 2004. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die nachfolgende Tabelle zeigt, welche Testorte aus den einzelnen Kantonen gewählt werden und dient auch der Kapazitätsplanung für die Testorte.

Wohnkanton	Testort								Total
	AARAU	BASEL	BELLIN-ZONA	BERN	CHUR	FRIBOURG	ST. GALLEN	ZÜRICH	
AG	57	13	0	1	0	0	0	45	116
AI	0	0	0	0	0	0	1	1	2
AR	0	0	0	0	0	0	8	0	8
BE	0	1	0	220	0	3	0	1	225
BL	0	83	0	2	0	0	0	0	85
BS	0	74	0	1	0	0	0	1	76
FR	0	0	0	22	1	44	0	0	67
GE	0	0	0	0	0	6	0	0	6
GL	0	0	0	0	1	0	0	11	12
GR	0	1	4	1	32	0	0	5	43
LU	18	5	0	19	0	0	0	34	76
NE	0	0	0	1	0	12	0	0	13
NW	5	0	0	0	0	0	0	2	7
OW	1	1	0	0	0	0	0	3	5
SG	0	0	0	1	11	0	88	12	112
SH	0	0	0	0	0	0	1	16	17
SO	13	13	0	23	0	0	0	2	51
SZ	1	1	0	0	0	0	0	20	22
TG	0	1	0	0	0	0	19	35	55
TI	0	0	38	0	0	0	1	0	39
UR	0	2	0	0	0	0	0	4	6
VD	0	0	0	1	0	33	0	0	34
VS	0	0	0	13	0	5	0	1	19
ZG	0	0	1	1	0	0	1	21	24
ZH	1	9	2	1	0	2	0	327	342
FL	0	0	0	1	4	0	6	0	11
Ausland	0	5	1	2	1	7	5	7	28
Gesamt	96	209	46	310	50	112	130	548	1501

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

5.5 Testabsolvierung nach Alter und Geschlecht

Disziplin Humanmedizin		Geschlecht				Total
		männlich		weiblich		
		Anzahl	%	Anzahl	%	
Geburtsjahre	1951-1981	45	42.9%	60	57.1%	105
	1982-1984	110	43.5%	143	56.5%	253
	1985-1988	253	33.8%	496	66.2%	749
	Total	408	36.9%	699	63.1%	1107
Disziplin Veterinärmedizin		Geschlecht				Total
		männlich		weiblich		
		Anzahl	%	Anzahl	%	
Geburtsjahre	1951-1981	3	9.4%	29	90.6%	32
	1982-1984	7	12.5%	49	87.5%	56
	1985-1988	16	10.3%	140	89.7%	156
	Total	26	10.7%	218	89.3%	244
Disziplin Zahnmedizin		Geschlecht				Total
		männlich		weiblich		
		Anzahl	%	Anzahl	%	
Geburtsjahre	1951-1981	7	33.3%	14	66.7%	21
	1982-1984	21	65.6%	11	34.4%	32
	1985-1988	33	34.0%	64	66.0%	97
	Total	61	40.7%	89	59.3%	150
Alle Bewerbungen		Geschlecht				Total
		männlich		weiblich		
		Anzahl	%	Anzahl	%	
Geburtsjahre	1951-1981	55	34.8%	103	65.2%	158
	1982-1984	138	40.5%	203	59.5%	341
	1985-1988	302	30.1%	700	69.9%	1002
	Total	495	33.0%	1006	67.0%	1501

Tabelle 9: Testabsolventen nach Geburtsjahr und Geschlecht 2005.

In Human- und Zahnmedizin bleibt der Frauenanteil bei ca. 60% stabil, in Veterinärmedizin weiter bei ca. 90%. Die Zunahme der Bewerberzahlen rekrutiert sich vor allem aus dem Personenkreis der Jüngeren, die sofort nach der Maturität Medizin studieren wollen (Tabelle 9, Abbildungen 11-15).

In absoluten Zahlen nehmen die Bewerbungen von älteren Personen allerdings auch nicht ab (vergleiche Abbildung 15).

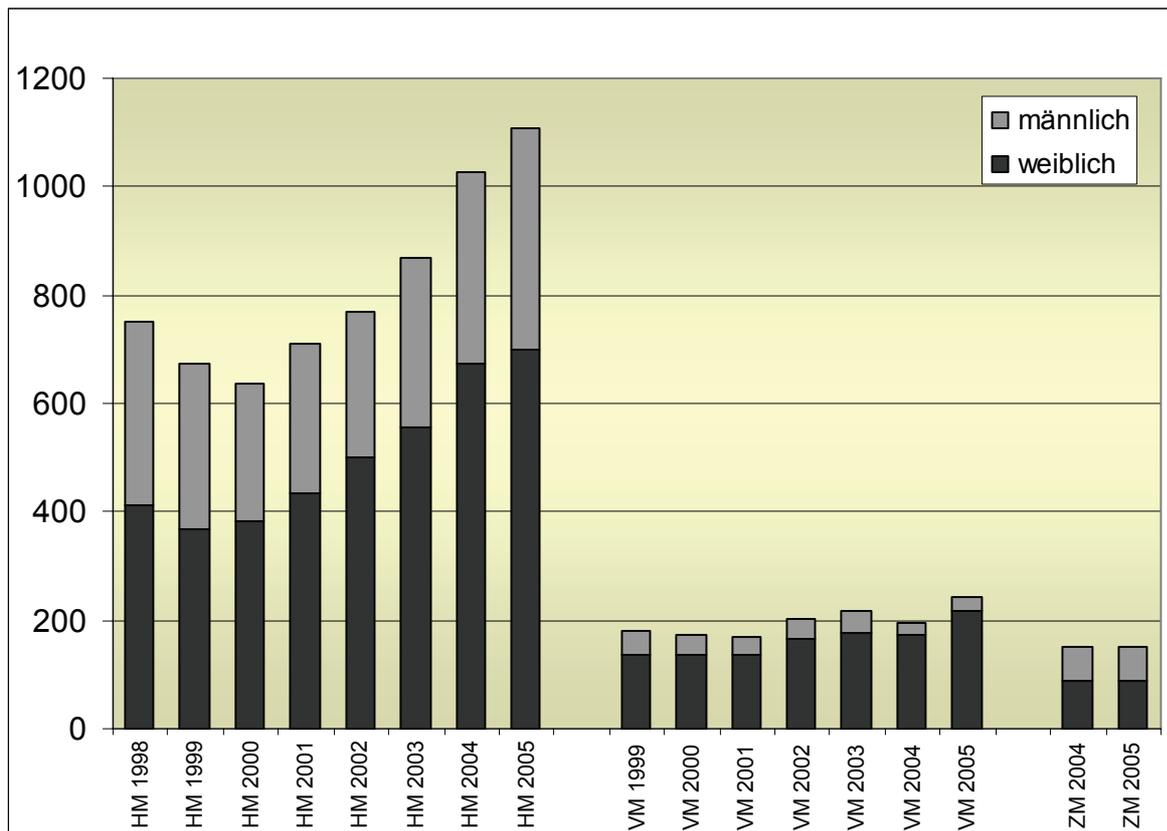


Abbildung 11: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM), Veterinärmedizin (VM) und Zahnmedizin (ZM), Jahre mit NC, Absolutzahlen.

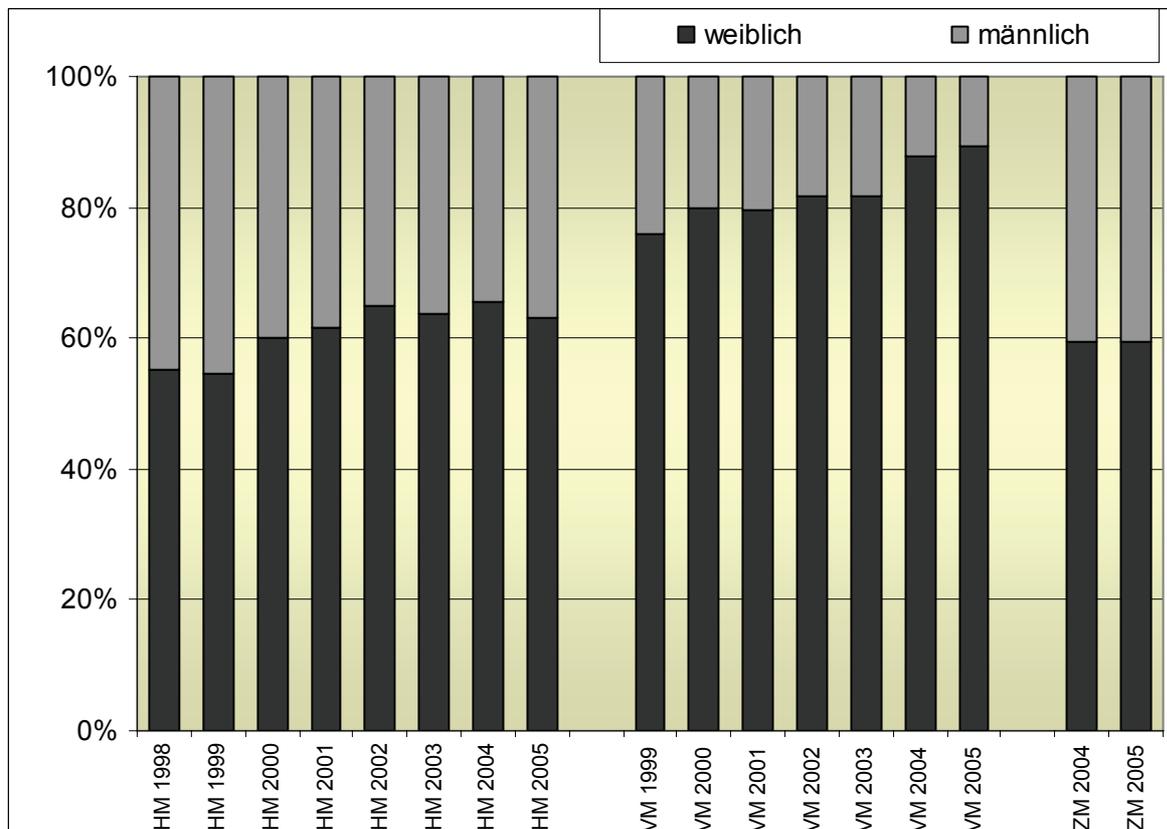


Abbildung 12: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM), Veterinärmedizin (VM) und Zahnmedizin (ZM), Jahre mit NC, bezogen auf 100% pro Jahr und Disziplin.

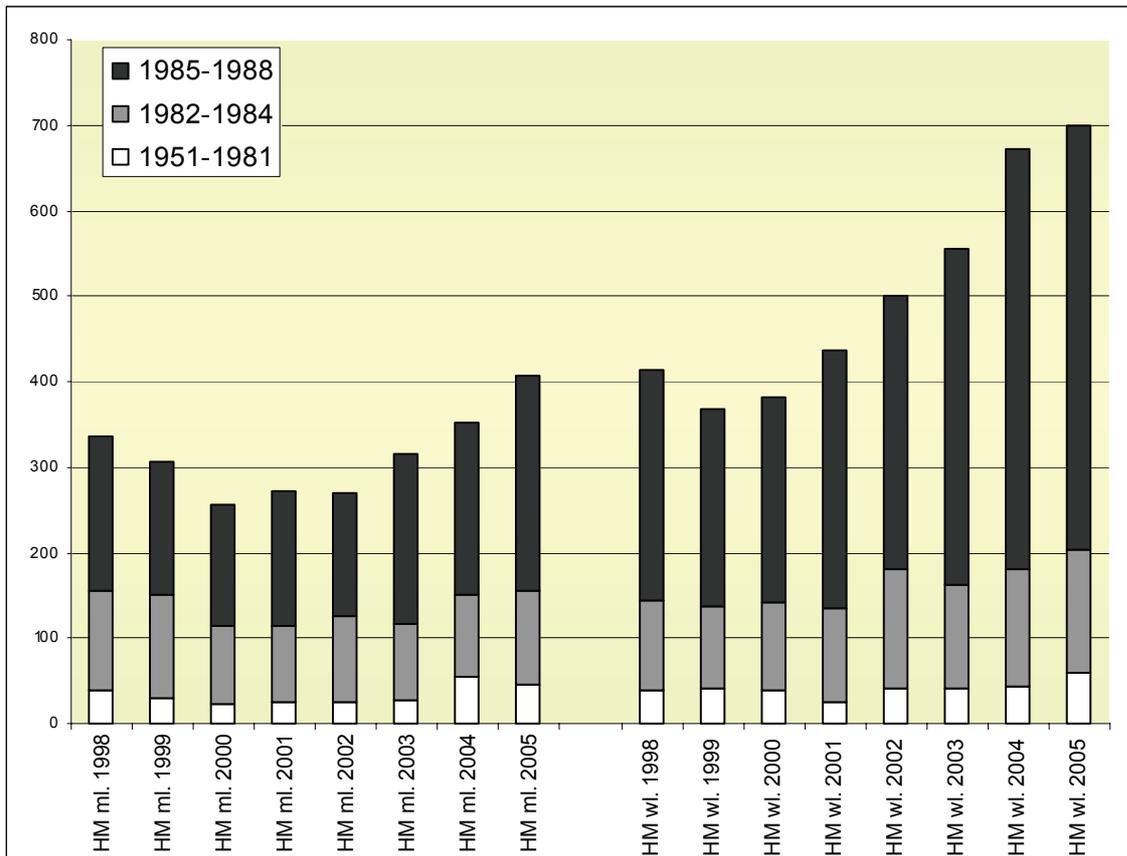


Abbildung 13: Zusammensetzung der Bewerbergruppen nach Geschlecht mi: männlich, wl: weiblich und Alter (Geburtsjahre), Absolutzahlen.

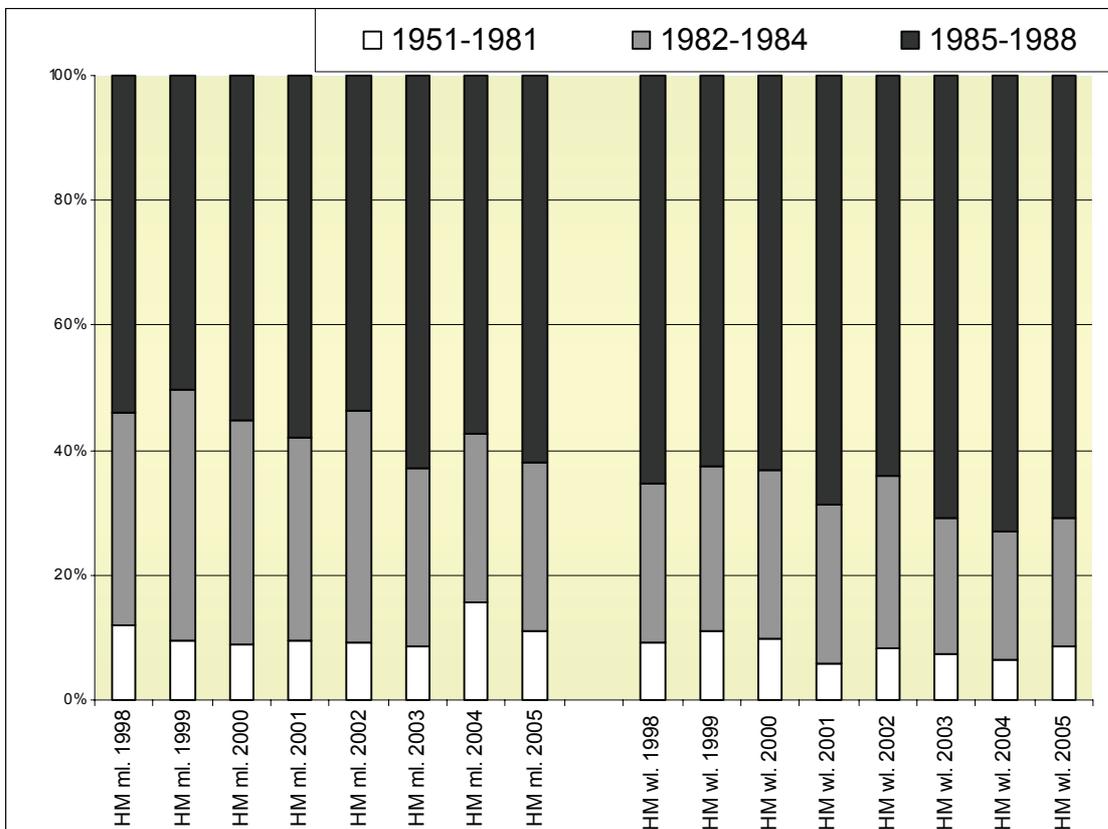


Abbildung 14: Zusammensetzung der Bewerbergruppen nach Geschlecht mi: männlich, wl: weiblich und Alter (Geburtsjahre) für Humanmedizin bezogen auf 100% pro Kohorte.

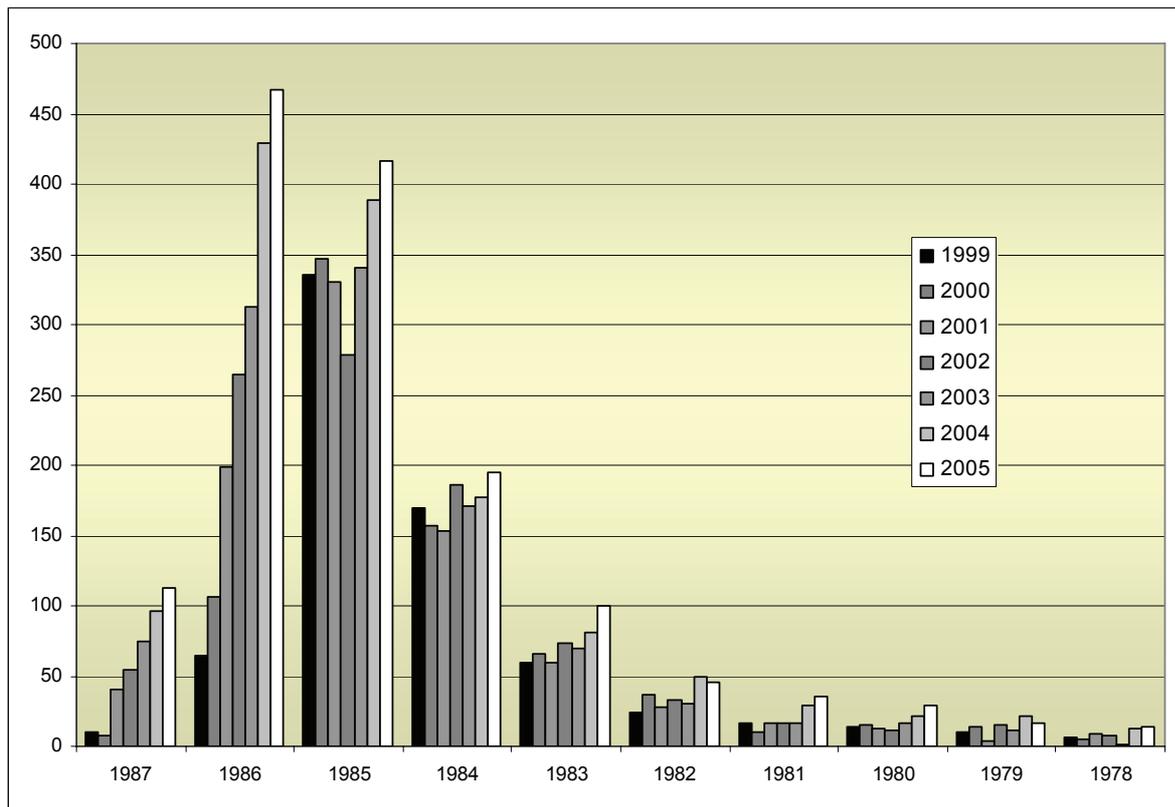


Abbildung 15: Bewerbungen nach Geburtsjahren für 2005 und Entsprechungen für die Jahre 1999 bis 2004 (jeweils um entsprechende Jahre versetzt).

5.6 Übernahme des Testergebnisses aus Vorjahren

Bekanntlich kann man das Ergebnis des Vorjahres in das Folgejahr übertragen lassen oder neu zum Test antreten (wobei dann immer das neue Ergebnis zählt).

Die Zahl der Übernahmen der Testergebnisse aus dem Vorjahr bleibt in etwa konstant (Tabelle 10). Dabei ist zu beachten, dass in Human- und Veterinärmedizin wegen der zunehmenden Kapazitätsüberschreitung die Grenzwerte des Testwertes für die Zulassung von Jahr zu Jahr steigen. Es kam zu Situationen, in welchen ein Testwert, der im Vorjahr noch ausreichend war, sich im Folgejahr als zu gering für eine Zulassung erwies. Deshalb treten viele Personen lieber erneut zum EMS an – vielleicht auch, weil die Wahrscheinlichkeit für eine Verbesserung beim zweiten Testantritt bisher immer ziemlich hoch war. Bei 8 Personen hat 2005 der übertragene Testwert nicht für eine Zulassung ausgereicht.

Die Anfragen nach einer optimalen Strategie, die uns zwischen Februar und Mai erreichen, nehmen zu – dabei können seitens des ZTD aber keine „garantierten“ Ratschläge gegeben werden. Hier hilft nur die Faustregel, dass man aufgrund der Februar-Zahlen in etwa prognostiziert, wie viel Prozent der Kandidaten einen Studienplatz in der jeweiligen Disziplin erhalten werden, wobei die Rückzüge zwischen Februar und der Testanmeldung zu diesem Zeitpunkt nur geschätzt werden können. Da die Testwerte und hier besonders die Prozentränge jedes Jahr das gleiche bedeuten, kann man dann zumindest grob abwägen, ob der Prozentrang noch berücksichtigt würde und wie sicher das in etwa ist.

Disziplin	Jahr	N	Minimum	Maximum	Mittelwert	Stand.-abw.	Zugelassen 1. Runde
Human- medizin	2000	33	85	119	102.9	7.25	33
	2001	27	94	119	103.0	7.28	27
	2002	40	83	118	104.0	9.16	40
	2003	47	91	121	103.7	8.27	44 (3 nicht)
	2004	37	90	119	105.9	7.70	31 (6 nicht)
	2005	41	73	120	105.7	8.22	35 (6 nicht)
Vete- rinär- medizin	2000	5	97	107	102.0	3.81	5
	2001	5	89	112	102.8	9.52	5
	2002	3	94	109	101.3	7.51	3
	2003	11	86	118	102.9	9.0	10 (1 nicht)
	2004	13	80	106	99.1	7.1	12 (1 nicht)
	2005	8	85	117	105.2	10.93	6 (2 nicht)
Zahn- medizin	2004	1			93.0		1
	2005	11	78	99	91.7	7.10	11

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

5.7 Erneute Testteilnahme 2005

Erste Testteil- nahme	Zweite Testteilnahme						
	1999	2000	2001	2002	2003	2004	2005
1998	12	2	1	1	1	1	0
1999		19	3	0	0	2	1
2000			12	0	3	1	1
2001				11	3	1	2
2002					32	13	1
2003						78	21
2004							151
Gesamt	12	21	16	12	39	96	177

Tabelle 11: Wiederholte Testteilnahme: Anzahl der Personen.

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

Verschärfung des NC nimmt die Zahl der erneuten Antritte zu – das Risiko, dass ein knapp über der Zulassungsgrenze liegender Wert in einem Jahr dann im nächsten Jahr nicht mehr reicht, wird grösser. Bisher war die Wahrscheinlichkeit eines besseren 2. Testergebnisses auch deutlich höher. Dieser Unterschied scheint sich weiter zu nivellieren – es treten zwar immer noch mehr Verbesserungen auf, die Wahrscheinlichkeit einer Verschlechterung wird relativ aber grösser.

	N	Minimum	Maximum	Mittelwert	Stand.- abweichung
Testwert 1. Teilnahme	177	74	116	93.7	6.6
Testwert 2. Teilnahme	177	79	120	101.2	8.4
Veränderungen	177	-20	34	7.6	9.3

Tabelle 12: Testwertvergleich zwischen erster und zweiter Testteilnahme.

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. Durch die bessere Vorbereitung beim ersten Mal ist das Niveau der optimalen Leistung bereits bei der ersten Teilnahme erreicht.

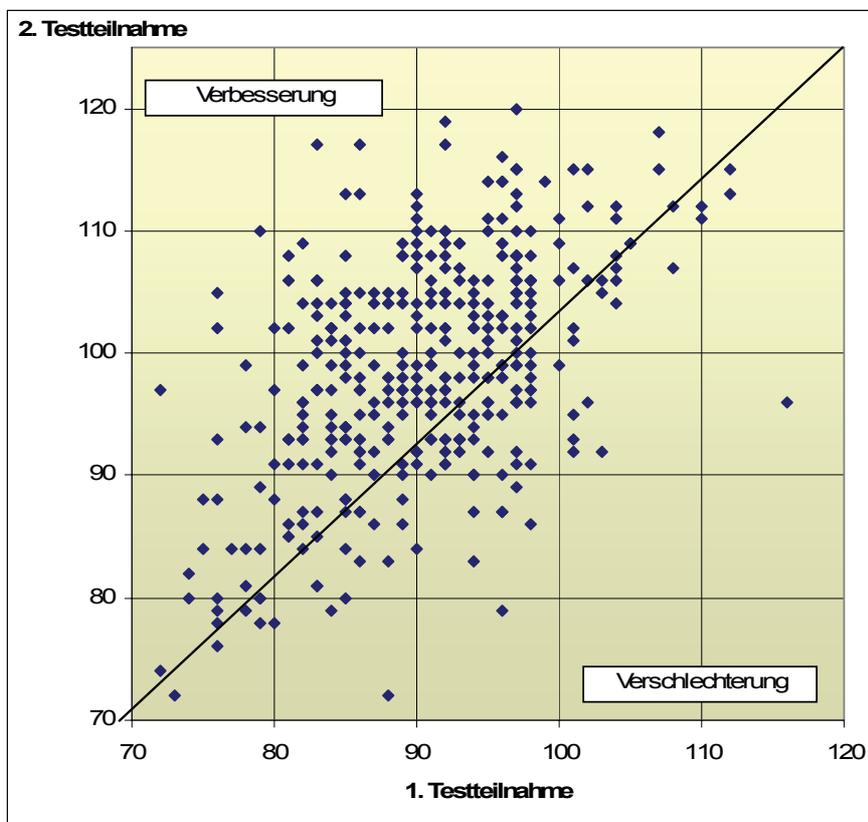


Abbildung 16: Testwer-

te für Personen, die zweimal am Test teilnahmen (1998 bis 2005).

5.8 Sprachgruppen

Nach wie vor treten vor allem deutschsprachige Kandidaten zum EMS an, für die französisch- und italienischsprachigen existieren Alternativen das Studium ohne NC, in der gewünschten Sprache vollständig zu absolvieren. Die Zahl der französisch- und italienischsprachigen Kandidaten bleibt über die Jahre in etwa vergleichbar.

Nach wie vor spielt auch die Chancengleichheit der Sprachgruppen eine bedeutende Rolle:

- Aufwendiges und angemessenes Verfahren zur Herstellung der Äquivalenz der Sprachversionen (die Evaluation zeigt, dass sprachbedingte Unterschiede nicht wahrscheinlich sind),
- vergleichbares Vorbereitungsmaterial in der jeweiligen Sprache (Test Info, veröffentlichte Originalversion),
- seit 2005 verstärkte Orientierung auf Probelaufe in der jeweiligen Sprache, was durch eine Kurz-Anleitung für die Testleiter dieser Probelaufe seitens des ZTD unterstützt wird.

			Geburtsjahre			Total
			1951-1981	1982-1984	1985-1988	
Humanmedizin	deutsch	männlich	44	102	226	372
		weiblich	59	131	462	652
		Total	103	233	688	1024
	französisch	männlich	1	6	14	21
		weiblich		9	20	29
		Total	1	15	34	50
	italienisch	männlich		2	13	15
		weiblich	1	3	14	18
		Total	1	5	27	33
Veterinärmedizin	deutsch	männlich	2	5	9	16
		weiblich	21	41	96	158
		Total	23	46	105	174
	französisch	männlich	1	2	6	9
		weiblich	8	8	35	51
		Total	9	10	41	60
	italienisch	männlich			1	1
		weiblich			9	9
		Total			10	10
Zahnmedizin	deutsch	männlich	7	20	33	60
		weiblich	13	11	61	85
		Total	20	31	94	145
	französisch	männlich		1		1
		weiblich			1	1
		Total		1	1	2
	italienisch	männlich				
		weiblich	1		2	3
		Total	1		2	3

Tabelle 13: Sprachgruppen 2005 nach Alter und Geschlecht aufgeschlüsselt.

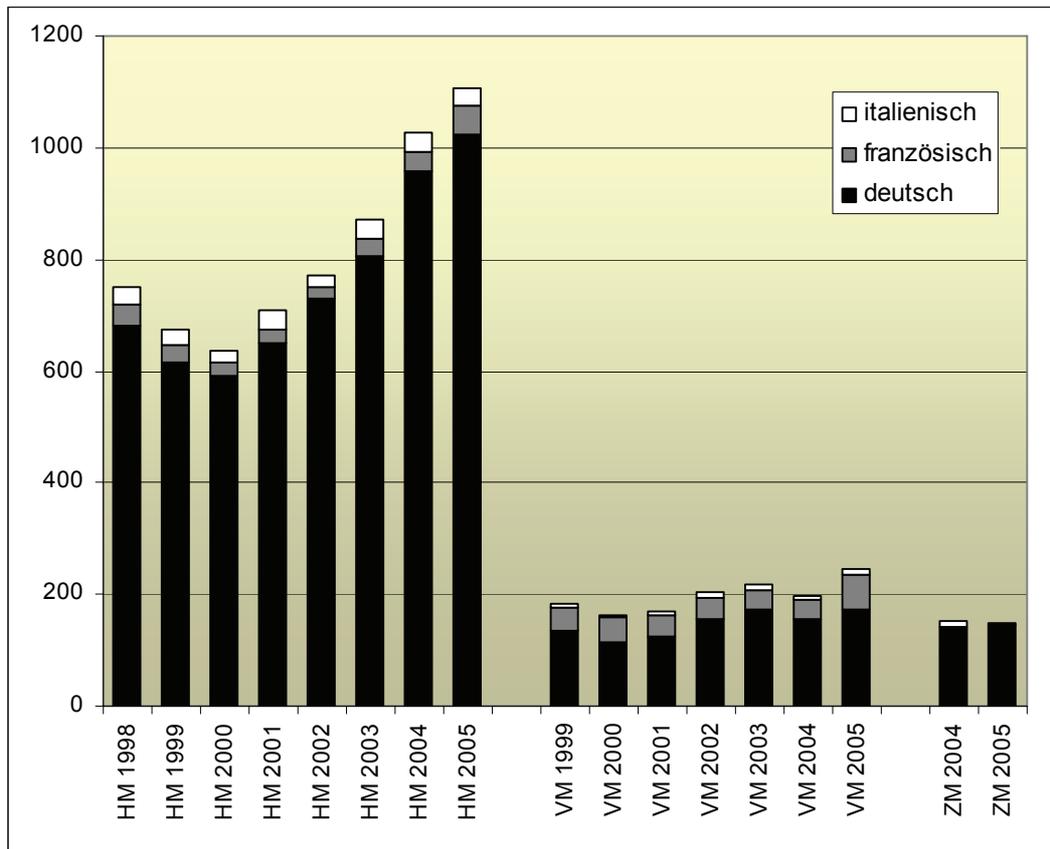


Abbildung 17: Vergleich der Sprachgruppen 1998 bis 2005, Absolutzahlen.

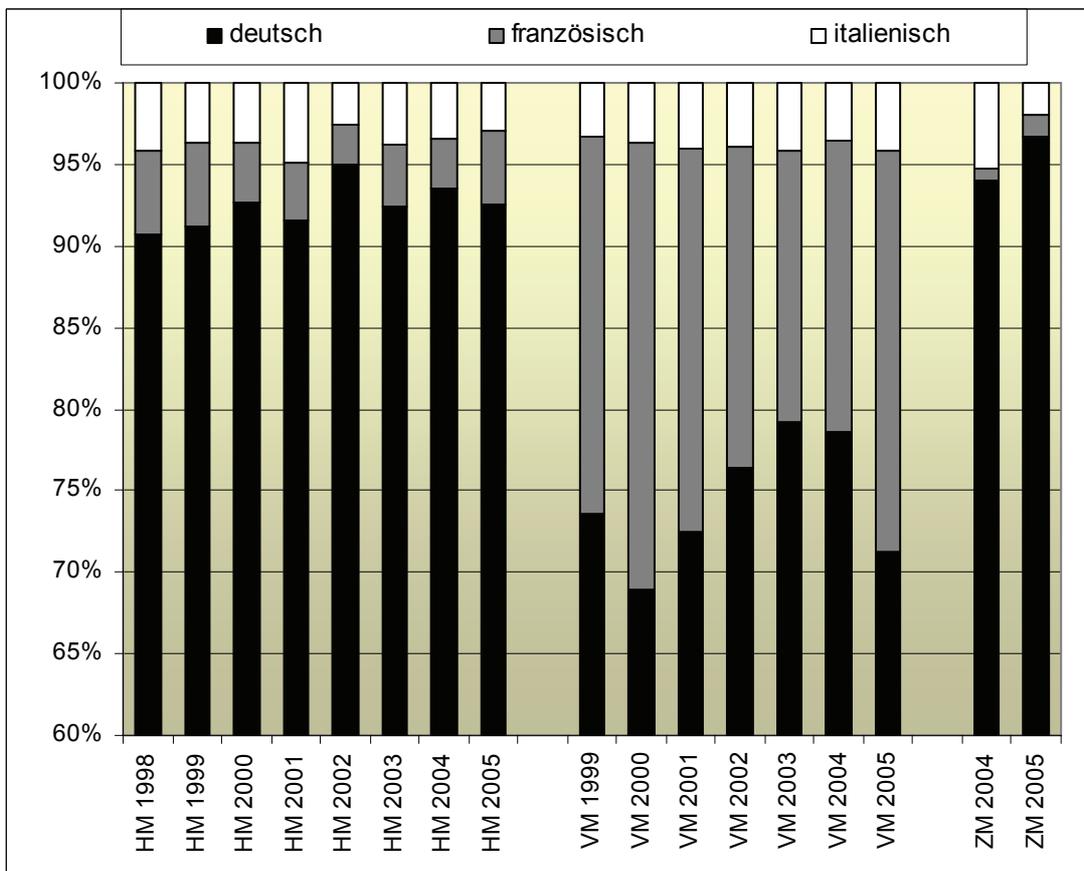


Abbildung 18: Vergleich der Sprachgruppen 1998 bis 2005, Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%).

5.9 Alter und Maturitätsjahr

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

Die Personen werden später analog der Vorjahre zu Vergleichszwecken in drei Altersgruppen aufgeteilt. Die älteste Gruppe wird wiederum danach unterschieden, ob die Maturitätsprüfung unmittelbar nach der Mittelschule absolviert wurde oder noch eine Wartezeit dazwischen bestand. Bei 2 Personen fehlt eine Angabe zum Maturitätsjahr, diese Personen werden hier nicht berücksichtigt.

Geb.- Jahr	Maturitätsjahr																		Total	
	83	84	86	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04		05
51																		1		1
58																			1	1
59																			1	1
63																		2		2
64	2																			2
65		1																1		2
66			1																1	2
67								1											2	3
70																			1	1
71					2	1													1	4
72				1			1									1		1		4
73						1													3	4
74									2	1								1		4
75									2	1				1			1	2	2	9
76														1		1		2	3	7
77											3	4	1	1				1	3	13
78											1	5	2				3	3		14
79												2	6	1	2			3	3	17
80													2	6	5		2	4	10	29
81													1		10	7	2	3	13	36
82														1	8	12	11	3	11	46
83														1	2	21	28	31	17	100
84															1	35	99	60		195
85																14	213	189		416
86																2	76	389		467
87																	1	1	111	113
88																			6	6
Total	1	2	1	1	2	1	1	1	2	5	2	2	0	8	4	6	7	14	14	1499

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

6 Beschreibung des verwendeten Eignungstests

6.1 Aufbau des Tests

Der EMS besteht ab 2005 aus zehn Untertests, die zu einem Gesamtwert verrechnet werden. „Planen und Organisieren“ ist als neuer Untertest dazugekommen. Die Tabelle 15 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Untertests. Aufgabenbeispiele finden Sie ab Seite 110.

Bezeichnung der Untertests	Aufgaben	Max. Punktzahl	Bearbeitungszeit
Muster zuordnen	20	20	18 min
Medizinisch-naturwissenschaftliches Grundverständnis	20	20	50 min
Schlauchfiguren	20	20	12 min
Quantitative und formale Probleme	20	20	50 min
Textverständnis	18	18	45 min
Konzentriertes und sorgfältiges Arbeiten	Gesondertes Blatt mit 1600 Zeichen	20	8 min
	Pause		1 Std.
<i>Lernphase zu den Gedächtnistests</i>			
Figuren lernen	Es werden die Vorlagen zum Einprägen gezeigt		4 min
Fakten lernen			6 min
Planen und Organisieren	20	19*	60 min
<i>Gedächtnistests:</i>			
Figuren reproduzieren	20	20	5 min
Fakten reproduzieren	20	20	7 min
Diagramme und Tabellen	20	20	50 min
Gesamttest	198	198	ca. 5 Std.
Gesamtdauer (inkl. Pause) 9.30 bis ca. 17.30 Uhr			

Tabelle 15: Struktur und Ablauf des EMS 2005; *Eine Aufgabe wurde nicht gewertet.

Beim EMS handelt es sich um eine Adaptation des deutschen Tests für Medizinische Studiengänge (TMS) mit Weiterentwicklungen für die Schweiz. Die Struktur des TMS mit ursprünglich 9 Untertests hat sich in Deutschland im Rahmen von 12 Testeinsätzen bewährt (1986 zwei Testtermine und von 1987 bis 1996 je ein Termin pro Jahr). In der Schweiz wird der Test seit 1998 erfolgreich eingesetzt – bis 2003 mit identischer Struktur wie beim deutschen TMS.

Seit 2004 wird beim Konzentrationstest jährlich eine Vorlage verwendet, die vorher nicht bekannt ist (Zeichen und Durchstreichregel) – nur der Typ der Anforderung und die Zeitdauer

bleiben jeweils gleich. Dadurch werden Effekte von exzessivem Üben vermieden und der Test misst wirklich „Konzentriertes und sorgfältiges Arbeiten“.

Im Jahr 2005 wurde die Teststruktur in der Schweiz erstmals grundsätzlich revidiert. Aufgrund einer Anforderungsanalyse der neuen Studienbedingungen wurde ein neuer Untertest „Planen und Organisieren“ aufgenommen. Diese Anforderung wurde in erster Priorität umgesetzt. Damit die Testlänge vergleichbar bleibt, wird zukünftig in der Schweiz auf nicht gewertete Einstreuaufgaben verzichtet. Sie wurden bis 2004 vor allem beibehalten, um die Testanforderung auch hinsichtlich der Länge mit der deutschen TMS-Anforderung identisch zu halten. Da die Aktualität des EMS für die Studienanforderungen den wichtigeren Wert darstellt, die Testzeit sich aber nicht verlängern lässt, wurde diese Änderung vorgenommen. Alle Aufgaben werden jetzt auch gewertet.

Die Vorteile des EMS lassen sich folgendermassen zusammenfassen:

- Auswahl der Testanforderungen aus einer grösseren Menge möglicher Studienanforderungen aufgrund einer Anforderungsanalyse – ständige Anpassung an sich verändernde Bedingungen;
- Wissenschaftlicher Nachweis der Vorhersagbarkeit von Studienerfolg, was die gesetzlich geforderte Zulassung nach der Eignung zum Studium erlaubt;
- Konstruktion der Aufgaben durch Experten UND anschliessende empirische Überprüfung, damit die Aufgaben alle Gütekriterien erfüllen, die gewünschte Fähigkeit tatsächlich messen und optimal „schwierig“ sind;
- Beachtung, dass für die Beantwortung der Aufgaben kein spezielles fachliches Vorwissen notwendig ist, sondern tatsächlich die „Studierfähigkeit“ als aktuelle Fähigkeit zur Wissensaneignung und Problemlösung gemessen wird. Dadurch ist der Test auch wenig trainierbar, was sich positiv auf die Sozialverträglichkeit auswirkt (kein zusätzlicher Gewinn durch zusätzliche Trainingskurse nachgewiesen, wenn die empfohlene Vorbereitung erfolgt);
- Es ist ein ökonomisches und genau kapazitätsentsprechendes Zulassungsverfahren möglich, die Rechtsfähigkeit des Verfahrens hat sich bei mehreren Überprüfungen (auch gerichtlich) bestätigt.

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

Die Erprobung neuer Aufgaben für sechs der neun Untertests (siehe Tabelle 15) erfolgte in Deutschland im Rahmen sogenannter „**Einstreuaufgaben**“. Nur bei ausreichender Bewährung wurden solche Aufgaben in nachfolgenden Testversionen für die Werteberechnung verwendet.

Im Unterschied zu vielen „Übungsaufgaben“, die im sogenannten Trainingsmarkt im Umlauf sind, sind die echten EMS-Aufgaben empirisch geprüft, so dass sie bezüglich Lösungseindeutigkeit und Schwierigkeit optimal sind.

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 beim „Textverständnis“.

Beim Test „Konzentriertes und sorgfältiges Arbeiten“ müssen seit 2004 insgesamt 1600 Zeichen der Reihe nach bearbeitet werden – 400 davon sind anzustreichen. Es können in der zur Verfügung stehenden Zeit in der Regel nicht alle Zeichen bearbeitet werden. Die Position des **letzten angestrichenen** Zeichens bestimmt, wie viele Zeichen als bearbeitet gewertet werden. Alle übersehenen und fälschlich angestrichenen Zeichen **vor** diesem letzten bearbeiteten Zeichen zählen als Fehler und diese werden von der Menge der insgesamt angestrichenen Zeichen abgezogen. Die verbleibende Menge sind die „Richtigen“, die dann in eine Skala zwischen 0 und 20 transformiert werden, um mit den anderen Tests gleichgewichtig zum Punktwert addiert zu werden.

Alle Punkte der Untertests werden zu einer Summe addiert (**Punktwert**, vgl. Abbildung 19). Dieser Wert hat den Nachteil, dass er nicht zwischen Tests verschiedener Jahre vergleichbar ist. 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 schlechteren Testwert erreicht als die entsprechende Person.

Beim „Planen und Organisieren“ werden 2005 nur 19 Aufgaben gewertet, weil eine Aufgabe die Qualitätskriterien nicht erfüllt hat – mehr dazu siehe Kapitel 8.3.

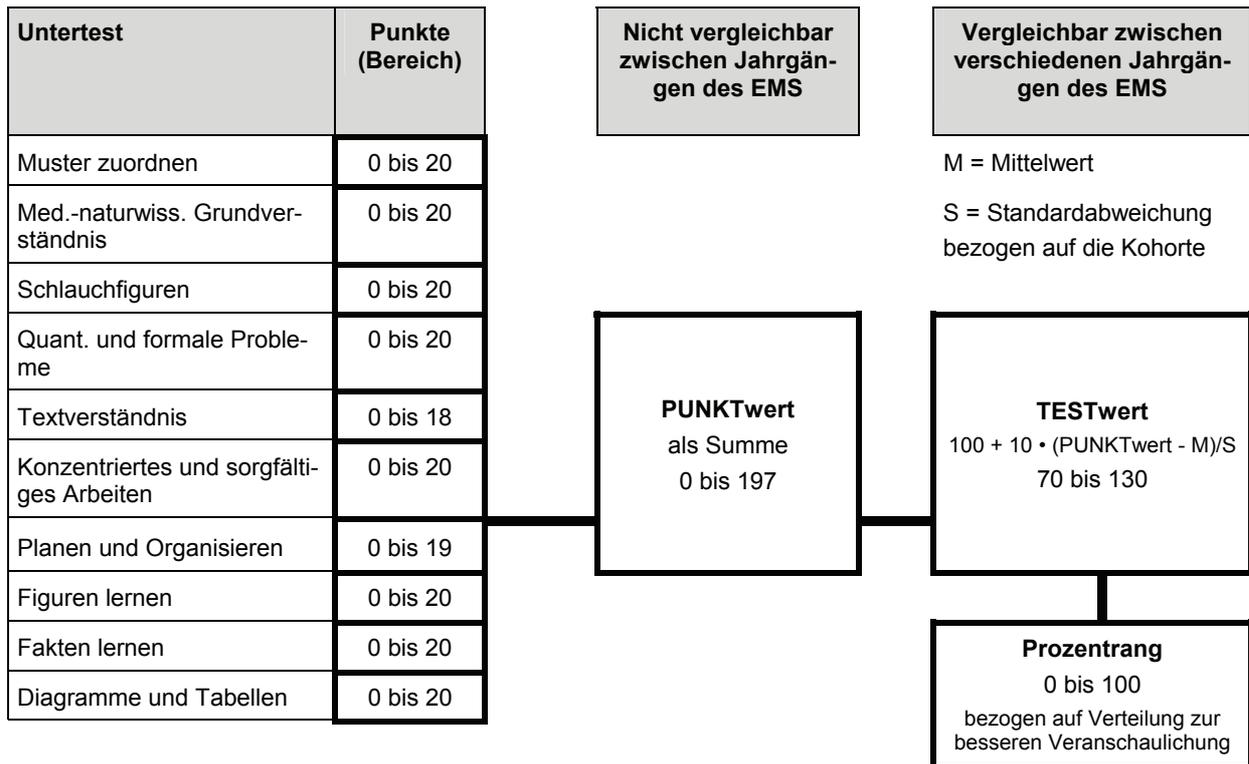


Abbildung 19: Punktwerte der einzelnen Untertests 2005 und ihre Zusammenführung über den Punktwert zum Testwert.

6.3 Mittlerer Rangplatz der Untertests

Ziel des EMS ist es, eine kapazitätsentsprechende Vergabe der Studienplätze zu erreichen. Es war vorgekommen, dass mehr Personen den gleichen Testwert aufwiesen als Plätze von der Kapazität her noch zur Verfügung standen.

Aus diesem Grunde hat der Beirat EMS beschlossen, ab 2005 weiter den Testwert als Zulassungskriterium zu verwenden. Bei **gleichem Testwert** werden die Personen in der Reihenfolge des mittleren Rangplatzes aller Untertests berücksichtigt und es werden nur noch so viele Personen zugelassen, wie freie Plätze tatsächlich vorhanden sind. Der mittlere Rangplatz wurde auf dem Testbescheid mitgeteilt. In der Praxis gibt es pro Untertest für jeden Punktwert einen Rangplatz-Wert. Der Rangplatz ist umso niedriger, je höher die Punktzahl ist und je weniger Personen insgesamt einen besseren Punktwert erreicht haben. Rangplätze gewichten besser, wie viele Personen besser oder schlechter als eine bestimmte Person sind.

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

Vom Gesetz gefordert ist ein Parameter, welcher die Studieneignung vorhersagt. Dies erfüllt der bisher verwendete Testwert nachweislich. Beim Testwert werden die Punktwerte der einzelnen Untertests addiert und der Gesamtwert wird durch die Standardisierung auf $M = 100$ und $S = 10$ zwischen den Jahren vergleichbar. Das Gewicht der Untertests in diesem Gesamtwert wird letztendlich von der Varianz in diesen Untertests beeinflusst. Da die für die

Studieneignung prognoserelevantesten Untertests zugleich die höchste Streuung aufweisen, erhöht dieser Effekt auch die Prognosekraft des Testwertes.

Zwecks Umrechenbarkeit der Rangplätze wird ein Rangplatz auf der Basis von 1000 Personen verwendet und über die Untertests gemittelt.

- Der Rangplatz bewegt sich 2005 zwischen 61.5 und 968. Niedrige Werte stehen für bessere Leistungen.
- Er korreliert mit dem Testwert mit -0.994 (2004 war es -0.992), was praktisch für eine Äquivalenz spricht – siehe auch Abbildung 20.
- Er differenziert allerdings besser zwischen den Personen und lässt eine kapazitätsgenaue Auswahl zu.

Das 2005 erstmals angewandte Prinzip, bei gleichem Testwert den Rangplatz zu verwenden, kann daher beibehalten werden.

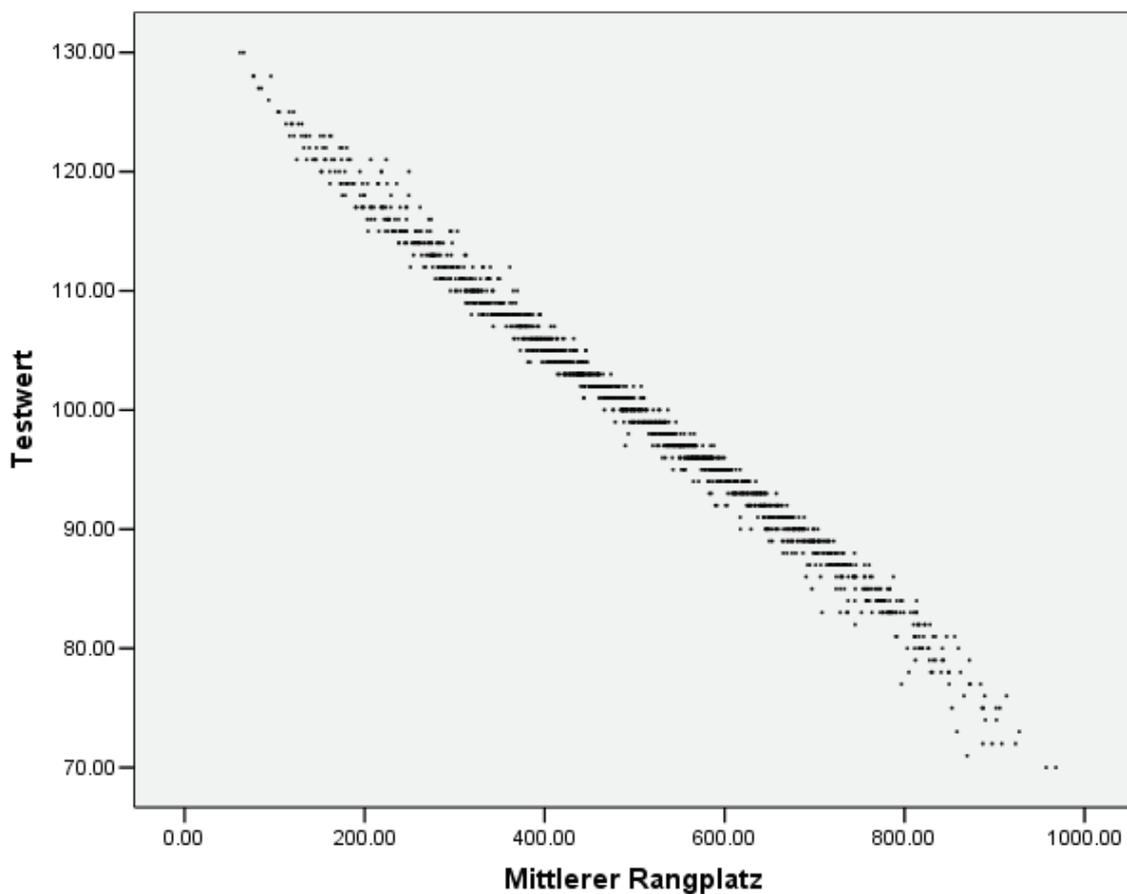


Abbildung 20: Beziehung zwischen Testwert und mittlerem Rangplatz.

7 Testanwendung in der Schweiz 2005

7.1 Verteilungsprüfung

Der Punktwert folgt wiederum einer Normalverteilung (Tabelle 16). Metrische Verfahren und die Standardisierung zum Testwert durch lineare z-Transformation sind daher zulässig. Die Punktwerte der Einzeltests sind auch dieses Jahr nicht normalverteilt (Abbildung 21 und Abbildung 22).

	Punktwert	Med.-naturw. Gr.	Quantitative und formale Probleme	Textver- ständnis	Diagramme und Tabellen	Planen und Organisieren	Fakten lernen	Figuren lernen	Muster zuord- nen	Schlauch- figuren	Konzentr. u. sorgf. Arbeiten
Mittelwert	112.69	11.14	10.25	8.98	10.26	9.88	12.52	13.36	12.72	12.15	11.43
Stand.- abweich.	21.55	3.68	4.05	3.70	3.62	3.20	3.77	3.83	3.20	3.61	4.35
extremste Differenz	0.024	0.064	0.070	0.083	0.069	0.072	0.074	0.073	0.067	0.073	0.067
Kolmog.- Smirnov Z	0.87	2.34	2.55	3.05	2.53	2.64	2.72	2.68	2.44	2.68	2.47
Asymp. Sig.	0.434	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

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

Die in den nachfolgenden Abbildungen dargestellten Verteilungen der Punktwerte machen deutlich, dass alle Punktskalen sehr gut differenzieren. Im Test „Konzentriertes und sorgfältiges Arbeiten“ ist erneut keine Extremhäufung hoher Punktzahlen vorhanden, was wir bis 2003 auf die hohe Trainingsabhängigkeit zurückgeführt haben. Die Strategie, Zeichen und Durchstreichregel vorher nicht bekannt zu geben, hat sich wie 2004 erneut bewährt. Bei „Figuren lernen“ und „Fakten lernen“ kommt es zu einer leichten Häufung der höchsten Punktzahlen, welche den Ergebnissen der Vorjahre entspricht. Hier wird langfristig ebenfalls an Versionen zu arbeiten sein, welche den Schwierigkeitsgrad erhöhen.

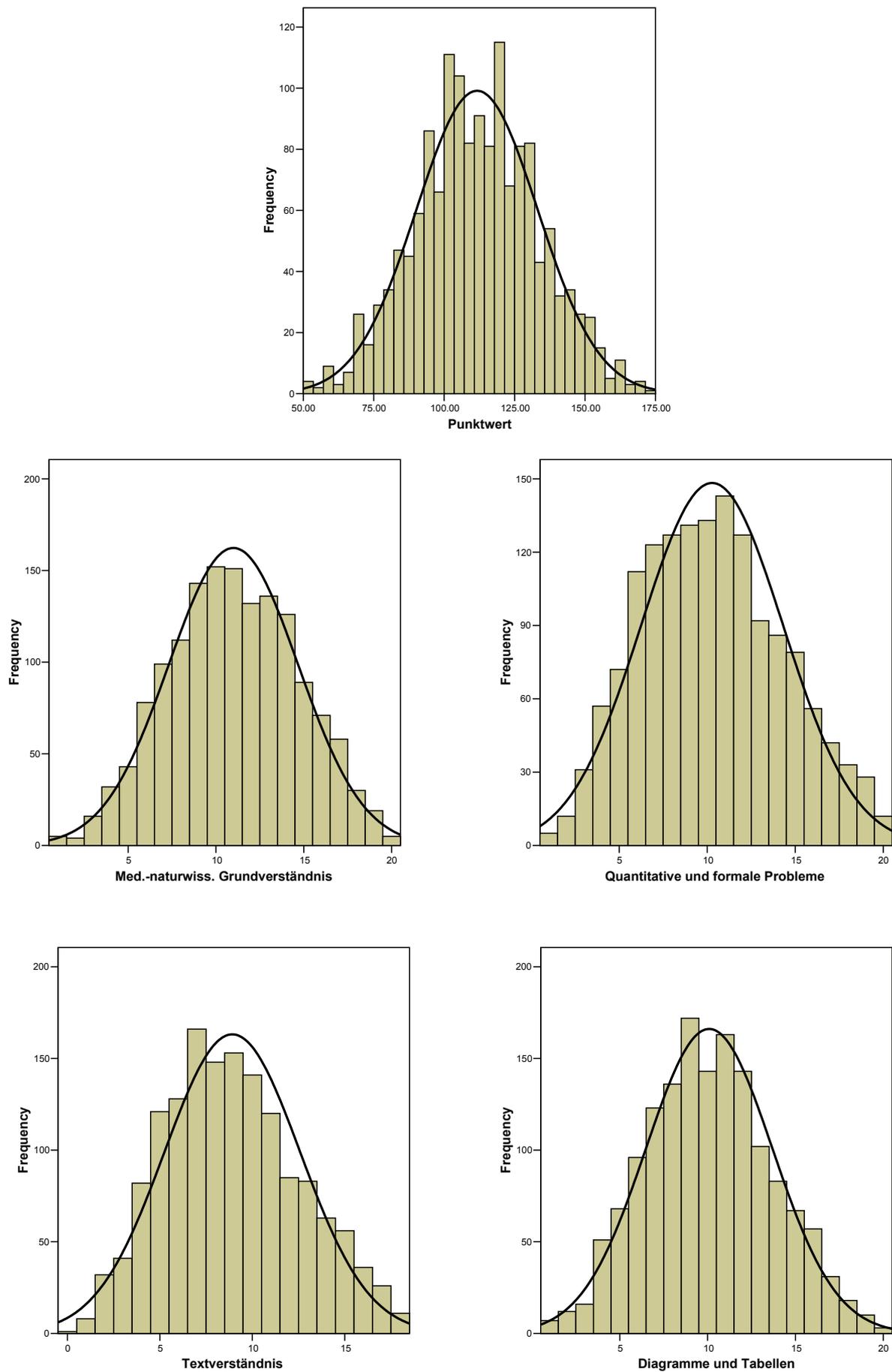


Abbildung 21: Häufigkeitsverteilungen für Punktwert und Punkte der Untertests (1).

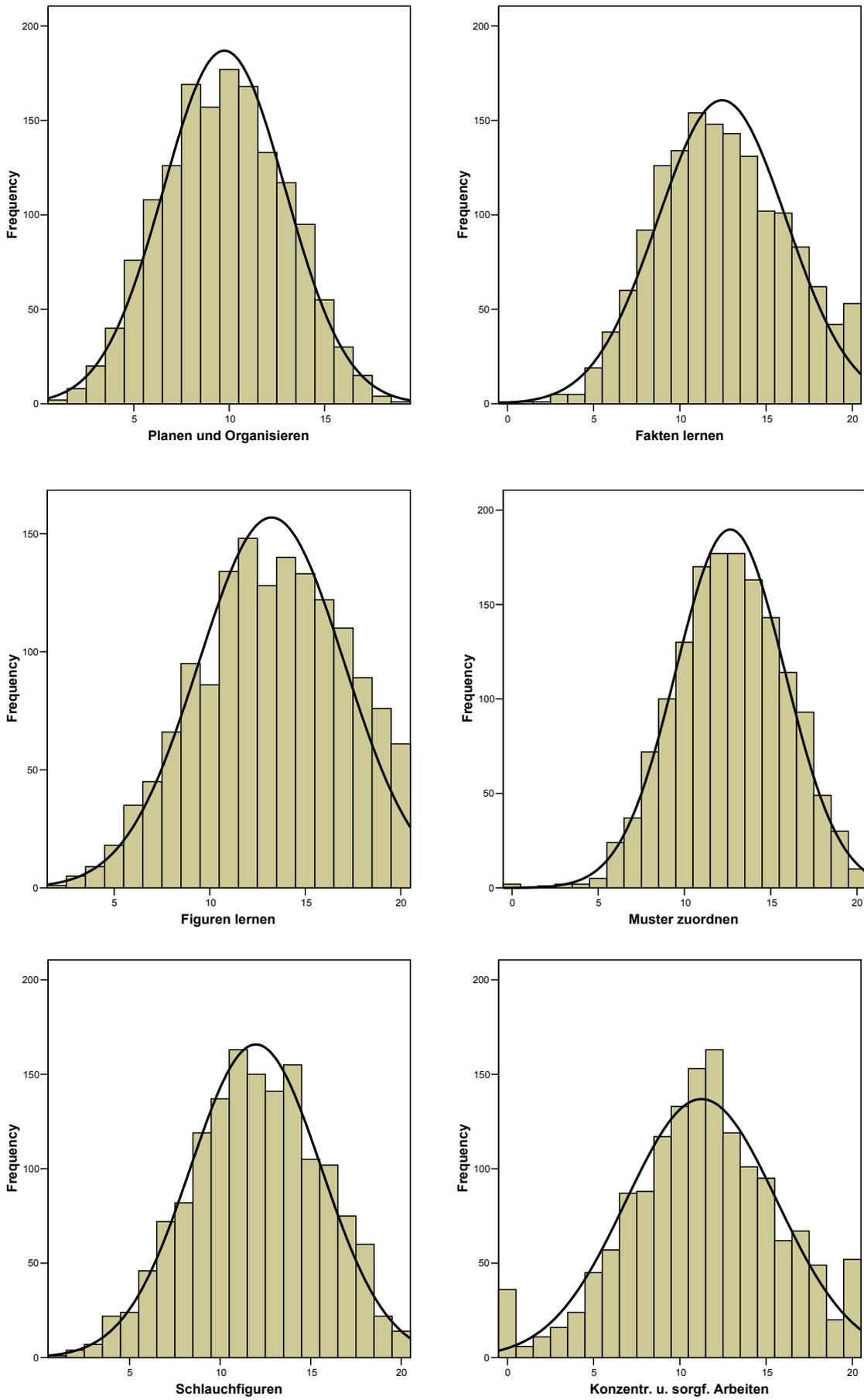


Abbildung 22: Häufigkeitsverteilungen für Punkte der Untertests (2).

7.2 Vergleich der Testfassungen 1998 bis 2005

Durch die Transformation der Punktwerte in eine Skala – unter Berücksichtigung des jeweiligen Mittelwertes und der Standardabweichungen pro Jahr – können die **Testwerte** aus verschiedenen Jahren miteinander verglichen werden. Dieses Vorgehen wird durch die zwischen den Jahren übereinstimmenden Gütekriterien zusätzlich legitimiert.

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

	2005		2004		2003		2002		2001		2000		1999		1998	
	D	CH	D	CH	D	CH	D	CH	D	CH	D	CH	D	CH	D	CH
Muster zuordnen	10.4	12.74	10.3	10.5	10.2	10.8	10.9	11.3	9.7	10.1	9.3	10.3	11	11.6	12.2	13.2
Med.-naturwiss. Grundverst.	11	11.32	10	10.4	10.2	10.5	9.5	9.9	10.7	10.1	10.8	10.7	11.2	11.3	11.6	12.5
Schlauchfiguren	10.9	12.33	11.8	12.1	12	12.7	11.8	12.1	12.2	13.3	12.3	13.2	11.9	12.9	11.8	13.3
Quant. und form. Probleme	9.82	10.57	9.4	9.9	9.6	10.5	9.9	11.1	9.8	10.9	10.2	11.8	9.7	11.6	9.9	11.9
Textverständnis	8.88	9.22	9	8.9	8.9	10.2	8.6	9.2	8.6	8.5	9.3	9	8.9	8.9	10.2	10.3
Figuren lernen	11.5	13.44	11.5	13	10.7	13.6	11.8	15.7	9.2	10.6	9.3	10.2	11.8	12.7	11.6	12.6
Fakten lernen	11.1	12.60	11.3	12.4	10.3	11.4	9.7	10.7	9.3	10.8	8.9	10.1	11	11.8	11.2	11.6
Diagramme und Tabellen	10.1	10.49	10.1	10.7	10.7	11.7	10.5	10.8	10	10.3	9.7	10.4	10.3	10.5	11	11.4
Total (ohne Konzent. u. sorgf. Arbeiten / Planen & Org.)	83.7	92.7	83.4	87.9	82.6	91.4	82.7	90.9	79.5	84.6	79.8	85.7	85.8	91.3	89.5	96.8

Tabelle 17: Mittelwerte (Punkte) der Untertests 1998 bis 2005; deutsche Sprachgruppe Schweiz und Schätzung aufgrund der Item-Schwierigkeitswerte aus Deutschland (HM: Humanmedizin) – für 8 Untertests liegen Werte vor.

Auch 2005 liegt das Gesamtergebnis aus der Schweiz (deutsche Sprachgruppe, Humanmedizin) über den Werten aus Deutschland. Mit neun Punkten Unterschied (ohne „Konzentriertes und sorgfältiges Arbeiten“ sowie „Planen und Organisieren“), ist diese Differenz trotz stetiger Zunahme der Teilnehmerzahlen etwas höher als in anderen Jahren ausgefallen. Eine mögliche Mitursache besteht im Wegfall der Einstreuaufgaben – es ist nicht mehr möglich, dass Testteilnehmer bei Aufgaben, die dann nicht gewertet werden, Zeit verlieren.

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

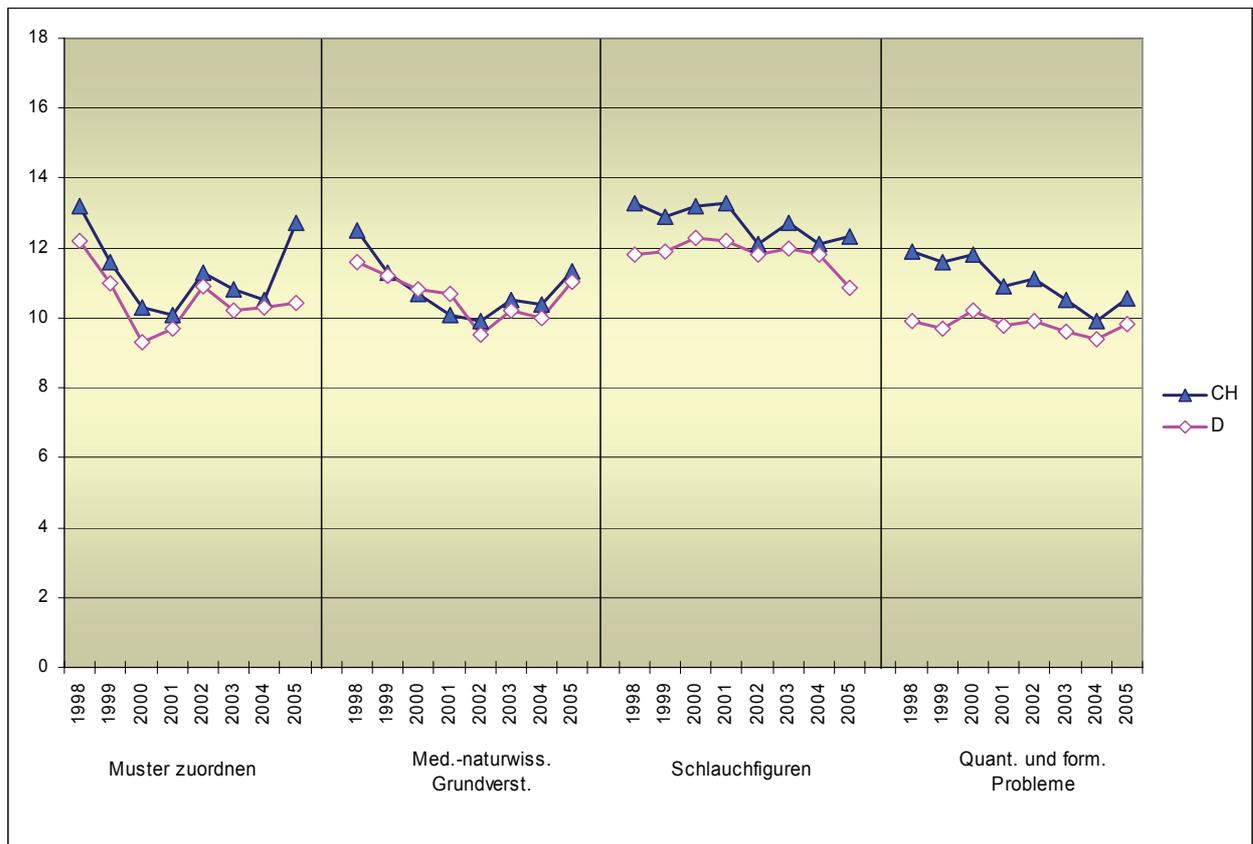


Abbildung 23: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2005 (Untertests 1-4).

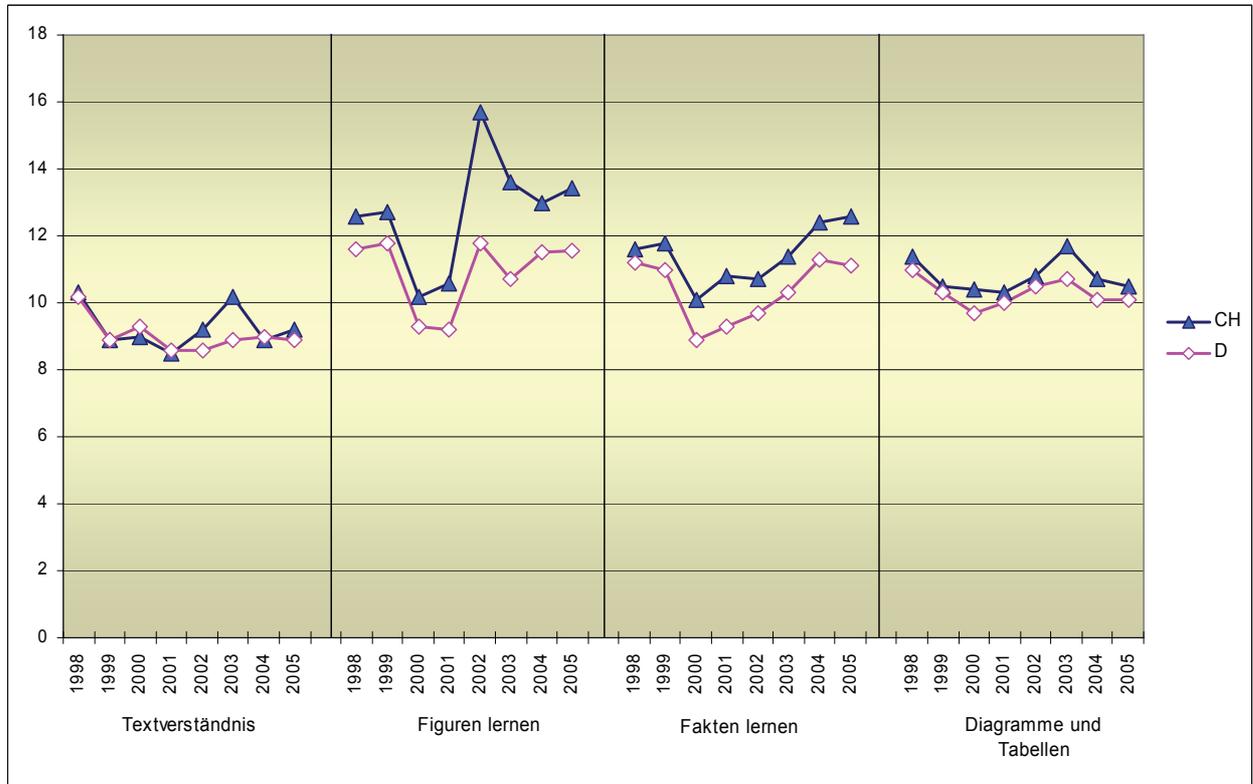


Abbildung 24: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig) und Vergleichswerte Deutschland 1998 bis 2005 (Untertests 5-8).

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 möglicherweise auf Unterschiede im Schulsystem zurückzuführen sind. Auch die Strategien, wie entsprechende Aufgaben gelöst werden, können sich unterscheiden und sich auf die Ergebnisse auswirken.
- Unterschiede in Leistungs-Voraussetzungen (z.B. Motivation, Belastbarkeit, Ausdauer) können zu unterschiedlichen Resultaten beitragen.

Unterschiedliche Repräsentativität der Stichproben für die Sprachgruppen

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

Testbedingte Unterschiede

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

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

1. Einen Ausgleich nur bei sprachabhängigen Tests vorzunehmen. Dies sind (neu) sechs von zehn Untertests. Der Sprachausgleich beschränkt sich demnach auf die eher sprachabhängigen Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“, „Diagramme und Tabellen“, sowie neu „Planen und Organisieren“.
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 nunmehr sechs **sprachabhängigen Untertests** („Textverständnis“, „Diagramme und Tabellen“, „Fakten lernen“, „Quantitative und formale Probleme“, „Medizinisch-naturwissenschaftliches Grundverständnis“, „Planen und Organisieren“) zu erwarten, während die vier weitgehend **sprachunabhängigen Untertests** („Muster zuordnen“, „Figuren lernen“, „Schlauchfiguren“, „Konzentriertes und sorgfältiges Arbeiten“) davon praktisch nicht betroffen sein können.

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 Wahrschein-

lichkeit 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.

7.3.1 Sprachvergleich für die Untertests

Wie in den Vorjahren treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Untertests auf. Auch dieses Jahr liegen die Punktzahlen der deutschsprachigen Teilnehmer mehrheitlich über jenen der französisch- und italienischsprachigen Kandidaten. In den Untertests „Quantitative und formale Probleme“, „Textverständnis“ und „Muster zuordnen“ erzielten die italienischsprachigen Teilnehmer die besten Ergebnisse.

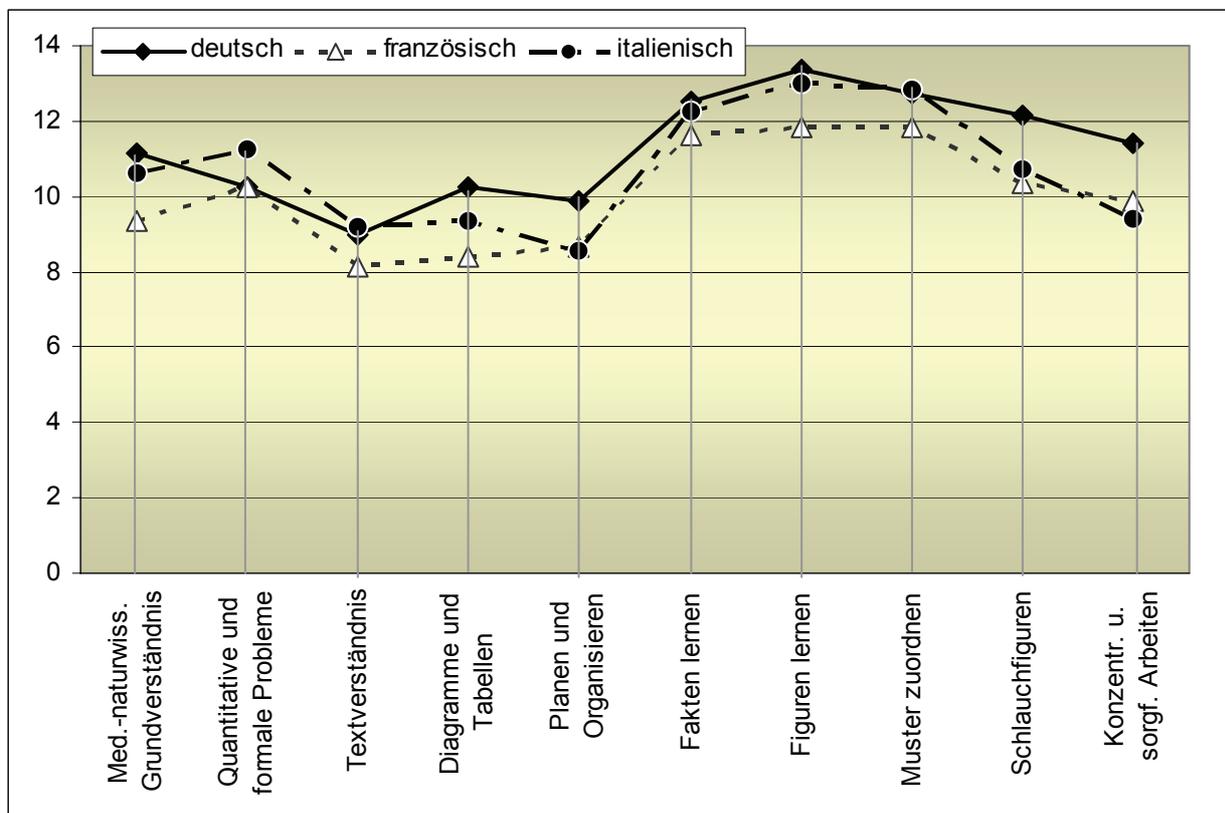


Abbildung 26: Mittelwerte der Untertests für die Sprachgruppen 2005 (unkorrigiert).

Die Abbildung 27 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern für alle Testanwendungen seit 1998. Die senkrechten Linien bezeichnen die Spannweite (Minimum bis Maximum) der Differenzen über sämtliche Jahrgänge, die Verbindungslinie steht für die Differenzen aus dem aktuellen Jahrgang. Größere Differenzwerte sprechen für bessere Ergebnisse der deutschen Sprachgruppe. Aus Abbildung 28 werden die höheren Punktzahlen der italienischsprachigen Teilnehmer in den Untertests „Muster zuordnen“, „Quantitative und formale Probleme“ und „Textverständnis“ ersichtlich.

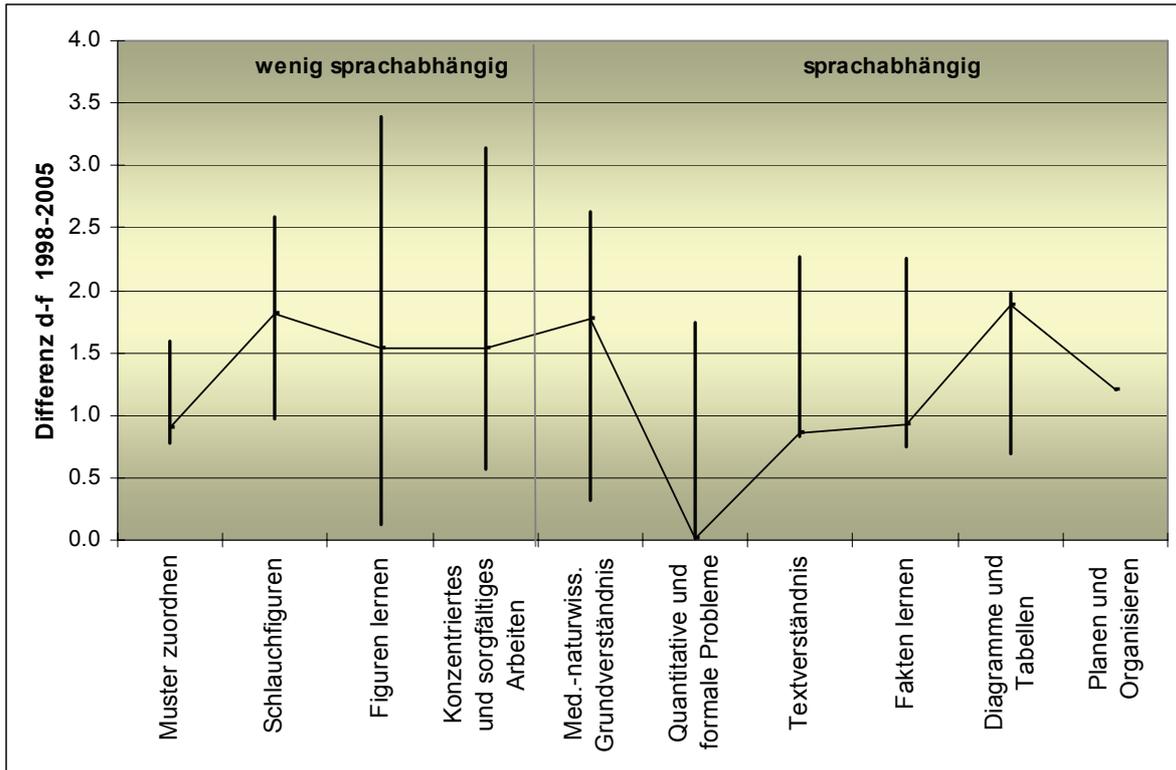


Abbildung 27: Differenz Testwert **deutsch-französisch** 2005 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2005.

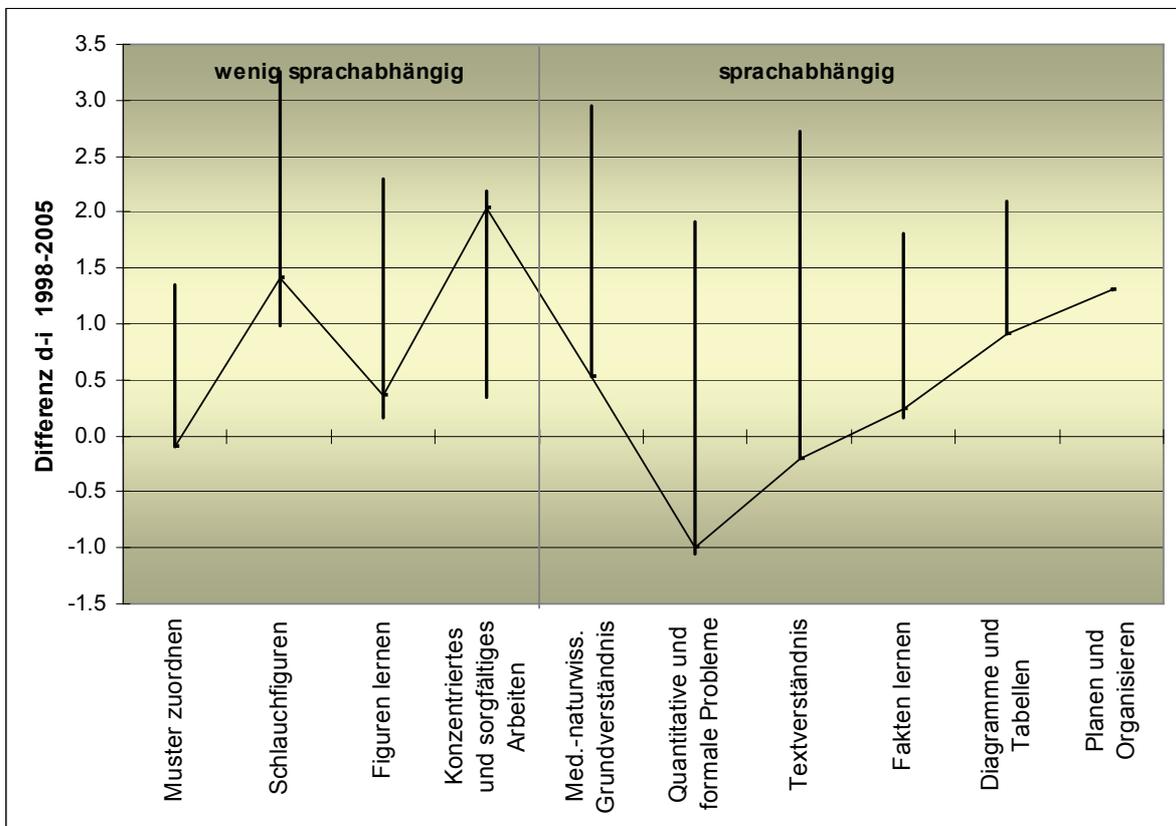


Abbildung 28: Differenz Testwert **deutsch-italienisch** 2005 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2005 (vertikale Linien).

7.3.2 Darstellung des Korrekturverfahrens

Die Identifikation und der Ausgleich testbedingter Unterschiede geschehen wie folgt:

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

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

7.3.2.1 Welche Untertests ausgleichen?

Die folgenden Tabellen zeigen die Sprachvergleiche für die eher sprachabhängigen Untertests. Referenz ist die deutschsprachige Testversion. Die französischsprachigen Teilnehmer (Tabelle 18) weisen in den Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Textverständnis“, „Fakten lernen“, „Diagramme und Tabellen“ sowie „Planen und Organisieren“ signifikant geringere Ergebnisse als die deutschsprachigen Teilnehmer auf.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	1343	11.14	3.68	0.000	Ja
	f	112	9.37	3.45		
Quantitative und formale Probleme	d	1343	10.25	4.05	0.968	Nein
	f	112	10.23	4.00		
Textverständnis	d	1343	8.98	3.70	0.007	Ja
	f	112	8.12	3.16		
Fakten lernen	d	1343	12.52	3.77	0.012	Ja
	f	112	11.60	3.52		
Diagramme und Tabellen	d	1343	10.26	3.62	0.000	Ja
	f	112	8.38	3.14		
Planen und Organisieren	d	1343	9.88	3.20	0.000	Ja
	f	112	8.67	2.95		

Tabelle 18: Signifikanzprüfung deutsche vs. französische Sprachform.

Die italienische Form (Tabelle 19) weist nur im Untertest „Planen und Organisieren“ einen signifikanten Unterschied auf – für die übrigen Untertests erfolgt kein Sprachausgleich.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	1343	11.14	3.68	0.332	Nein
	i	46	10.61	3.69		
Quantitative und formale Probleme	d	1343	10.25	4.05	0.102	Nein
	i	46	11.24	3.85		
Textverständnis	d	1343	8.98	3.70	0.722	Nein
	i	46	9.17	3.78		
Fakten lernen	d	1343	12.52	3.77	0.666	Nein
	i	46	12.28	2.74		
Diagramme und Tabellen	d	1343	10.26	3.62	0.908	Nein
	i	46	9.35	3.19		
Planen und Organisieren	d	1343	9.88	3.20	0.006	Ja
	i	46	8.57	3.30		

Tabelle 19: Signifikanzprüfung deutsche vs. italienische Sprachform.

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

7.3.2.2 Identifikation von DIF-Items und Bestimmung des Korrekturwertes

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

$$\Delta = 13 - 4z.$$

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

Die aus den Delta-Werten abgeleitete Regressionsgerade $Y = AX + B$ beschreibt die Beziehung zwischen den interessierenden Sprachversionen. Eine graphische Darstellung der Delta-

Werte (Delta-Plot) würde im Idealfall eine ellipsenförmige Anordnung von Punkten entlang der Diagonalen ergeben (Abbildung 29). Dies würde bedeuten, dass sowohl die Itemschwierigkeiten wie auch deren Reihenfolge in beiden Sprachversionen vergleichbar sind.

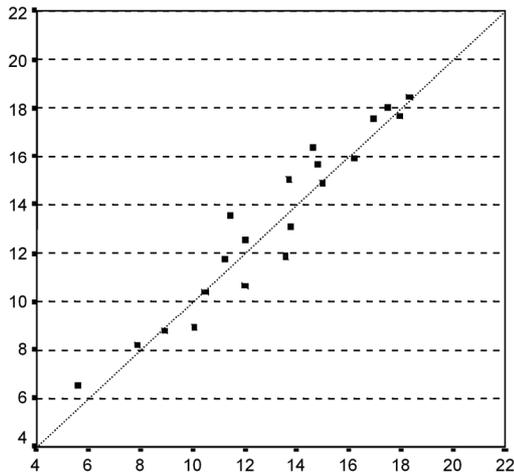


Abbildung 29: Beispiel für einen Delta-Plot.

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

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

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

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

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

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

7.3.2.3 Medizinisch-naturwissenschaftliches Grundverständnis

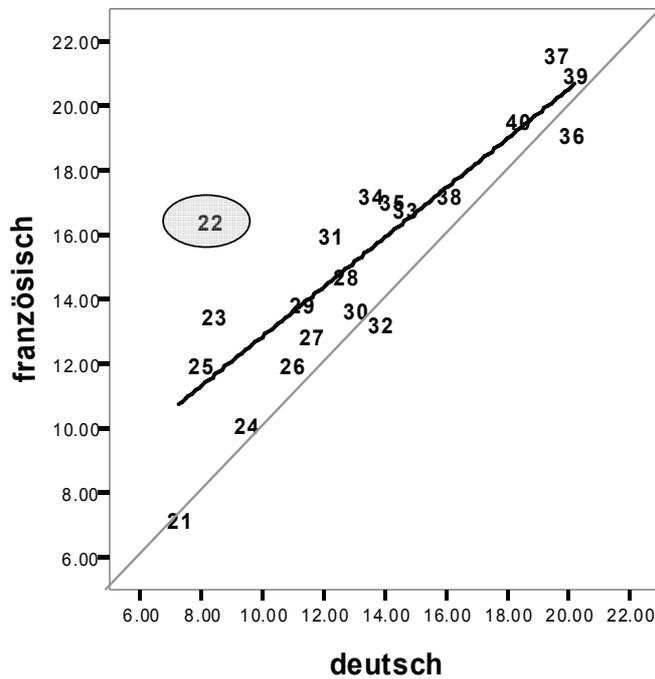


Abbildung 30: Delta-Plot für „Medizinisch-naturwissenschaftliches Grundverständnis“ (deutsch-französisch).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
21	0.81	0.81	0.74	7.26	7.17	8.89	2.82			
22	0.77	0.42	0.78	8.26	16.39	7.87	-3.89		0.35	
23	0.76	0.54	0.72	8.36	13.46	9.40	-1.50			
24	0.72	0.69	0.72	9.44	10.10	9.40	1.82			
25	0.78	0.61	0.65	7.96	11.99	10.93	-0.58			
26	0.65	0.61	0.59	10.95	11.99	12.46	1.24			
27	0.63	0.57	0.74	11.54	12.83	8.89	0.94			
28	0.58	0.49	0.52	12.70	14.71	13.99	0.14			
29	0.64	0.53	0.70	11.23	13.88	9.91	-0.09			
30	0.57	0.54	0.57	12.98	13.67	12.97	1.14			
31	0.60	0.44	0.59	12.19	15.97	12.46	-1.17			
32	0.53	0.55	0.67	13.81	13.25	10.42	1.99			
33	0.49	0.40	0.41	14.64	16.81	16.55	-0.34			
34	0.54	0.38	0.50	13.48	17.23	14.50	-1.38			
35	0.51	0.39	0.43	14.16	17.02	16.04	-0.80			
36	0.26	0.30	0.20	20.09	19.12	21.65	1.15			
37	0.28	0.20	0.22	19.58	21.63	21.14	-1.16			
38	0.43	0.38	0.41	16.09	17.23	16.55	0.21			
39	0.26	0.22	0.17	20.19	21.00	22.16	-0.29			
40	0.34	0.29	0.28	18.34	19.54	19.61	-0.25			

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

7.3.2.4 Textverständnis

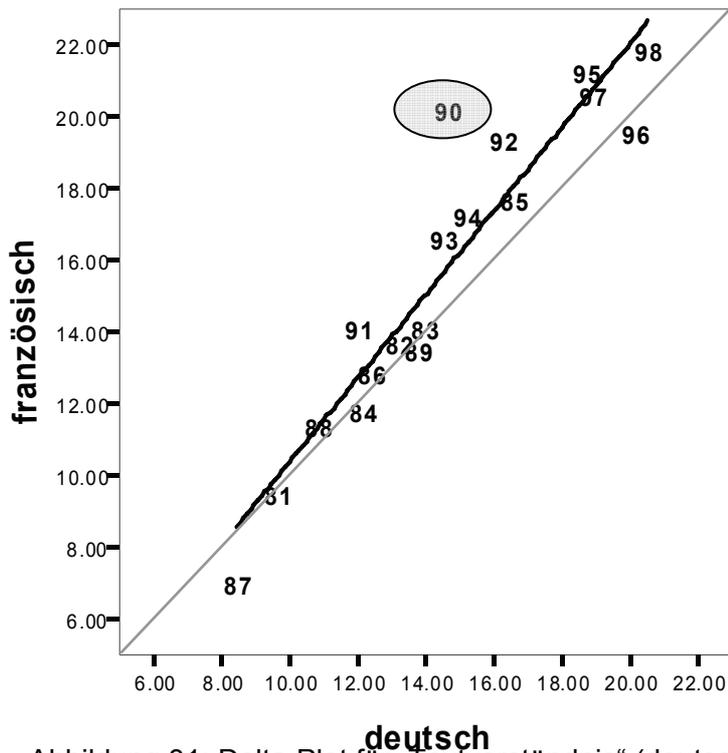
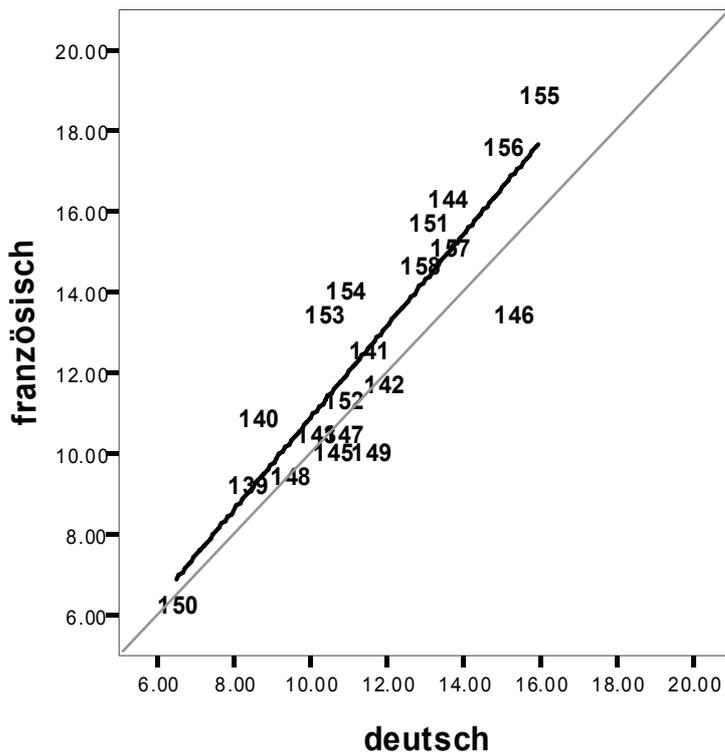


Abbildung 31: Delta-Plot für „Textverständnis“ (deutsch-französisch).

Item	Schwierigkeit			Δ -Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
81	0.71	0.71	0.63	9.64	9.47	11.44	0.31			
82	0.56	0.54	0.63	13.18	13.67	11.44	0.27			
83	0.52	0.52	0.59	13.95	14.09	12.46	0.58			
84	0.60	0.62	0.63	12.10	11.78	11.44	0.68			
85	0.41	0.37	0.39	16.56	17.65	17.06	0.24			
86	0.59	0.57	0.54	12.40	12.83	13.48	0.22			
87	0.76	0.82	0.80	8.45	6.96	7.36	1.05			
88	0.66	0.63	0.70	10.84	11.36	9.91	0.00			
89	0.53	0.54	0.65	13.73	13.46	10.93	0.82			
90	0.49	0.26	0.41	14.65	20.16	16.55	-2.85		0.23	
91	0.61	0.52	0.61	11.98	14.09	11.95	-0.92			
92	0.43	0.29	0.61	16.23	19.33	11.95	-1.11			
93	0.50	0.41	0.50	14.53	16.60	14.50	-0.62			
94	0.47	0.38	0.48	15.20	17.23	15.02	-0.52			
95	0.32	0.21	0.33	18.71	21.21	18.59	-0.45			
96	0.26	0.29	0.26	20.12	19.54	20.12	1.72			
97	0.31	0.24	0.17	18.85	20.58	22.16	0.06			
98	0.24	0.19	0.24	20.53	21.84	20.63	0.52			

Tabelle 21: DIF-Analyse „Textverständnis“.

7.3.2.5 Fakten lernen



Trotz signifikanter Unterschiede zwischen deutsch- und französischsprachigen Teilnehmern bezüglich der Ergebnisse im Unter-test „Fakten lernen“ treten auf Itemebene keine „unsystematischen“ Differenzen auf, welche auf Probleme bei der Übersetzung oder anderweitige Einschränkungen beim Verständnis der einzelnen Fragen hindeuten würden.

Abbildung 32: 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
139	0.76	0.72	0.63	8.34	9.26	11.44	-0.18			
140	0.75	0.65	0.74	8.61	10.94	8.89	-1.09			
141	0.63	0.58	0.61	11.52	12.62	11.95	-0.01			
142	0.61	0.62	0.54	11.91	11.78	13.48	0.84			
143	0.69	0.67	0.76	10.14	10.52	8.38	0.34			
144	0.54	0.42	0.48	13.55	16.39	15.02	-0.97			
145	0.67	0.69	0.67	10.58	10.10	10.42	0.95			
146	0.47	0.54	0.59	15.28	13.46	12.46	2.26			
147	0.66	0.67	0.50	10.84	10.52	14.50	0.87			
148	0.71	0.71	0.61	9.46	9.47	11.95	0.52			
149	0.63	0.69	0.52	11.56	10.10	13.99	1.68			
150	0.84	0.85	0.83	6.49	6.33	6.85	0.36			
151	0.56	0.45	0.48	13.08	15.76	15.02	-0.91			
152	0.66	0.63	0.59	10.84	11.36	12.46	0.31			
153	0.68	0.54	0.65	10.32	13.46	10.93	-1.47			
154	0.65	0.52	0.76	10.88	14.09	8.38	-1.46			
155	0.44	0.31	0.46	15.96	18.91	15.53	-0.82			
156	0.48	0.37	0.46	15.00	17.65	15.53	-0.71			
157	0.54	0.47	0.65	13.59	15.13	10.93	-0.12			
158	0.57	0.49	0.76	12.85	14.71	8.38	-0.39			

Tabelle 22: DIF-Analyse „Fakten lernen“.

7.3.2.6 Diagramme und Tabellen

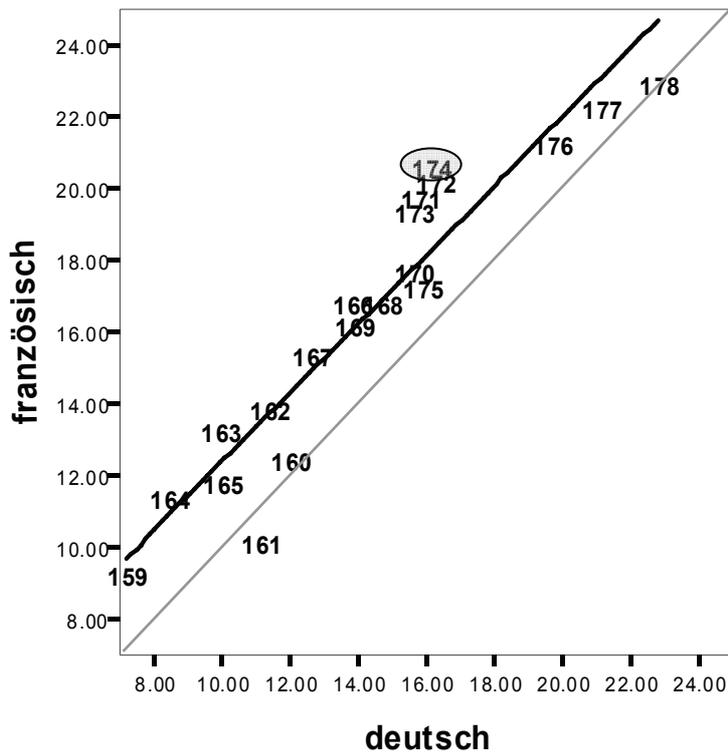


Abbildung 33: Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch).

Item	Schwierigkeit			Δ -Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
159	0.81	0.72	0.83	7.21	9.26	6.85	0.31			
160	0.61	0.59	0.63	12.01	12.41	11.44	1.37			
161	0.64	0.69	0.65	11.14	10.10	10.93	2.43			
162	0.63	0.53	0.72	11.35	13.88	9.40	-0.15			
163	0.69	0.55	0.67	9.95	13.25	10.42	-0.66			
164	0.76	0.63	0.67	8.41	11.36	10.42	-0.37			
165	0.69	0.62	0.74	10.02	11.78	8.89	0.44			
166	0.53	0.40	0.50	13.83	16.81	14.50	-0.55			
167	0.58	0.46	0.37	12.64	15.34	17.57	-0.31			
168	0.49	0.40	0.30	14.71	16.81	19.10	0.06			
169	0.53	0.43	0.59	13.88	16.18	12.46	-0.06			
170	0.45	0.37	0.33	15.61	17.65	18.59	0.08			
171	0.44	0.28	0.28	15.82	19.74	19.61	-1.28			
172	0.43	0.26	0.30	16.23	20.16	19.10	-1.31			
173	0.45	0.29	0.41	15.63	19.33	16.55	-1.11			
174	0.43	0.24	0.28	16.14	20.58	19.61	-1.67		0.19	
175	0.44	0.38	0.39	15.89	17.23	17.06	0.58			
176	0.28	0.21	0.22	19.69	21.21	21.14	0.33			
177	0.22	0.17	0.22	21.12	22.26	21.14	0.57			
178	0.15	0.14	0.24	22.80	22.89	20.63	1.28			

Tabelle 23: DIF-Analyse „Diagramme und Tabellen“.

7.3.2.7 Planen und Organisieren

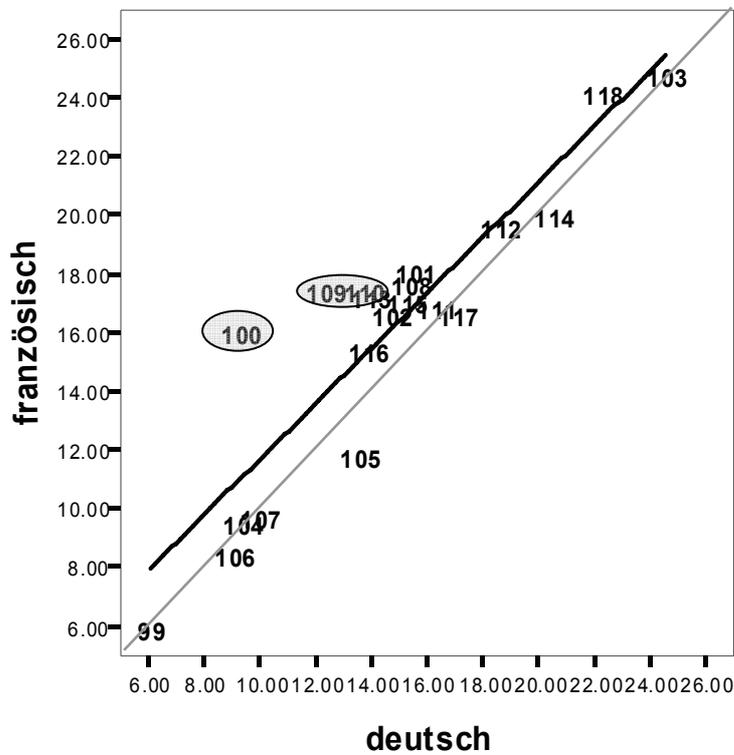


Abbildung 34: Delta-Plot für „Planen und Organisieren“ (deutsch-französisch).

Abbildung 34: Delta-Plot für „Pla-

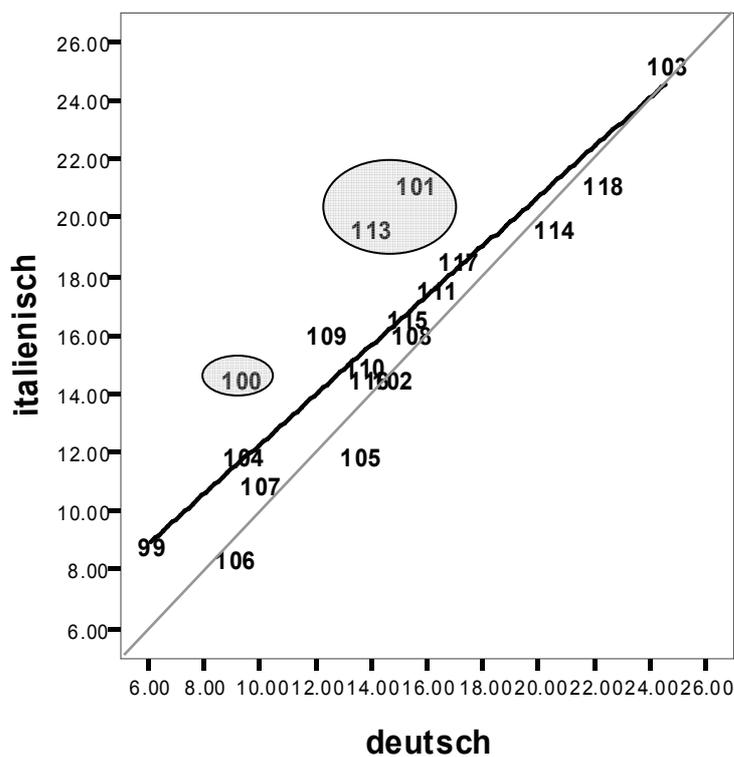


Abbildung 35: Delta-Plot für „Planen und Organisieren“ (deutsch-italienisch).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
99	0.86	0.87	0.74	6.05	5.91	8.89	1.49	0.05		
100	0.72	0.44	0.50	9.27	15.97	14.50	-3.62	-2.17	0.29	0.22
101	0.45	0.35	0.22	15.56	18.07	21.14	-0.82	-3.18		0.24
102	0.49	0.41	0.50	14.67	16.60	14.50	-0.37	1.32		
103	0.07	0.06	0.04	24.55	24.78	25.22	0.47	-0.50		
104	0.72	0.71	0.61	9.36	9.47	11.95	1.16	-0.16		
105	0.54	0.62	0.61	13.53	11.78	11.95	2.36	2.53		
106	0.73	0.76	0.76	9.08	8.43	8.38	1.73	2.39		
107	0.69	0.71	0.65	10.00	9.68	10.93	1.46	1.04		
108	0.46	0.37	0.43	15.40	17.65	16.04	-0.63	0.62		
109	0.59	0.38	0.43	12.36	17.44	16.04	-2.56	-1.34	0.22	
110	0.53	0.38	0.48	13.71	17.44	15.02	-1.64	0.31	0.16	
111	0.42	0.40	0.37	16.28	16.81	17.57	0.58	0.01		
112	0.33	0.29	0.24	18.60	19.54	20.63	0.20	-0.83		
113	0.52	0.38	0.28	13.94	17.23	19.61	-1.33	-3.06		0.24
114	0.24	0.27	0.28	20.49	19.95	19.61	1.19	1.17		
115	0.47	0.39	0.41	15.20	17.02	16.55	-0.31	0.09		
116	0.53	0.46	0.50	13.90	15.34	14.50	0.02	0.82		
117	0.39	0.41	0.33	17.05	16.60	18.59	1.26	-0.27		
118	0.17	0.09	0.22	22.29	24.15	21.14	-0.62	1.16		

Tabelle 24: DIF-Analyse „Planen und Organisieren“.

7.3.2.8 Effekte der Korrektur

Die durch den Ausgleich resultierenden Bonuspunkte können Tabelle 25 entnommen werden. Durchschnittlich werden den französischsprachigen Teilnehmern 0.92 Punkte, den italienischsprachigen Teilnehmern 0.40 Punkte gutgeschrieben. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der Bonus bei beiden Gruppen beträgt also 0 oder 1 Punkt.

		n	Minimum	Maximum	m	s
Franz.	Med. naturw. Grundverständnis	112	.00	.35	.2031	.174
	Textverständnis	112	.00	.23	.1704	.101
	Diagramme und Tabellen	112	.00	.19	.1442	.082
	Planen und Organisieren	112	.00	.67	.4006	.205
	GESAMT	112	.16	1.44	.9184	.345
Ital.	Planen und Organisieren	46	.00	.60	.3983	.166
	GESAMT	46	.00	.60	.3983	.166

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

Die Mittelwertsdifferenzen für jeden Untertest **nach** dem Sprachausgleich sind der Tabelle 26 zu entnehmen.

Die Ergebnisse multipler Mittelwertvergleiche können ebenfalls der ersten Spalte entnommen werden. Differenzen zwischen den Teilnehmern sind dort beispielsweise in der Form

D > (F, I) angegeben. Das Beispiel würde bedeuten, dass die deutschsprachigen Kandidaten signifikant höhere Werte erzielt haben als die beiden anderen Sprachgruppen.

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Testwert ** D > F	deutsch	1343	100.46	9.99	72	129
	französisch	112	95.08	9.49	76	117
	italienisch	46	98.07	8.11	81	114
Muster zuordnen * D > F	deutsch	1343	12.72	3.20	0	20
	französisch	112	11.82	2.88	5	19
	italienisch	46	12.83	2.21	9	18
Med.-naturwiss. Grundverständnis ** D > F	deutsch	1343	11.14	3.68	1	20
	französisch	112	9.57	3.36	1.35	18
	italienisch	46	10.61	3.69	1	19
Schlauchfiguren ** D > F	deutsch	1343	12.15	3.61	1	20
	französisch	112	10.34	3.30	2	18
	italienisch	46	10.74	3.15	4	18
Quantitative und for- male Probleme <i>homogen</i>	deutsch	1343	10.25	4.05	1	20
	französisch	112	10.23	4.00	3	19
	italienisch	46	11.24	3.85	6	20
Textverständnis <i>homogen</i>	deutsch	1343	8.98	3.70	0	18
	französisch	112	8.29	3.11	2.23	17.23
	italienisch	46	9.17	3.78	1	17
Figuren lernen ** D > F	deutsch	1343	13.36	3.83	2	20
	französisch	112	11.82	3.73	4	20
	italienisch	46	13.00	2.86	8	19
Fakten lernen * D > F	deutsch	1343	12.52	3.77	0	20
	französisch	112	11.60	3.52	5	20
	italienisch	46	12.28	2.74	7	18
Diagramme und Tabel- len ** D > F	deutsch	1343	10.26	3.62	1	20
	französisch	112	8.52	3.12	1.19	15
	italienisch	46	9.35	3.19	3	17
Konzentriertes und sorgf. Arbeiten ** D > (F,I)	deutsch	1343	11.43	4.35	0	20
	französisch	112	9.89	4.16	0	20
	italienisch	46	9.39	4.65	0	20
Planen und Organisieren ** D > F	deutsch	1343	9.88	3.20	1	19
	französisch	112	9.07	2.83	2.67	16
	italienisch	46	8.96	3.24	2.6	15

Tabelle 26: Testwert und Punktwerte der Untertests für die Sprachgruppen.

**/*: signifikanter Einfluss des Faktors „Sprache“ bei varianzanalytischer Prüfung (1%-, bzw. 5%-Niveau); daneben sind signifikante Unterschiede bei multiplem Mittelwertvergleich angegeben. In französischer und italienischer Sprachgruppe korrigierte Punktwerte.

In Abbildung 36 und Abbildung 37 sind die Verteilungen der (ungerundeten) Bonuspunkte für die bisherigen Testdurchführungen dargestellt (1998 wurde kein Sprachausgleich vorgenommen). Die senkrechten Linien bezeichnen die Spannweite in jedem Jahr (Minimum bis Maximum), der jeweilige Mittelwert wird durch die waagerechte Markierung gekennzeichnet.

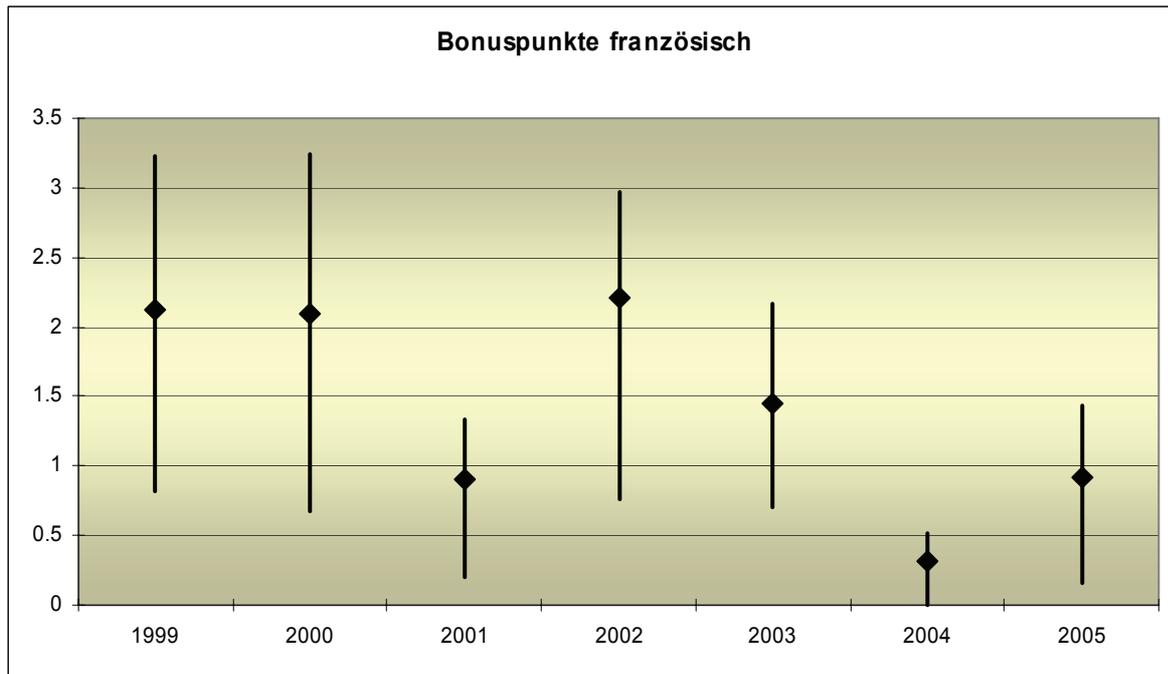


Abbildung 36: Auswirkungen der Korrektur über die Jahre für die französischsprachigen Kandidaten (Minimum – Mittelwert – Maximum).

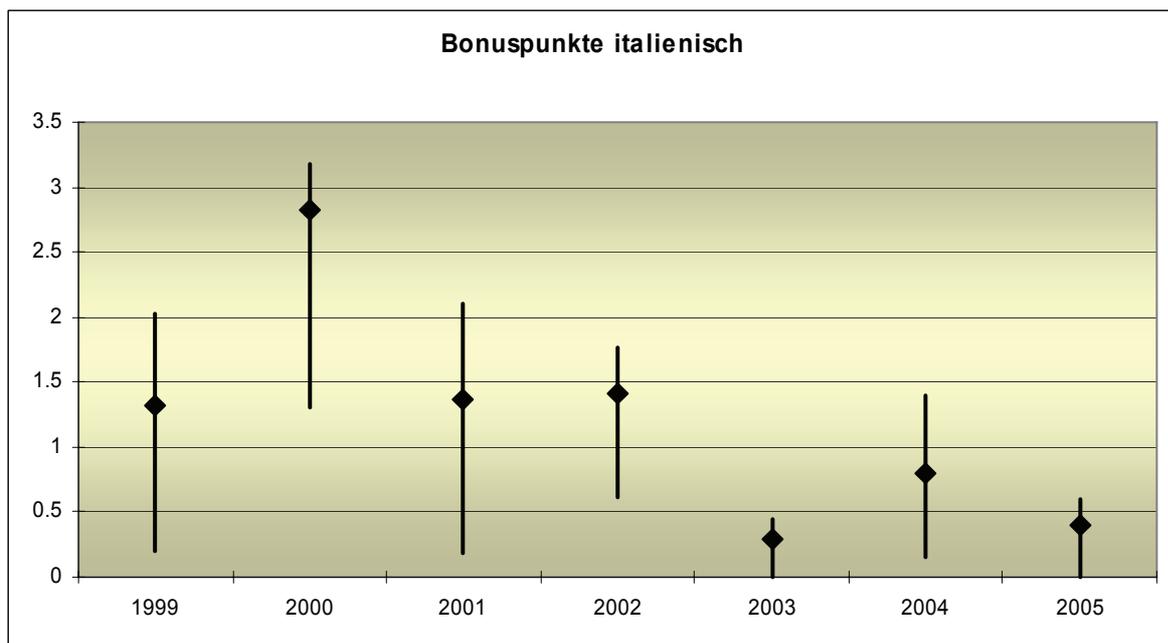


Abbildung 37: Auswirkungen der Korrektur über die Jahre für die italienischsprachigen Kandidaten (Minimum – Mittelwert – Maximum).

7.3.2.9 Evaluation: Gleichbehandlung der Sprachgruppen

Die Zulassungsquoten nach Sprachgruppe sind in der Tabelle 27 und der Tabelle 28 dargestellt. In der Humanmedizin unterscheiden sich die Zulassungsquoten nicht signifikant, in der Veterinärmedizin ist die Zulassungsquote der deutschsprachigen Teilnehmer höher als die der übrigen Kandidaten.

Disziplin	Sprache	nicht zugelassen	zugelassen	Total
Humanmedizin	deutsch	488 47.7%	536 52.3%	1024 100%
	französisch	27 54.0%	23 46.0%	50 100%
	italienisch	17 51.5%	16 48.5%	33 100%
	Total	532 48.1%	575 51.9%	1107 100%
Veterinärmedizin	deutsch	43 24.7%	141 75.3%	174 100%
	französisch	35 58.3%	25 41.7%	60 100%
	italienisch	6 60.0%	4 40.0%	10 100%
	Total	84 34.4%	160 65.6%	244 100%

Tabelle 27: Zulassung nach Sprache des Tests.

Disziplin		Wert	df	Asymp. Sig. (2seitig)
Humanmedizin	Pearson Chiquadrat	.931	2	.628
	Likelihood Ratio	.931	2	.628
Veterinärmedizin	Pearson Chiquadrat	25.361	2	.000
	Likelihood Ratio	24.635	2	.000

Tabelle 28: Prüfung auf Verteilungsunterschiede für Sprachgruppen getrennt nach Disziplinen: Kein signifikanter Unterschied für Humanmedizin, signifikanter Unterschied für Veterinärmedizin.

In den folgenden Abbildungen sind die Zulassungsquoten für Humanmedizin und Veterinärmedizin dargestellt. Hohe Werte bedeuten eine geringere Zulassungsquote. Für die Zahnmedizin konnte jedem Teilnehmer mit gültigem Testergebnis ein Studienplatz zugewiesen werden.

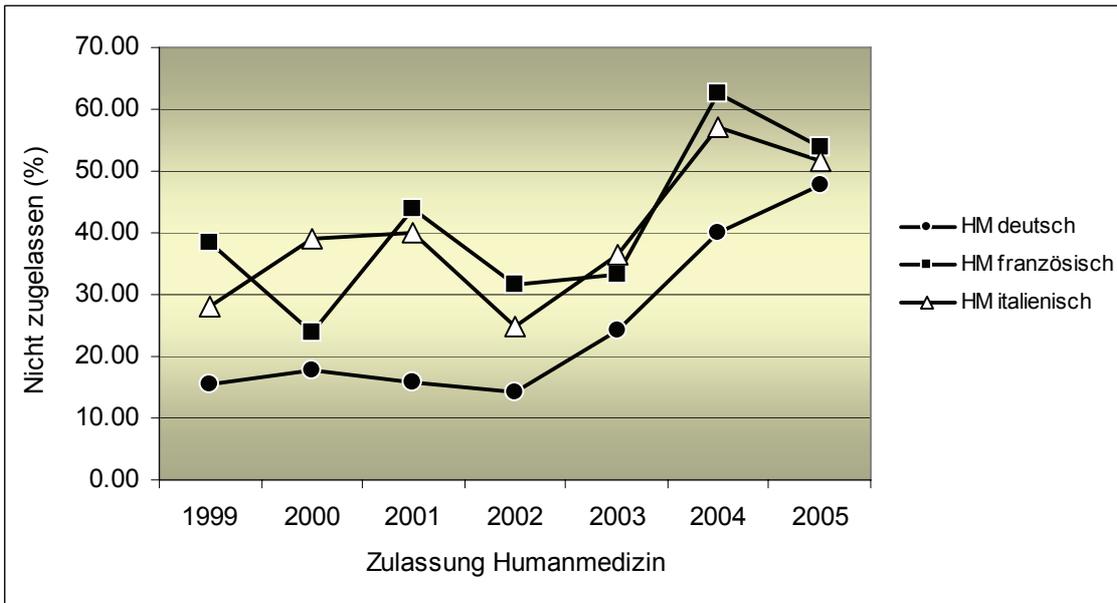


Abbildung 38: Zulassungsquoten Humanmedizin nach Sprache 1999 bis 2005.

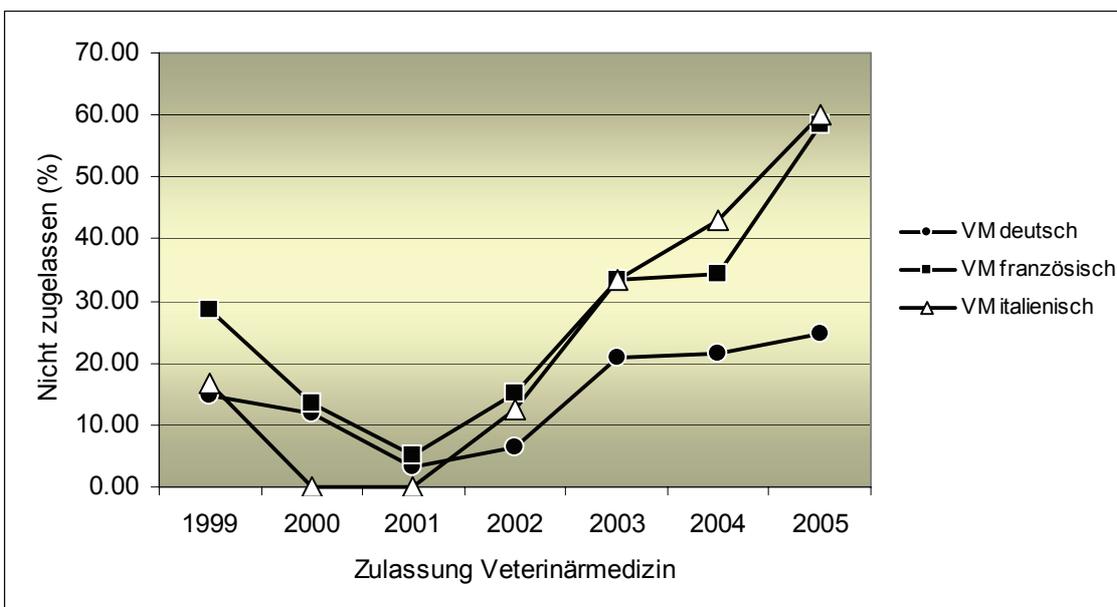


Abbildung 39: Zulassungsquoten Veterinärmedizin nach Sprache 1999 bis 2005.

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

Die Differenzen zwischen den französischsprachigen Teilnehmern und der deutschsprachigen Referenzgruppe sowie jene zwischen den italienischsprachigen Teilnehmern und der Referenzgruppe werden nach folgender Formel berechnet:

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

m ist der Mittelwert und s die Standardabweichung der entsprechenden Sprachgruppe. Die resultierenden Werte „Delta z “ drücken die Abweichung der jeweiligen Sprachgruppe von der

deutschen Sprachgruppe in Standardabweichungen aus. 0.5 hiesse, dass die betroffene Sprachgruppe um eine halbe Standardabweichung schlechter ist als die deutschsprachige Referenzgruppe.

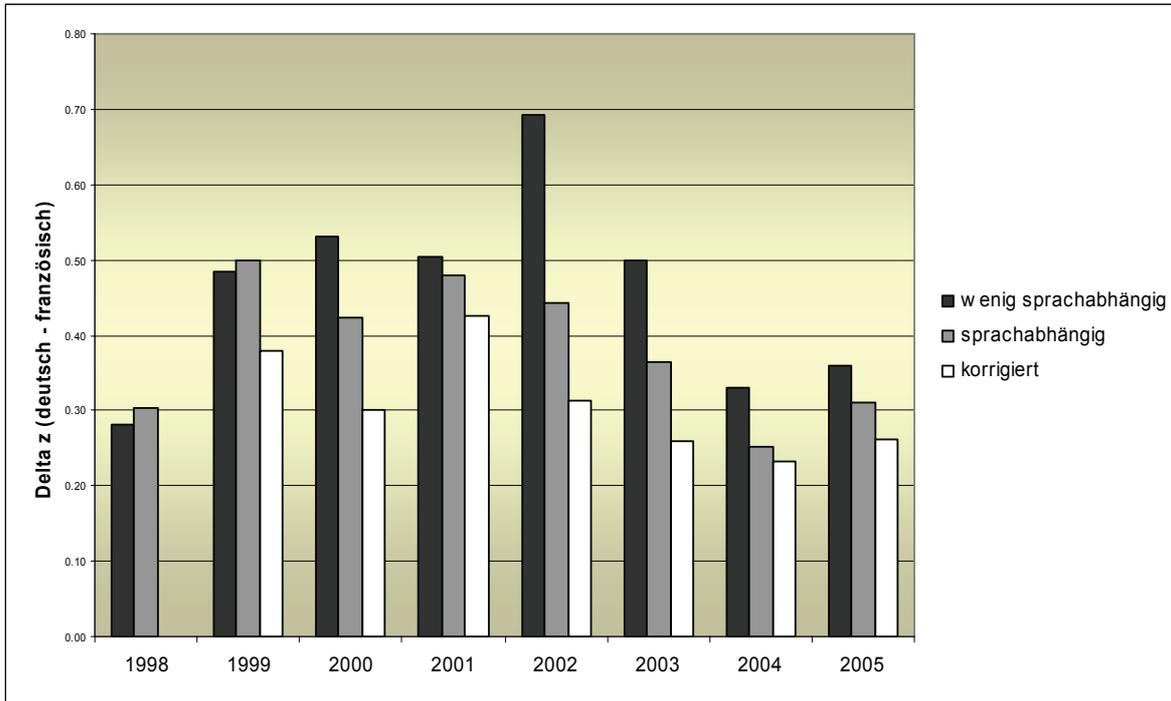


Abbildung 40: Abweichung (Delta z) für den Testwert zwischen deutscher und französischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

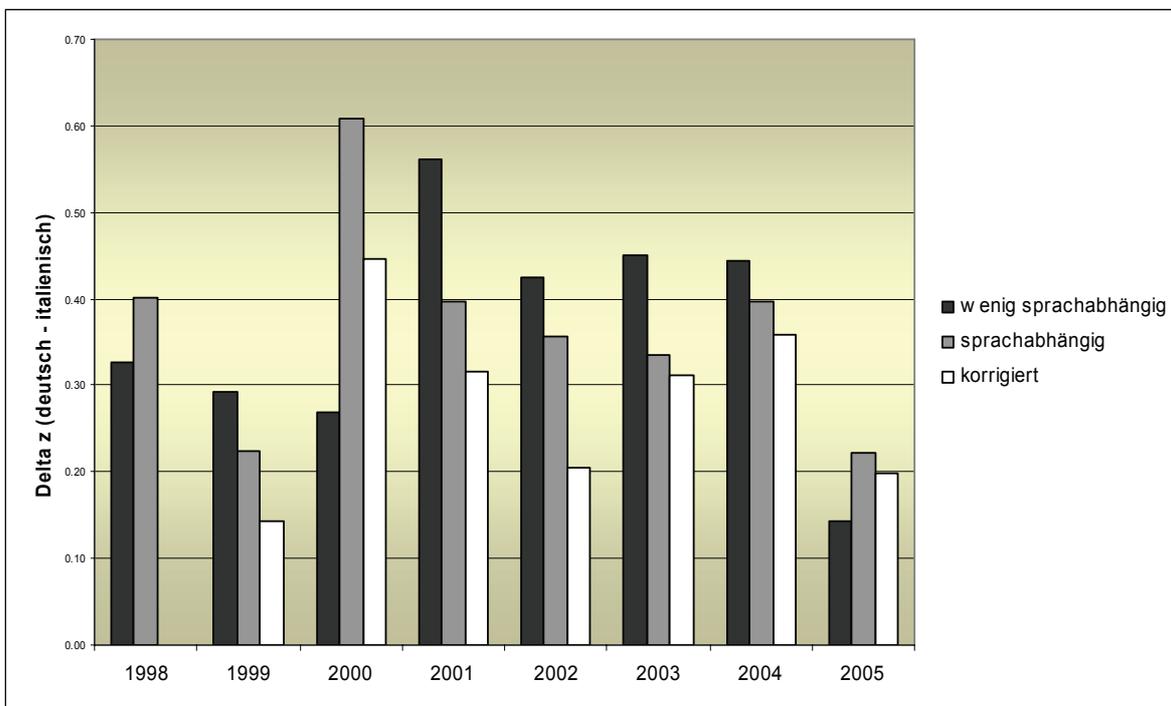


Abbildung 41: Abweichung (Delta z) für den Testwert zwischen deutscher und italienischer Sprachgruppe. Für die sprachabhängigen Untertests sind korrigierte und unkorrigierte Werte dargestellt.

Erneut sind in der französischen Gruppe (Abbildung 40) die Unterschiede in den wenig sprachabhängigen Untertests grösser als jene der sprachabhängigen Untertests. Die Korrektur verringert diese noch, so dass eine Benachteiligung ausgeschlossen werden kann. Bei den italienischsprachigen Teilnehmern fallen die Unterschiede in den wenig sprachabhängigen Untertests geringer aus, dort werden durch den Sprachausgleich nach 2000 (1998 fand kein Ausgleich statt) zum zweiten Mal möglicherweise vorhandene Unterschiede verringert.

7.4 Vergleichbarkeit der Testlokale

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

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testwert	<i>Zwischen Gruppen</i>	1446.158	15	96.411	0.966	0.489
	<i>Innerhalb Gruppen</i>	132421.693	1327	99.790		
	<i>Total</i>	133867.851	1342			
Muster zuordnen	<i>Zwischen Gruppen</i>	141.051	15	9.403	0.919	0.542
	<i>Innerhalb Gruppen</i>	13574.797	1327	10.230		
	<i>Total</i>	13715.848	1342			
Med.-naturwiss. Grundverständnis	<i>Zwischen Gruppen</i>	191.884	15	12.792	0.945	0.512
	<i>Innerhalb Gruppen</i>	17957.380	1327	13.532		
	<i>Total</i>	18149.264	1342			
Schlauchfiguren	<i>Zwischen Gruppen</i>	214.453	15	14.297	1.098	0.353
	<i>Innerhalb Gruppen</i>	17285.863	1327	13.026		
	<i>Total</i>	17500.316	1342			
Quant. und forma- le Probleme	<i>Zwischen Gruppen</i>	355.371	15	23.691	1.455	0.114
	<i>Innerhalb Gruppen</i>	21605.061	1327	16.281		
	<i>Total</i>	21960.432	1342			
Textverständnis	<i>Zwischen Gruppen</i>	88.915	15	5.928	0.430	0.971
	<i>Innerhalb Gruppen</i>	18282.322	1327	13.777		
	<i>Total</i>	18371.238	1342			
Figuren lernen	<i>Zwischen Gruppen</i>	169.364	15	11.291	0.767	0.715
	<i>Innerhalb Gruppen</i>	19531.039	1327	14.718		
	<i>Total</i>	19700.404	1342			
Fakten lernen	<i>Zwischen Gruppen</i>	256.326	15	17.088	1.207	0.259
	<i>Innerhalb Gruppen</i>	18780.588	1327	14.153		
	<i>Total</i>	19036.914	1342			
Diagramme und Tabellen	<i>Zwischen Gruppen</i>	257.037	15	17.136	1.315	0.185
	<i>Innerhalb Gruppen</i>	17288.362	1327	13.028		
	<i>Total</i>	17545.398	1342			
Konzentr. u. sorgf. Arbeiten	<i>Zwischen Gruppen</i>	284.821	15	18.988	1.003	0.449
	<i>Innerhalb Gruppen</i>	25120.695	1327	18.930		
	<i>Total</i>	25405.516	1342			
Planen und Orga- nisieren	<i>Zwischen Gruppen</i>	242.221	15	16.148	1.590	0.069
	<i>Innerhalb Gruppen</i>	13475.261	1327	10.155		
	<i>Total</i>	13717.482	1342			

Tabelle 29: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den (deutschsprachigen) Testlokalen.

7.5 Vergleich der Geschlechter

Die Forderung nach Gleichbehandlung der Geschlechter konnte in den bisherigen Testdurchführungen jeweils als erfüllt betrachtet werden. Chancengleichheit bedeutet, dass bei gleicher Eignung die gleichen Chancen auf eine Zulassung bestehen müssen. Beide Teilgruppen sind unterschiedlich repräsentativ in der Bewerberkohorte vertreten. Der Männeranteil in der Humanmedizin sank zwischen 1999 bis 2002 immer weiter ab (Abbildung 42). 2003 war erstmals wieder ein geringer Anstieg des Anteils an Kandidaten zu verzeichnen, der aber 2004 weiter abfiel. 2005 wird der Anteil auf diesem Niveau stabilisiert.

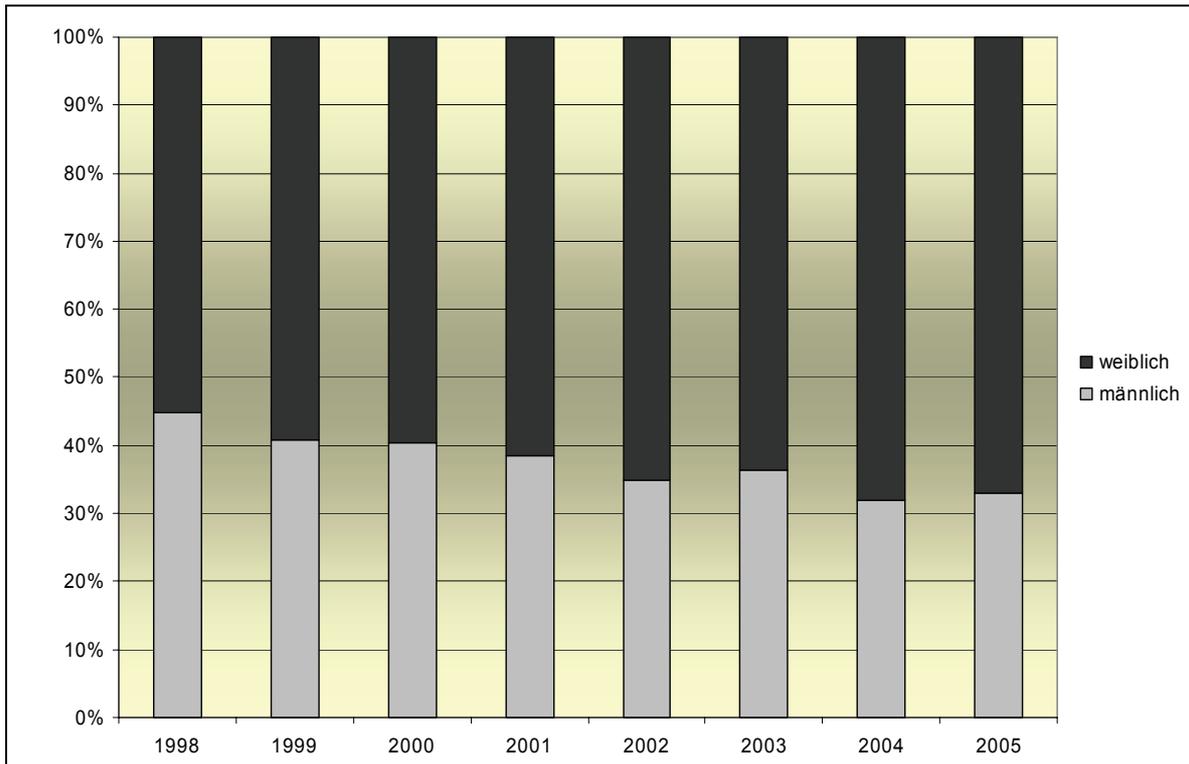


Abbildung 42: Bewerberkohorten (Humanmedizin) 1998 bis 2005 nach Geschlecht.

Die mittleren Testwerte (in Klammer die Standardabweichung) nach Geschlecht, Sprache und Disziplin können Tabelle 30 entnommen werden.

		Sprache			Disziplin			Gesamt
		deutsch	französisch	italienisch	HM	VM	ZM	
Geschlecht	männlich	101.52 (10.18)	95.87 (10.85)	103.00 (8.33)	102.16 (10.15)	94.54 (10.10)	97.72 (9.04)	101.21 (10.23)
	weiblich	99.93 (9.85)	94.78 (9.09)	95.43 (6.76)	100.24 (9.94)	98.14 (9.57)	95.70 (8.43)	99.38 (9.93)

Tabelle 30: Statistiken für den Testwert (in Klammern die Standardabweichung) nach Geschlechtern.

		Humanmedizin			Veterinärmedizin			Zahnmedizin		
		Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch
männlich	m	102.34	97.35	104.54	95.92	93.82	80.68	97.94	84.39	
	s	10.18	10.67	6.01	9.94	10.29	.	8.99	.	
	n	372	21	15	16	9	1	60	1	0
weiblich	m	100.39	99.40	96.91	100.45	92.12	92.17	95.71	97.37	95.98
	s	10.04	9.55	6.35	9.28	7.82	6.74	8.50	.	8.38
	n	652	29	18	158	51	9	85	1	3

Tabelle 31: Kombinierte Statistiken für den Testwert nach Geschlechtern, Sprache und Disziplin.

Sprache	Disziplin	Levene's Test Gleichheit der Varianzen			t-Test Gleichheit des Mittelwerts		
			F	Sig.	t	df	Sig.
deutsch	Humanmedizin	Varianzen homogen	.155	.694	2.975	1022	.003
	Veterinärmedizin	Varianzen homogen	.005	.945	-1.847	172	.066
	Zahnmedizin	Varianzen homogen	.003	.953	1.522	143	.130
französisch	Humanmedizin	Varianzen homogen	.357	.553	-.714	48	.479
	Veterinärmedizin	Varianzen homogen	.398	.530	.573	58	.569
italienisch	Humanmedizin	Varianzen homogen	.003	.957	3.523	31	.001

Tabelle 32: Prüfung Varianzhomogenität und Mittelwert (t-Test) für Testwert zwischen Geschlechtern. Französische Sprachgruppe Zahnmedizin und italienische Sprachgruppe Veterinär- und Zahnmedizin wg. Fallzahl nicht berechnet.

Abbildung 43 zeigt den Vergleich zwischen männlichen Bewerbern und Kandidatinnen, aufgeteilt nach Disziplin, Abbildung 44 den Boxplot für die Testwerte nach Sprache und Geschlecht geschichtet.

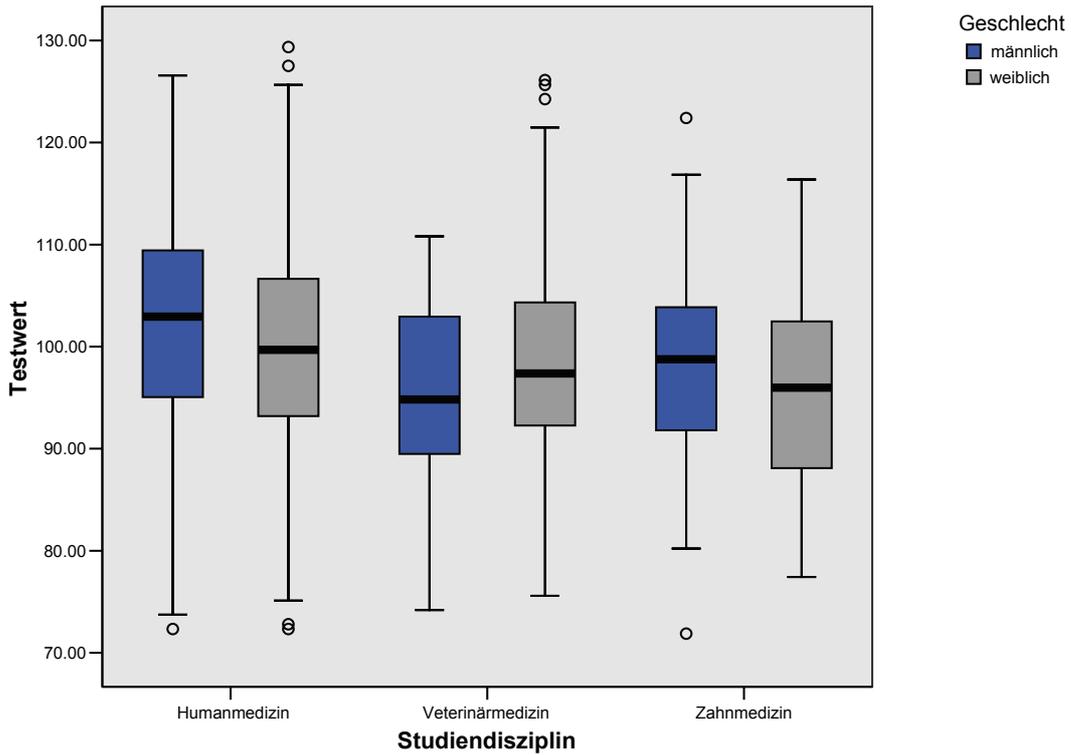


Abbildung 43: Boxplot für den Testwert nach Geschlecht und Disziplin (Humanmedizin und Veterinärmedizin).

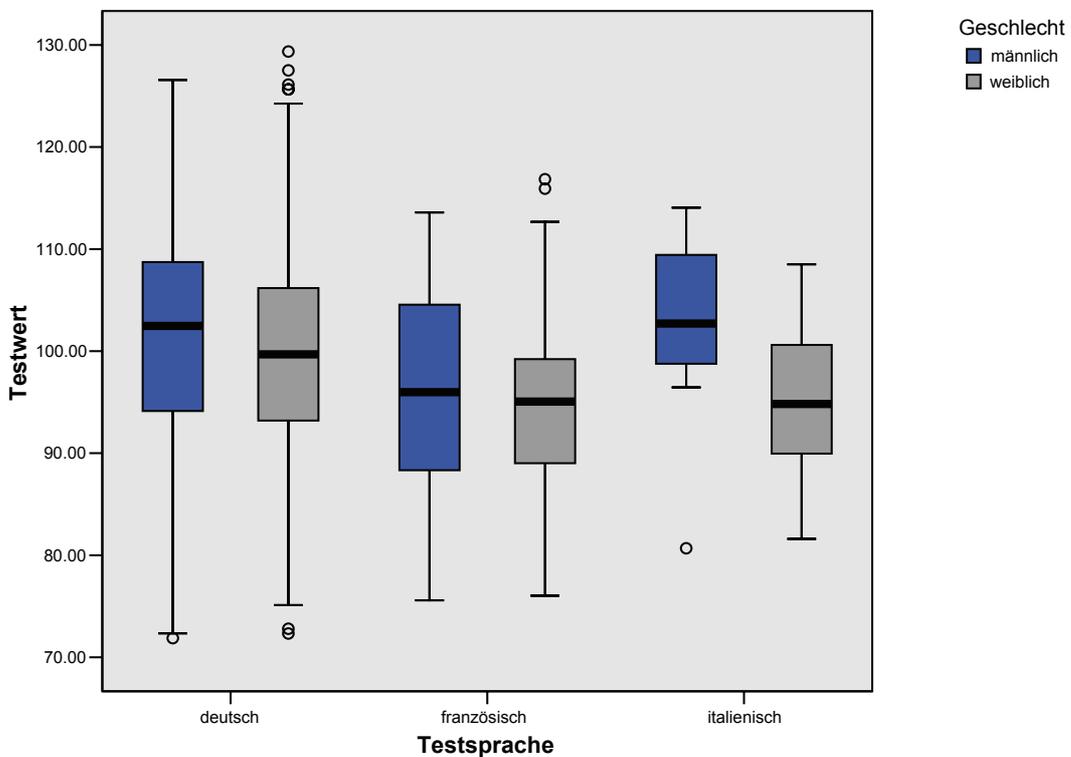


Abbildung 44: Boxplot für die Testwerte nach Geschlecht, getrennt für die drei Sprachgruppen.

Untertest	Geschlecht	m	s	Levene's Test Varianz			t-test Mittelwert		
					F	Sig.	t	df	Sig.
Muster zuordnen	männlich	12.36	3.27	<i>homogen</i>	0.72	0.40	-2.93	1341	0.003
	weiblich	12.90	3.15						
Med.-naturwiss. Grundverst.	männlich	11.54	3.75	<i>homogen</i>	1.05	0.31	2.82	1341	0.005
	weiblich	10.94	3.63						
Schlauchfiguren	männlich	13.01	3.69	<i>homogen</i>	0.55	0.46	6.23	1341	0.000
	weiblich	11.72	3.50						
Quantitative und formale Probleme	männlich	11.63	4.08	<i>homogen</i>	3.05	0.08	9.12	1341	0.000
	weiblich	9.56	3.85						
Textverständnis	männlich	9.64	3.72	<i>homogen</i>	0.47	0.49	4.68	1341	0.000
	weiblich	8.64	3.65						
Figuren lernen	männlich	12.59	3.93	<i>homogen</i>	2.12	0.15	-5.25	1341	0.000
	weiblich	13.75	3.72						
Fakten lernen	männlich	11.78	3.75	<i>homogen</i>	0.04	0.84	-5.17	1341	0.000
	weiblich	12.90	3.72						
Diagramme und Tabellen	männlich	11.07	3.77	<i>homogen</i>	3.69	0.06	5.94	1341	0.000
	weiblich	9.85	3.47						
Konzent. u. sorgfältiges Arbeiten	männlich	11.12	4.38	<i>homogen</i>	0.03	0.86	-1.34	1341	0.179
	weiblich	11.46	4.43						
Planen & Organisieren	männlich	9.89	3.35	<i>homogen</i>	1.80	0.18	0.12	1341	0.908
	weiblich	9.87	3.12						

Tabelle 33: Mittelwerte (m) und Standardabweichungen (s) geschlechtsspezifisch und Ergebnisse der Prüfungen auf Varianzhomogenität und Mittelwertsunterschiede.

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

Beim neuen Untertest „Planen und Organisieren“ treten keine Geschlechterunterschiede bezüglich des erreichten Punktwerts auf.

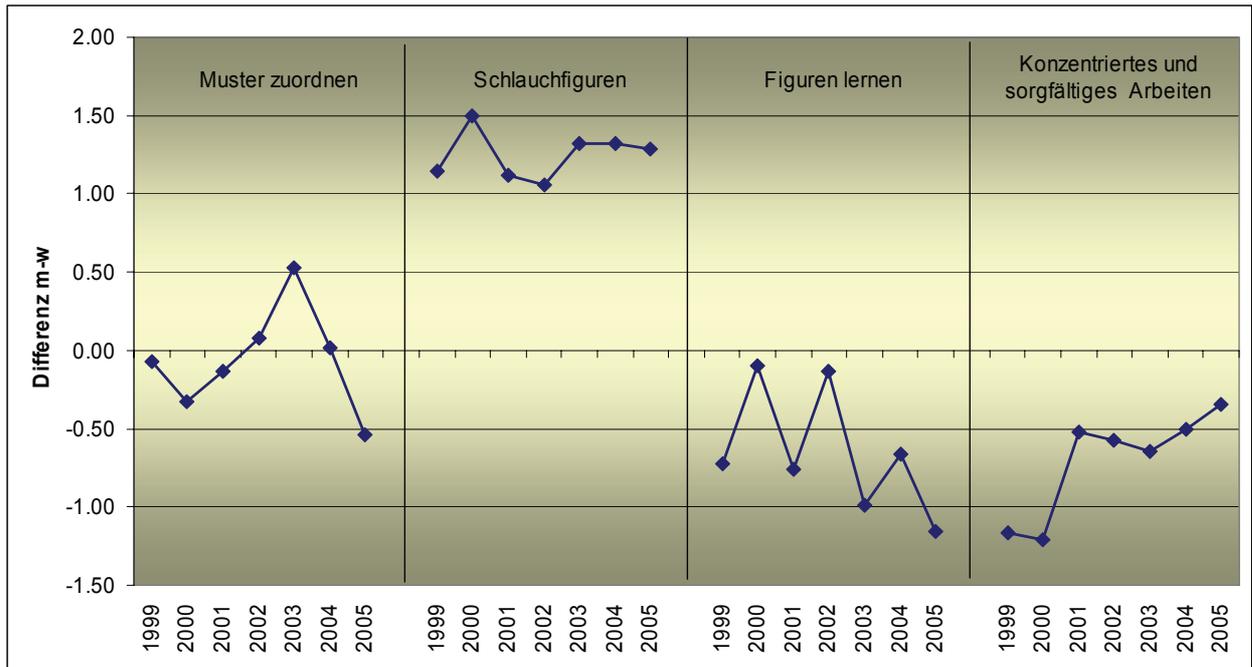


Abbildung 45: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (wenig sprachabhängige Untertests).

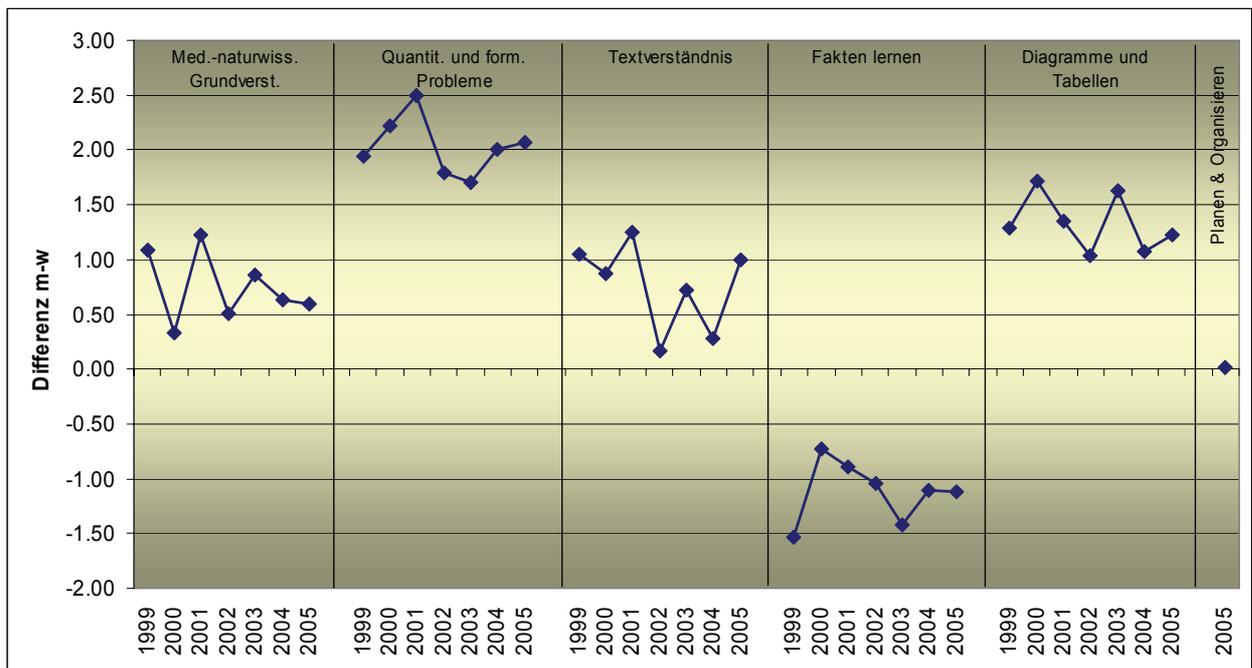


Abbildung 46: Differenzen für Punktwerte der Untertests geschlechtsspezifisch (sprachabhängige Untertests).

Bezüglich der Zulassung ist in der Humanmedizin ein signifikanter Unterschied zwischen den Geschlechtern festzustellen (Tabelle 35). Bei den Männern beträgt die Quote 59.7%, bei den Frauen 48.2% (Tabelle 34).

Disziplin	Geschlecht	nicht zugelassen	zugelassen	Total
Humanmedizin	männlich	150 40.3%	222 59.7%	372 100%
	weiblich	338 51.8%	314 48.2%	652 100%
	Total	488 47.7%	536 52.3%	1024 100%
Veterinärmedizin	männlich	6 37.5%	10 62.5%	16 100%
	weiblich	37 23.4%	121 76.6%	158 100%
	Total	43 24.7%	131 75.3%	174 100%
Zahnmedizin	männlich	4 6.7%	56 93.3%	60 100%
	weiblich	4 4.7%	81 95.3%	85 100%
	Total	8 5.5%	137 94.5%	145 100%

Tabelle 34: Zulassungsquoten geschlechtsspezifisch.

Disziplin		Wert	df	Asymp. Sig. (2-seitig)
Humanmedizin	Pearson Chiquadrat	12.597	1	.000
	Likelihood Ratio	12.657	1	.000
Veterinärmedizin	Pearson Chiquadrat	1.549	1	.213
	Likelihood Ratio	1.426	1	.232
Zahnmedizin	Pearson Chiquadrat	.259	1	.611
	Likelihood Ratio	.256	1	.613

Tabelle 35: Prüfung Unterschiede der Zulassung – geschlechtsspezifisch.

7.6 Vergleich der Kantone

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

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

7.6.1 Vergleich 2005 innerhalb der Deutschschweiz

hoch			mittel			niedrig		
Kanton	Maturitätsquote	Anzahl	Kanton	Maturitätsquote	Anzahl	Kanton	Maturitätsquote	Anzahl
BS	24.4%	76	OW	17.8%	5	AR	15.8%	8
BL	18.6%	85	BE	17.7%	225	TG	15.6%	55
ZH	18.3%	342	ZG	17.5%	24	GR	15.1%	43
			LU	17.1%	76	UR	15.0%	6
			NW	16.6%	7	SH	14.6%	17
			SZ	16.0%	22	SG	14.4%	112
						AI	14.3%	2
						SO	14.3%	51
						AG	13.8%	116
						GL	12.2%	12
Summe		503	Summe		359	Summe		322

Tabelle 36: Gruppierung der Kantone nach der Maturitätsquote (Basis 2003).

Die Einteilung der Kantone nach „hohen“, „mittleren“ und „niedrigen“ Maturitätsquoten beruht auf den Quoten aus dem Jahr 2003 (aktuell verfügbare Zahlen). Gegenüber früheren Jahren ergaben sich in der Kantonszuordnung einige Unterschiede. Nach wie vor kann jedoch kein systematischer Unterschied bezüglich des Testwerts beobachtet werden, welcher in Zusammenhang mit der kantonalen Maturitätsquote stehen würde.

				Zwischen Gruppen				
	Maturitäts- quote	Mittel- wert	Stand.- abw.	Quadrat- summe	df	MQ	F	Sig.
Testwert	hoch	100.37	9.99	53.02	2	26.51	0.26	0.770
	mittel	100.82	10.18	128950.86	1273	101.30		
	niedrig	100.37	10.05	129003.88	1275			
Muster zuordnen	hoch	12.74	3.18	7.24	2	3.62	0.36	0.699
	mittel	12.75	3.31	12850.33	1273	10.09		
	niedrig	12.58	3.05	12857.57	1275			
Medizin.-natur- wiss. Grund- verständnis	hoch	11.06	3.57	27.60	2	13.80	1.02	0.361
	mittel	11.38	3.81	17251.99	1273	13.55		
	niedrig	11.05	3.70	17279.59	1275			
Schlauchfiguren	hoch	12.09	3.55	32.07	2	16.03	1.21	0.298
	mittel	12.38	3.68	16834.93	1273	13.22		
	niedrig	11.98	3.70	16867.00	1275			
Quantitat. u. formale Probleme	hoch	10.12	4.13	19.97	2	9.99	0.61	0.543
	mittel	10.33	4.00	20795.90	1273	16.34		
	niedrig	10.40	3.97	20815.87	1275			
Textverständnis	hoch	8.86	3.68	19.61	2	9.81	0.71	0.493
	mittel	9.11	3.77	17652.19	1273	13.87		
	niedrig	9.11	3.74	17671.80	1275			
Figuren lernen	hoch	13.31	3.91	9.38	2	4.69	0.32	0.728
	mittel	13.48	3.69	18837.20	1273	14.80		
	niedrig	13.27	3.90	18846.59	1275			
Fakten lernen	hoch	12.90	3.70	114.08	2	57.04	4.03	0.018
	mittel	12.19	3.73	18010.55	1273	14.15		
	niedrig	12.41	3.86	18124.63	1275			
Diagramme und Tabellen	hoch	10.20	3.61	12.95	2	6.48	0.49	0.613
	mittel	10.25	3.67	16840.92	1273	13.23		
	niedrig	10.43	3.64	16853.87	1275			
Konzentr. u. sorgf. Arb.	hoch	11.24	4.51	151.52	2	75.76	3.90	0.020
	mittel	10.84	4.33	24704.34	1273	19.41		
	niedrig	11.73	4.34	24855.87	1275			
Planen & Organisieren	hoch	9.66	3.19	44.297	2	22.148	2.165	.115
	mittel	10.11	3.12	13024.389	1273	10.231		
	niedrig	9.92	3.27	13068.686	1275			

Tabelle 37: Mittelwerte und Standardabweichungen für drei Gruppen der Kantone nach der Maturitätsquote.

Kanton		Testwert	Musterzuordnen	Medizin.-naturw. Grundv.	Schlauchfiguren	Quantit. u. form. Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arb.	Planen & Org.
AG	M	102.33	12.84	11.70	12.13	10.85	9.59	13.56	12.99	10.84	12.03	10.33
	S	10.84	3.00	3.77	3.90	4.20	4.03	4.20	3.97	3.74	4.01	3.30
BE	M	100.50	12.62	11.32	12.42	10.05	8.99	13.73	12.25	9.91	10.51	9.91
	S	10.73	3.37	3.91	3.89	4.10	3.71	3.71	3.76	3.73	4.28	3.10
BL	M	99.50	13.20	10.54	12.05	9.49	8.81	13.56	13.11	9.81	12.01	9.07
	S	9.98	2.87	3.50	3.60	4.26	3.27	3.93	3.80	3.67	4.14	3.22
BS	M	97.84	12.62	10.21	11.97	8.84	7.88	13.20	12.33	9.46	10.82	9.01
	S	8.74	3.22	3.42	3.39	3.76	3.59	4.05	3.69	3.36	4.70	2.61
LU	M	101.07	12.88	11.70	12.37	11.07	9.00	12.66	11.97	10.96	11.25	10.64
	S	9.24	3.14	3.70	3.36	3.87	3.75	3.44	3.68	3.53	4.15	2.86
SG	M	99.86	12.18	10.93	11.55	10.74	9.11	13.02	12.11	10.70	11.39	10.29
	S	9.02	3.03	3.45	3.67	3.72	3.58	3.51	3.70	3.52	4.43	3.15
ZH	M	101.28	12.65	11.40	12.20	10.58	9.10	13.35	13.00	10.49	11.15	10.00
	S	10.14	3.25	3.58	3.55	4.12	3.79	3.84	3.68	3.64	4.53	3.25

Tabelle 38: Vergleich der Kantone für Testwert und die einzelnen Untertests.

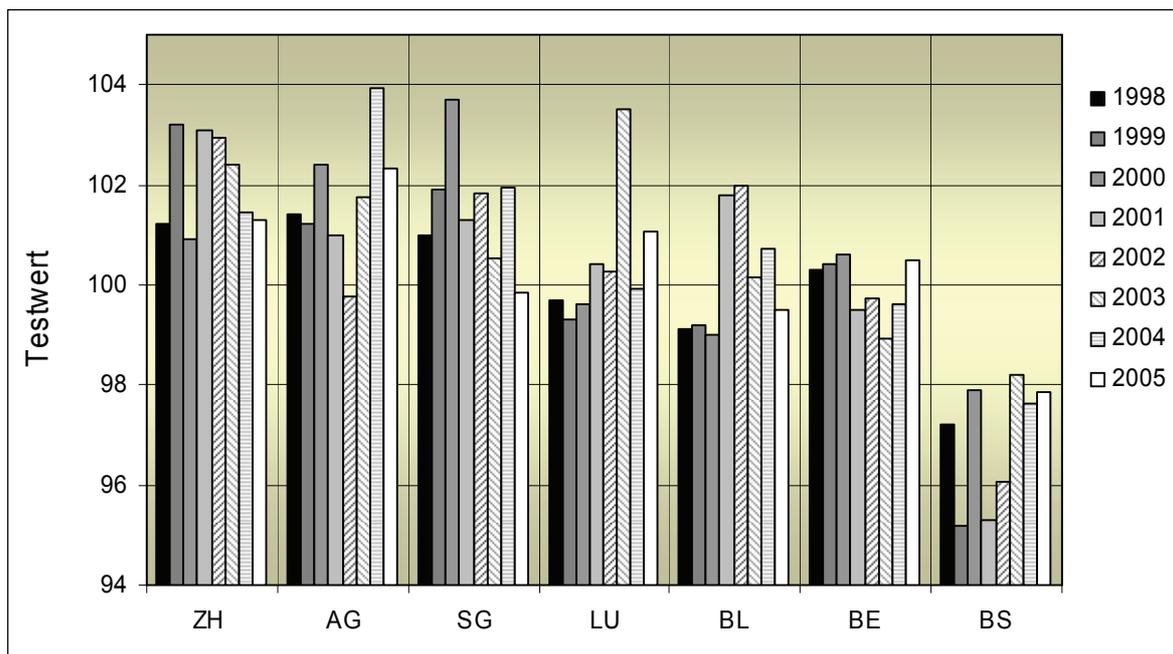


Abbildung 47: Testwert für Kantone im Vergleich.

Tabelle 38 zeigt Testwerte und Untertestpunktwerte der sieben Kantone mit den grössten Teilnehmerzahlen. Abbildung 47 stellt die Testwerte grafisch dar. Der Kanton Aargau erreichte erneut die höchsten mittleren Testwerte.

Besondere Beachtung findet jeweils das Abschneiden von Basel-Stadt. Eine nähere Betrachtung der Untertestergebnisse (Abbildung 48) macht deutlich, dass das tiefe Gesamtergebnis erneut in erster Linie auf die sprachabhängigen Untertests zurückzuführen ist.

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

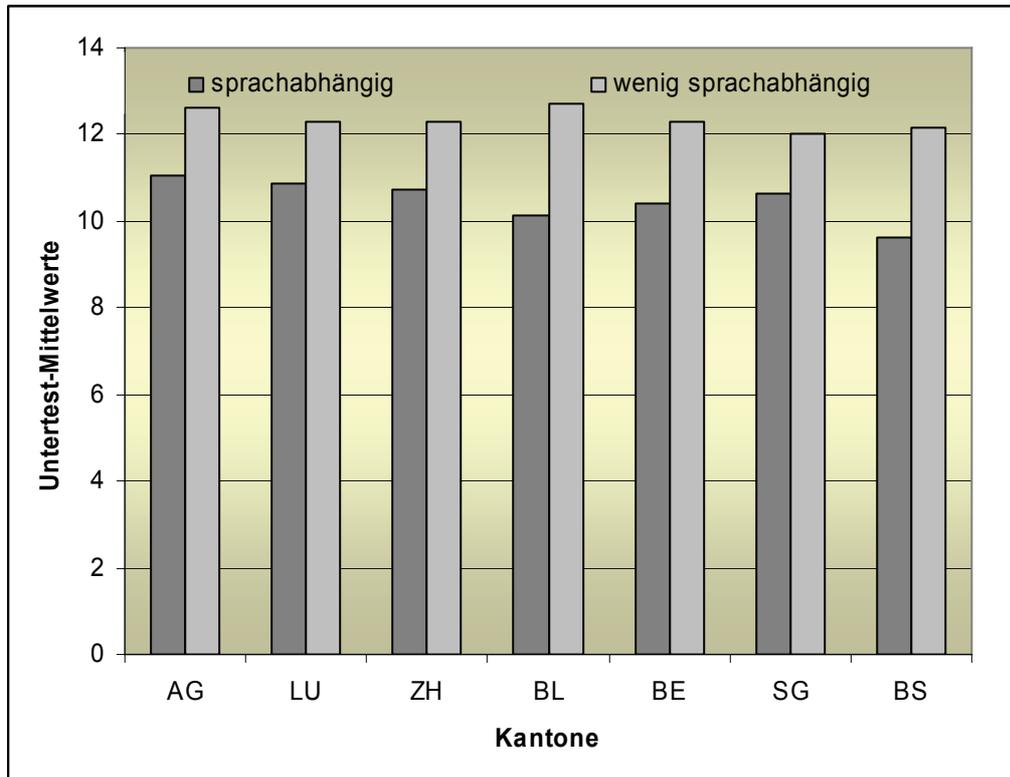


Abbildung 48: Mittelwerte sprachabhängiger und weniger sprachabhängiger Untertests 2005 für die Kantone im Vergleich.

Abbildung 49 stellt den Verlauf über die Jahre anhand der aggregierten Mittelwerte dar. Deutlich sichtbar werden die drei Gruppen ZH-SG-AG / LU-BL-BE / BS, wobei es innerhalb der Gruppen zu Rangverschiebungen kommt, nicht aber dazwischen. Aargau konnte den Aufwärtstrend bestätigen und hat mittlerweile St. Gallen überholt, Bern konnte nach schwächeren Jahrgängen den Abstand zu Luzern auf konstantem Niveau halten und hat gegenüber Basel-land aufgeholt.

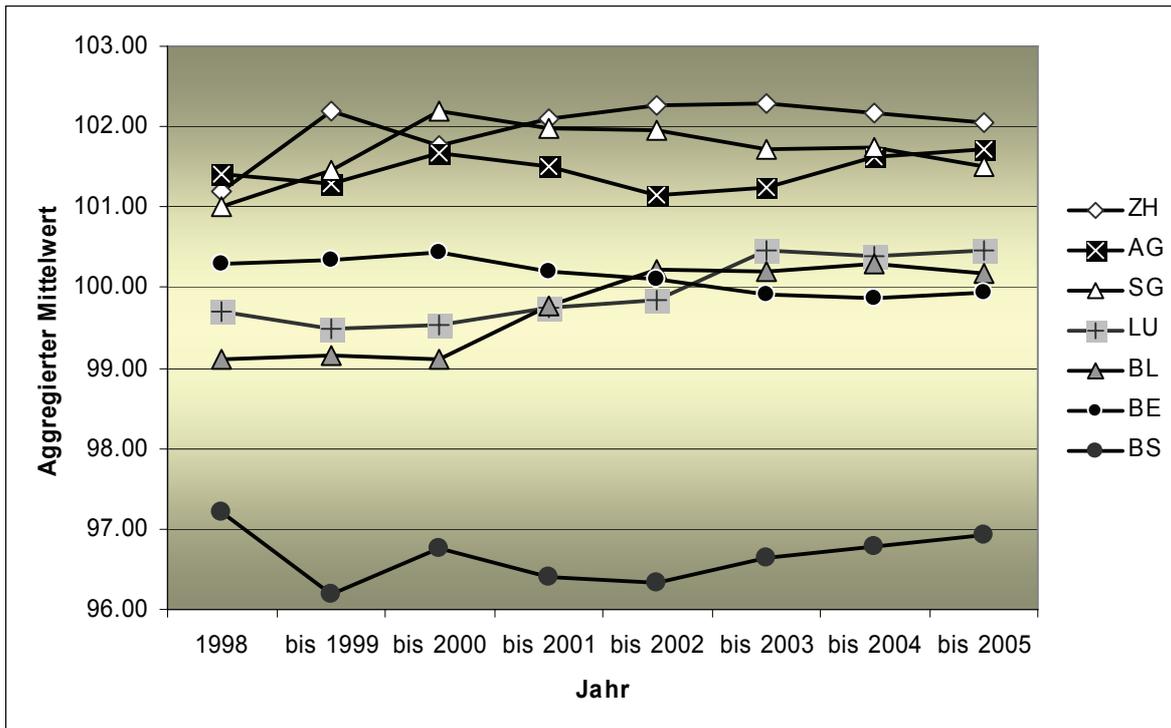


Abbildung 49: Aggregierte Testwerte für Kantone im Vergleich 1998 – 2005.

7.7 Vergleiche für Altersgruppen

Für die Überprüfung der Alterseffekte im Jahr 2005 wurden die Gruppen auf der Basis der Tabelle 39 gebildet. Die älteste Gruppe wurde nahe dem Median des Maturitätsalters geteilt.

Maturitätsjahr	Geburtsjahr			Total
	1951-1981	1982-1984	1985-1988	
1983	2			2
1984	1			1
1986	1			1
1990	1			1
1991	2			2
1992	2			2
1993	1			1
1994	1			1
1995	4			4
1996	2			2
1997	4			4
1998	11			11
1999	13			13
2000	9	2		11
2001	17	10		27
2002	9	34		43
2003	8	74	17	99
2004	24	133	290	447
2005	44	88	695	827
	156	341	1002	1499

* 2 Personen mit fehlenden Angaben

Tabelle 39: Gruppenbildung für das Jahr der Maturitätsprüfung bezogen auf die Geburtsjahre.

Erneut zeigen die älteren Jahrgänge mit später Maturität niedrigere Testwerte als alle anderen Gruppen (Tabelle 40 und Abbildung 50). Die älteren Jahrgänge mit früher Maturität unterscheiden sich hingegen statistisch nicht signifikant von den anderen Gruppen. Die besten Ergebnisse sind sowohl bezüglich des Testwerts wie auch der einzelnen Untertests bei den jüngsten Teilnehmern zu finden.

	Geburtsjahr und Maturität	Mittelwert	Standardabw.	Quadratsumme	df	MQ	F	Sig.	Homogene Gruppen
Testwert	1951-1981, frühe Maturität	99.24	11.47	1254.17	3	418.06	4.22	0.01	
	1951-1981, späte Maturität	97.21	9.54						
	1982-1984	99.21	10.17						
	1985-1988	100.58	9.79						
Muster zuordnen	1951-1981, frühe Maturität	11.66	2.99	111.48	3	37.16	3.76	0.01	
	1951-1981, späte Maturität	12.12	3.14						
	1982-1984	12.63	3.07						
	1985-1988	12.78	3.18						
Med.-naturwiss. Grundverständnis	1951-1981, frühe Maturität	11.56	3.69	123.87	3	41.29	3.05	0.03	
	1951-1981, späte Maturität	10.06	3.68						
	1982-1984	10.80	3.54						
	1985-1988	11.11	3.73						
Schlauchfiguren	1951-1981, frühe Maturität	11.79	3.96	27.44	3	9.15	0.70	0.55	
	1951-1981, späte Maturität	11.53	3.31						
	1982-1984	12.13	3.86						
	1985-1988	11.97	3.52						
Quantitative und formale Probleme	1951-1981, frühe Maturität	10.79	4.28	221.40	3	73.80	4.56	0.00	
	1951-1981, späte Maturität	9.19	3.63						
	1982-1984	9.86	3.95						
	1985-1988	10.48	4.06						
Textverständnis	1951-1981, frühe Maturität	9.17	3.78	63.42	3	21.14	1.57	0.19	
	1951-1981, späte Maturität	8.24	3.45						
	1982-1984	8.76	3.54						
	1985-1988	9.02	3.72						
Figuren lernen	1951-1981, frühe Maturität	11.94	3.70	256.72	3	85.57	5.94	0.00	
	1951-1981, späte Maturität	12.36	3.64						
	1982-1984	13.02	3.89						
	1985-1988	13.48	3.78						
Fakten lernen	1951-1981, frühe Maturität	12.14	4.25	19.04	3	6.35	0.46	0.71	
	1951-1981, späte Maturität	12.47	3.63						
	1982-1984	12.30	3.76						
	1985-1988	12.52	3.68						
Diagramme und Tabellen	1951-1981, frühe Maturität	10.49	4.22	172.92	3	57.64	4.48	0.00	
	1951-1981, späte Maturität	9.14	3.42						
	1982-1984	9.71	3.57						
	1985-1988	10.28	3.56						
Konzentriertes und sorgfältiges Arbeiten	1951-1981, frühe Maturität	10.54	4.66	151.87	3	50.62	2.67	0.05	
	1951-1981, späte Maturität	10.55	5.18						
	1982-1984	10.97	4.57						
	1985-1988	11.47	4.18						
Planen und Organisieren	1951-1981, frühe Maturität	9.17	3.48	192.04	3	64.01	6.31	0.00	
	1951-1981, späte Maturität	8.84	3.02						
	1982-1984	9.40	3.19						
	1985-1988	9.99	3.18						

Tabelle 40: Testwert und Punktwerte für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung; varianzanalytische Prüfung des Unterschiedes.

In Tabelle 40 sind in den letzten Spalten ausserdem die bezüglich des Mittelwertes homogenen Gruppen dargestellt. Schattierungen in der gleichen **Spalte** bedeuten, dass sich die grau markierten Gruppen NICHT voneinander unterscheiden.

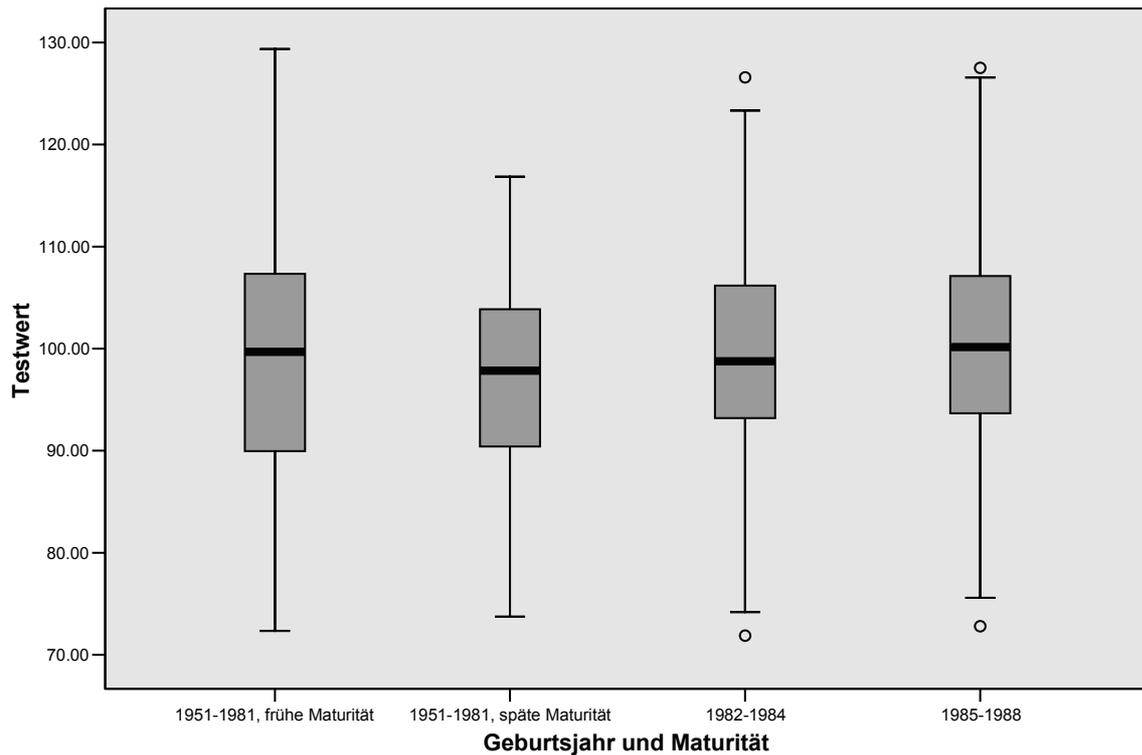


Abbildung 50: Boxplot für Testwert der Altersgruppen (Geburtsjahre) und Maturität.

Geburtsjahr und Maturität	nicht zugelassen	zugelassen	n
1951-1981, frühe Maturität	27 38.0%	44 62.0%	71 100.0%
1951-1981, späte Maturität	45 52.9%	40 47.1%	85 100.0%
1982-1984	153 44.9%	188 55.1%	341 100.0%
1985-1988	398 39.7%	604 60.3%	1002 100.0%
Total	623 41.6%	876 58.4%	1499* 100.0%

*2 Personen mit fehlenden Angaben

Tabelle 41: Zulassungsquoten für die drei Altersgruppen (nach Geburtsjahren) – für die älteste Gruppe aufgeteilt nach früher und später Maturitätsprüfung.

Hinsichtlich der Zulassungsquote (Tabelle 41) schneiden die ältesten Jahrgänge mit früher Maturität besser ab als alle anderen Gruppen.

7.8 Vergleiche nach Wunschuniversitäten

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

Disziplinen	Wunsch-Universität	Zulassung		Total
		nicht zugelassen	zugelassen	
Humanmedizin	Bern	134	157	291
		46.0%	54.0%	100.0%
	Basel	133	132	265
		50.2%	49.8%	100.0%
	Freiburg	44	32	76
	57.9%	42.1%	100.0%	
Zürich	221	254	475	
	46.5%	53.5%	100.0%	
Total	532	575	1107	
	48.1%	51.9%	100.0%	
Veterinärmedizin	Bern	54	65	119
		45.4%	54.6%	100.0%
	Zürich	30	95	125
		24.0%	76.0%	100.0%
	Total	84	160	244
34.4%		65.6%	100.0%	

Tabelle 42: Zulassungsquoten für Disziplin und Wunschuniversität (erste Wahl).

	Value	df	Asymp. Sig.
Pearson Chi-Square	3.265	3	.353
Likelihood Ratio	3.272	3	.352
Cramer's V	.056		.353

Tabelle 43: Signifikanz für Zulassungsquoten nach Wunschuniversität (HM).

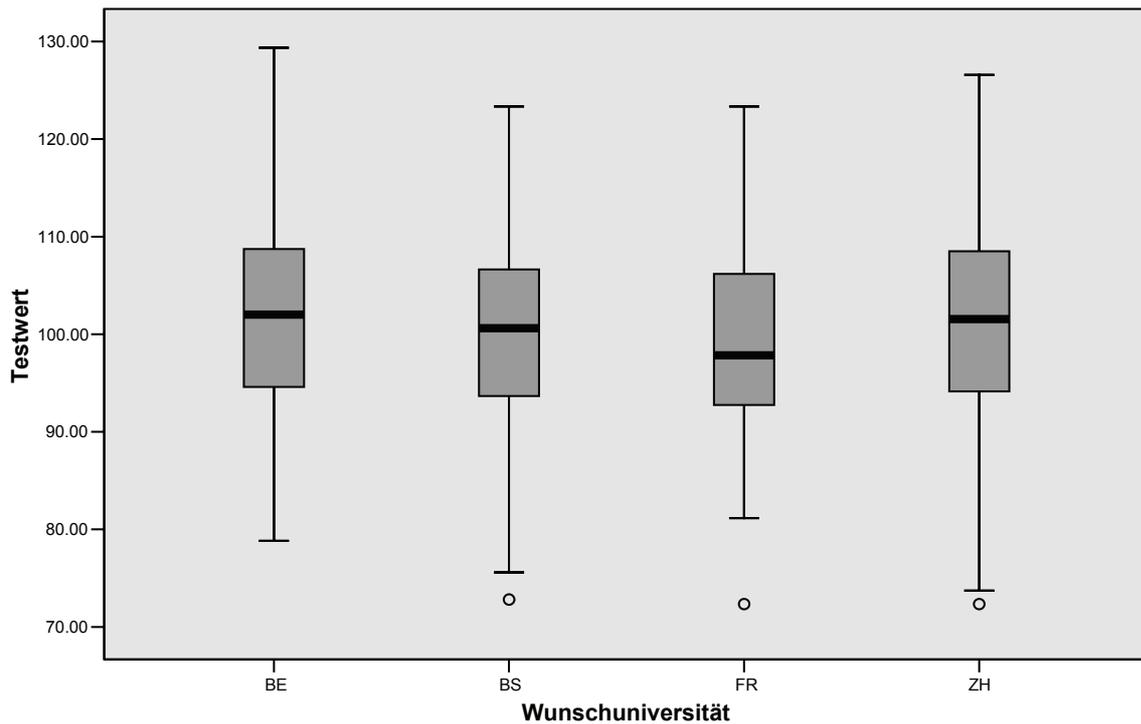


Abbildung 51: Boxplot für Testwerte aller Teilnehmer nach Wunschuniversitäten.

Wunschuniversität	Teilnehmer gesamt			zugelassen		
	N	Mittelwert	Standard-abw.	N	Mittelwert	Standard-abw.
Bern	291	101.69	9.98	157	109.04	6.30
Basel	265	100.05	9.57	132	107.83	5.29
Freiburg	76	99.25	10.62	32	109.35	5.86
Zürich	475	101.29	10.24	254	109.00	6.09
Total	1107	100.96	10.06	575	108.76	5.97

Tabelle 44: Testwerte nach Wunschuniversitäten, aller Teilnehmer und Zugelassenen (HM).

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	153.352	3	51.117	1.437	.231
Innerhalb der Gruppen	20311.322	571	35.571		
Gesamt	20464.675	574			

Tabelle 45: Signifikanzprüfung nach Wunschuniversitäten der zugelassenen Personen (HM).

Betrachtet man nur die zum Studium zugelassenen Personen, so treten zwischen den Universitäten keine signifikanten Unterschiede mehr auf (Tabelle 44 und Tabelle 45).

8 Ergebnisse zur Testgüte

8.1 Zuverlässigkeit

Die Reliabilität kann anhand zweier Koeffizienten verglichen werden. Die internen Konsistenzen (Cronbachs Alpha) schätzen die Messgenauigkeit anhand der Korrelationen jeder Aufgabe mit allen anderen des entsprechenden Untertests.

Untertest	Interne Konsistenz (Cronbachs Alpha)								
	Deutschland	Schweiz (d)							
		1998	1999	2000	2001	2002	2003	2004	2005
Muster zuordnen	.64 - .69	0.69	0.6	0.53	0.56	0.57	0.53	0.54	0.67
Med.-nat. Grundv.	.64 - .73	0.7	0.7	0.61	0.66	0.57	0.72	0.66	0.72
Schlauchfiguren	.76 - .80	0.71	0.74	0.71	0.73	0.68	0.74	0.72	0.72
Quant. u. form. Prob.	.69 - .71	0.76	0.75	0.76	0.74	0.71	0.67	0.71	0.76
Textverständnis	.68 - .71	0.73	0.75	0.74	0.7	0.68	0.78	0.67	0.73
Figuren lernen	.69 - .74	0.72	0.7	0.62	0.64	0.82	0.77	0.71	0.74
Fakten lernen	.62 - .70	0.68	0.72	0.64	0.69	0.7	0.73	0.72	0.72
Diagr. und Tab.	.67 - .71	0.67	0.75	0.63	0.61	0.62	0.7	0.66	0.7
Planen und Org.									0.62

Tabelle 46: Zuverlässigkeit 1998 bis 2005; Vergleich der Skalenkonsistenzen für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

Untertest	Reliabilität nach der Testhalbierungsmethode								
	Deutschland	Schweiz (d)							
		1998	1999	2000	2001	2002	2003	2004	2005
Muster zuordnen	.67 - .73	0.72	0.61	0.62	0.59	0.56	0.55	0.57	0.70
Med.-nat. Grundv.	.65 - .72	0.72	0.71	0.6	0.69	0.57	0.73	0.67	0.73
Schlauchfiguren	.75 - .82	0.75	0.74	0.74	0.72	0.67	0.76	0.74	0.73
Quant. u. form. Prob.	.70 - .75	0.78	0.75	0.74	0.74	0.71	0.68	0.72	0.76
Textverständnis	.69 - .74	0.75	0.7	0.75	0.75	0.71	0.78	0.7	0.75
Figuren lernen	.72 - .75	0.72	0.7	0.62	0.62	0.81	0.78	0.73	0.74
Fakten lernen	.68 - .73	0.7	0.7	0.61	0.72	0.69	0.64	0.72	0.73
Diagr. und Tab.	.68 - .72	0.66	0.62	0.61	0.63	0.64	0.62	0.68	0.71
Planen und Org.									0.63
Gesamter Test	.91 - .93	0.92	0.91	0.9	0.91	0.91	0.92	0.91	0.92

Tabelle 47: Zuverlässigkeit 1998 bis 2005; Vergleich der Reliabilitäten nach der Testhalbierungsmethode für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

Die Reliabilitäten nach der Testhalbierungsmethode (Teilung nach gerad- und ungeradzähligen Aufgaben) wurden ebenfalls verglichen (Tabelle 47). Wegen der Stichprobengrößen werden nur die Ergebnisse des deutschsprachigen Tests der Schweiz verwendet.

Die **Reliabilität des Punktwertes** insgesamt, geschätzt nach der Testhalbierungsmethode (gerade vs. ungerade Itemnummern), auf der Basis aller neun Untertests, liegt mit 0.92 wie gewohnt im sehr hohen Bereich.

Durch den neuen Untertest „Planen und Organisieren“ hat der Test insgesamt eine leichte Korrektur seiner Schwerpunkte erhalten. Das Weglassen des Untertests „Konzentriertes und sorgfältiges Arbeiten“ würde zu einer minimal höheren Gesamtkonsistenz führen. Das Ausmass dieser Auswirkung allein scheint jedoch kaum geeignet, eine derartig einschneidende Veränderung in der Struktur des Gesamttests zu rechtfertigen.

Untertest	Kons. ohne den U-test 1998	Kons. ohne den U-test 1999	Kons. ohne den U-test 2000	Kons. ohne den U-test 2001	Kons. ohne den U-test 2002	Kons. ohne den U-test 2003	Kons. ohne den U-test 2004	Kons. ohne den U-test 2005
Muster zuordnen	.82	.80	.80	.78	.79	.80	.77	.81
Med.-nat. Grundverständnis	.79	.78	.78	.77	.77	.78	.75	.79
Schlauchfiguren	.80	.79	.79	.78	.78	.79	.75	.80
Quant. u. form. Probleme	.80	.79	.78	.78	.77	.78	.75	.79
Textverständnis	.80	.79	.78	.77	.78	.77	.75	.79
Figuren lernen	.81	.80	.80	.78	.78	.79	.76	.81
Fakten lernen	.82	.81	.80	.79	.79	.80	.77	.81
Diagramme und Tabellen	.80	.78	.78	.77	.77	.78	.74	.79
Konz. u. sorgf. Arbeiten	.82	.81	.80	.79	.79	.80	.78	.83
Planen und Organisieren								.80
TESTPROFIL	.83	.81	.81	.80	.80	.81	.78	.82

Tabelle 48: Interne Konsistenzen, wenn der jeweilige Untertest weggelassen worden wäre und Konsistenz des Testprofils („standardized alpha“).

8.2 Binnenstruktur

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

Von besonderem Interesse dürfte dieses Jahr die Einordnung des neu eingeführten Untertests „Planen und Organisieren“ sein.

	Muster zuordnen	Med.-naturwis. Grundv.	Schlauchfiguren	Quantitat. und formale Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arbeiten	Planen und Organisieren	Punktwert
Muster zuordnen		0.22	0.38	0.20	0.19	0.37	0.27	0.22	0.32	0.18	0.52
Med.-naturw. Grundverst.	0.22		0.32	0.57	0.69	0.24	0.26	0.62	0.22	0.52	0.74
Schlauchfiguren	0.38	0.32		0.36	0.29	0.36	0.28	0.36	0.26	0.27	0.62
Quant. u. formale Probl.	0.20	0.57	0.36		0.53	0.19	0.21	0.61	0.20	0.49	0.71
Textverständnis	0.19	0.69	0.29	0.53		0.21	0.26	0.60	0.17	0.50	0.71
Figuren lernen	0.37	0.24	0.36	0.19	0.21		0.40	0.22	0.21	0.20	0.55
Fakten lernen	0.27	0.26	0.28	0.21	0.26	0.40		0.23	0.15	0.18	0.52
Diagramme und Tabellen	0.22	0.62	0.36	0.61	0.60	0.22	0.23		0.23	0.53	0.74
Konzent. u. sorgfält. Arbeiten	0.32	0.22	0.26	0.20	0.17	0.21	0.15	0.23		0.22	0.50
Planen und Organisieren	0.18	0.52	0.27	0.49	0.50	0.20	0.18	0.53	0.22		0.64

Tabelle 49: Korrelationen zwischen Punktwerten und Untertests.

Abbildung 52 zeigt, dass die Korrelationen mit der Gesamtpunktzahl für die einzelnen Untertests über die Jahre mehr oder weniger stabil bleiben.

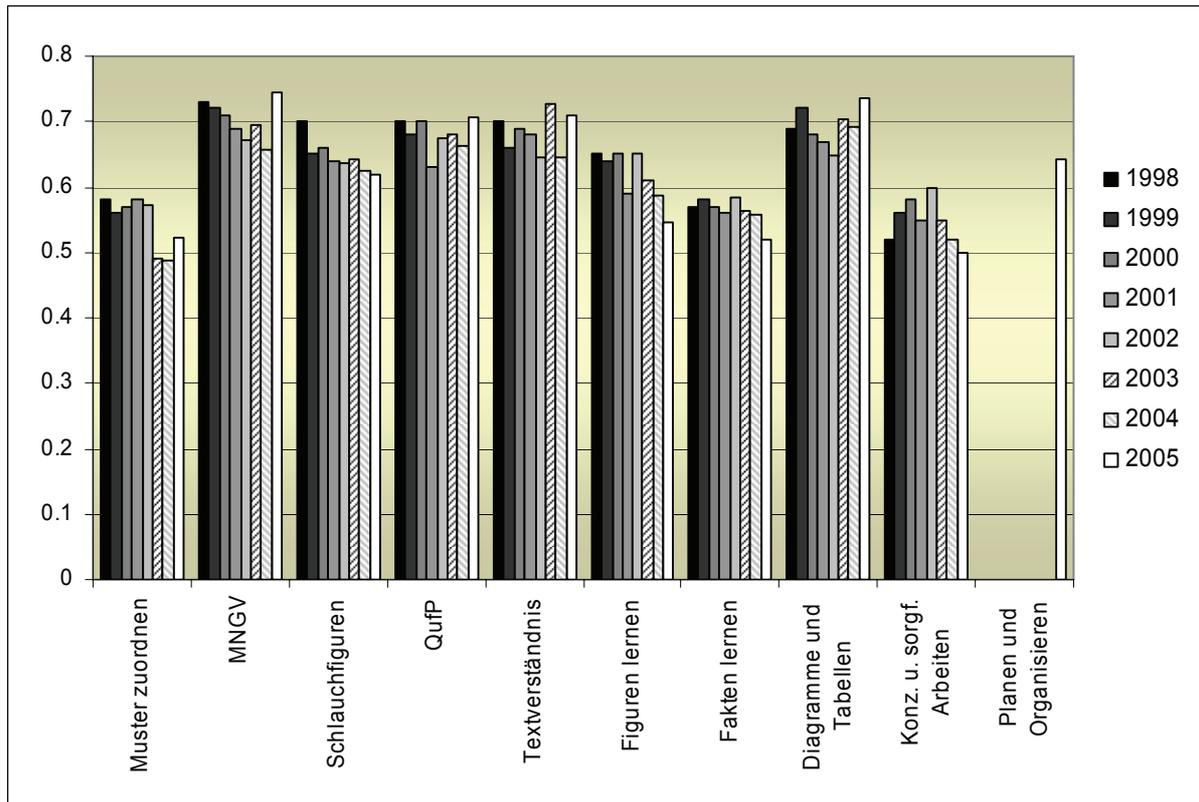


Abbildung 52: Korrelationen zwischen Punktwerten und Untertests nach Jahren.

Die Drei-Faktoren-Struktur bleibt ebenfalls stabil (Tabelle 50) – trotz des neuen Untertests.

Unrotierte Lösung															
Fakt.	Eigenwert					% Varianz					Kumuliert				
	2005	2004	2003	2002	2001	2005	2004	2003	2002	2001	2005	2004	2003	2002	2001
1	4.04	3.36	3.64	3.61	3.53	40.38	37.33	40.48	40.14	39.21	40.38	37.33	40.48	40.14	39.21
2	1.48	1.39	1.6	1.54	1.63	14.77	15.42	17.78	17.12	18.14	55.14	52.74	58.26	57.26	57.35
3	0.91	0.9	0.92	0.77	0.85	9.15	10.05	10.28	8.53	9.39	64.29	62.79	68.54	65.79	66.74
4	0.71	0.82	0.67	0.68	0.62	7.09	9.06	7.5	7.51	6.93	71.38	71.85	76.03	73.3	73.67
5	0.61	0.62	0.54	0.58	0.57	6.06	6.9	5.95	6.41	6.35	77.44	78.75	81.98	79.71	80.03
6	0.58	0.52	0.5	0.53	0.55	5.82	6.36	5.54	5.94	6.06	83.26	85.11	87.52	85.64	86.09
7	0.53	0.54	0.48	0.49	0.5	5.26	6.03	5.38	5.4	5.53	88.52	91.14	92.9	91.05	91.61
8	0.47	0.47	0.35	0.42	0.42	4.69	4.623	3.93	4.67	4.63	93.21	95.76	96.83	95.72	96.24
9	0.37	0.38	0.29	0.39	0.34	3.74	4.24	3.17	4.28	3.76	96.95	100	100	100	100
10	0.30					3.05					100				

Tabelle 50: Faktorenanalyse: Verlauf der unrotierten Lösung.

Die rotierten Lösungen sind in Tabelle 51 dargestellt. Die Zwei-Faktorenlösung klärt 55% der Gesamtvarianz auf, die Drei-Faktorenlösung 64%. Der Scree-Plot (Abbildung 53) stellt den Eigenwerteverlauf der Faktoren grafisch dar. Die Ein-Faktorenlösung bleibt deutlich hervor-gehoben.

Varimax-rotierte Lösungen														
Eigenwert					% Varianz					Kumuliert %				
2005	2004	2003	2002	2001	2005	2004	2003	2002	2001	2005	2004	2003	2002	2001
Zwei-Faktorenlösung														
3.26	2.63	2.87	2.66	2.72	32.58	29.26	31.9	29.6	30.19	32.58	29.26	31.9	29.6	30.19
2.27	2.11	2.37	2.49	2.44	22.70	23.48	26.36	27.66	27.16	55.28	52.74	58.26	57.26	57.35
Drei-Faktorenlösung														
3.27	2.56	2.84	2.61	2.71	32.74	28.39	31.57	29.05	30.06	32.74	28.39	31.57	29.05	30.06
1.73	1.93	1.83	1.67	1.82	17.31	21.43	20.34	18.53	20.21	50.05	49.81	51.91	47.57	50.27
1.43	1.17	1.5	1.64	1.48	14.33	12.97	16.62	18.22	16.47	64.38	62.79	68.54	65.79	66.74

Tabelle 51: Varianzanteile der einzelnen Faktorenlösungen (rotierte Lösungen).

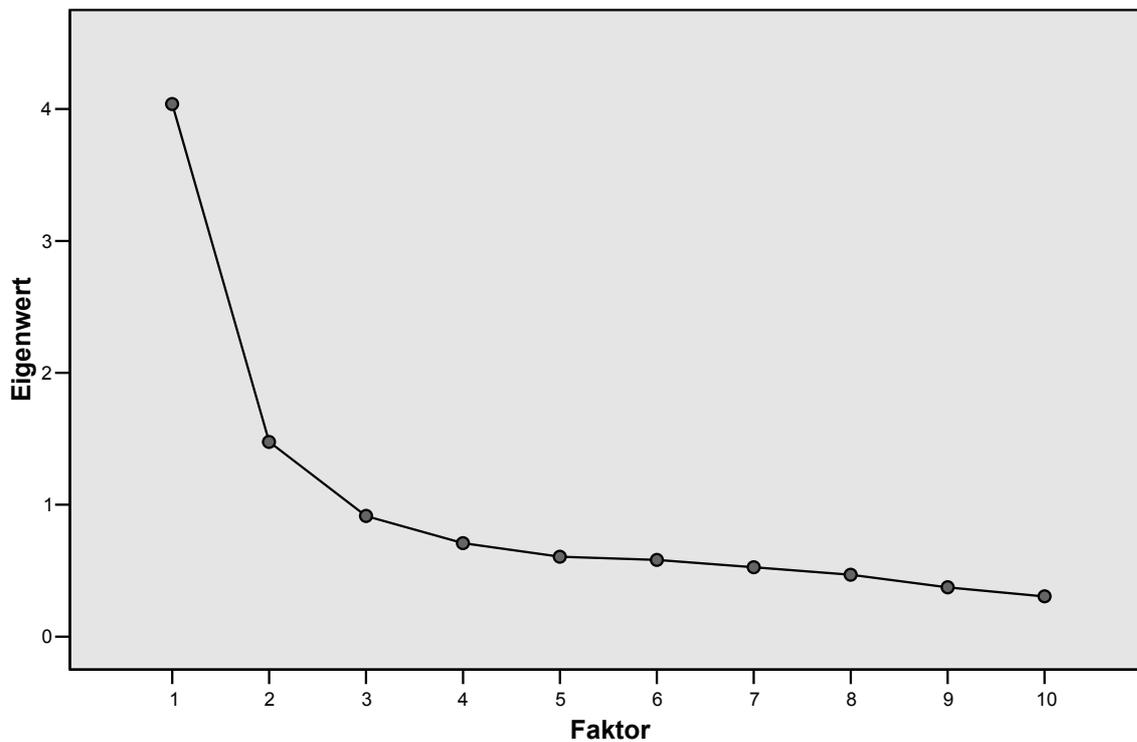


Abbildung 53: Scree-Plot (Eigenwerteverlauf) der Faktoren 1 bis 10 für die Untertests des EMS 2005.

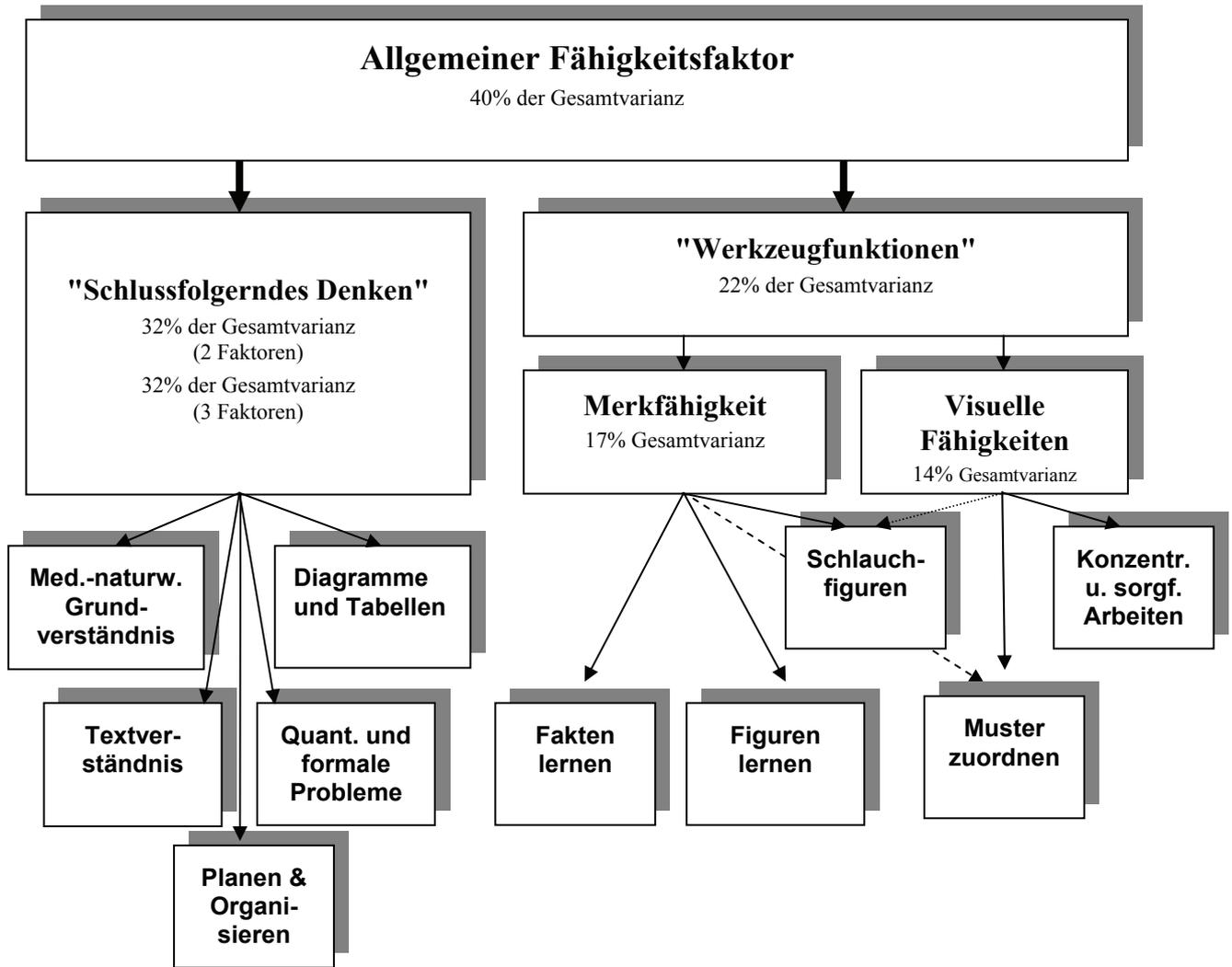


Abbildung 54: Struktur der Untertests des EMS, hierarchische Faktorenstruktur 2005.

	1998	1999	2000	2001	2002	2003	2004	2005
Muster zuordnen	.56	.53	.57	.57	.71	.48	.45	.46
Med.-naturwiss. Grundverst.	.76	.76	.74	.74	.70	.74	.71	.79
Schlauchfiguren	.69	.64	.65	.62	.69	.61	.61	.58
Quant. und formale Probleme	.70	.70	.72	.66	.68	.73	.70	.75
Textverständnis	.72	.70	.71	.72	.62	.77	.70	.76
Figuren lernen	.62	.59	.53	.56	.62	.55	.54	.48
Fakten lernen	.53	.52	.55	.52	.57	.51	.51	.47
Diagramme und Tabellen	.72	.75	.71	.72	.56	.76	.74	.78
Konzentr. u. sorgf. Arbeiten	.47	.51	.55	.49	.54	.48	.44	.41
Planen und Organisieren								.68

Tabelle 52: Ladungen im ersten Faktor der unrotierten Lösung 1998 bis 2005 („Generalfaktor“ bzw. allgemeiner Fähigkeitsfaktor).

	Faktor 1					Faktor 2					Kommunalitäten (h^2)				
	01	02	03	04	05	01	02	03	04	05	01	02	03	04	05
Diagramme und Tabellen	0.77	0.77	0.83	0.79	0.81	0.21					0.63	0.63	0.73	0.66	0.7
Med.-naturwiss. Grundverständnis	0.82	0.77	0.86	0.8	0.82		0.21				0.7	0.63	0.76	0.65	0.71
Textverständnis	0.82	0.82	0.83	0.7	0.81						0.7	0.68	0.74	0.63	0.68
Quant. und formale Probleme	0.81	0.79	0.77	0.7	0.78			0.2			0.66	0.65	0.65	0.6	0.64
Planen und Organisieren					0.72										0.54
Konzentr. u. sorgf. Arbeiten						0.79	0.82	0.79	0.33	0.51	0.62	0.68	0.62	0.2	0.28
Figuren lernen		0.2				0.66	0.7	0.78	0.74	0.74	0.46	0.52	0.65	0.56	0.56
Muster zuordnen		0.2				0.69	0.63	0.58	0.7	0.74	0.5	0.43	0.62	0.47	0.56
Fakten lernen			0.26			0.62	0.67	0.5	0.67	0.62	0.41	0.47	0.78	0.47	0.4
Schlauchfiguren	0.25	0.29	0.26		0.29	0.66	0.6	0.66	0.67	0.62	0.5	0.45	0.6	0.52	0.47

Tabelle 53: Ladungen und Kommunalitäten der Zwei-Faktorenlösung, varimaxrotiert.

Faktor	Schweiz								Deutschland		
	1		2		3		h ²		1	2	3
	2004	2005	2004	2005	2004	2005	2004	2005			
Diagramme und Tabellen	0.77	0.81	0.24				0.67	0.7	0.82		
Medizin.-naturwiss. Grundverständnis	0.81	0.82					0.68	0.72	0.81		0.2
Quantitat. u. formale Probleme	0.73	0.78	0.23				0.6	0.64	0.8	0.18	
Textverständnis	0.8	0.81					0.66	0.69	0.79		0.2
Planen und Organisieren		0.72						0.55			
Muster zuordnen			0.54	0.63	0.52	0.45	0.57	0.6		0.81	
Konzentr. u. sorgfält. Arbeiten	0.2		0.85	0.85			0.76	0.75		0.7	0.4
Schlauchfiguren	0.23			0.43	0.66	0.44	0.53	0.47	0.35	0.71	
Fakten lernen					0.74	0.8	0.58	0.67	0.21		0.87
Figuren lernen				0.21	0.76	0.77	0.6	0.65	0.13	0.47	0.64

Tabelle 54: Faktorenanalyse: Varimaxrotierte Drei-Faktorenlösung, Schweiz und Deutschland (deutsche Daten nach Blum, 1996, in Trost et al., 1998, S. 42).

In der Drei-Faktoren-Lösung werden die Faktoren durch die Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Konzentriertes und sorgfältiges Arbeiten“ und „Fakten lernen“ charakterisiert.

8.3 Ein neuer Untertest: Planen und Organisieren

Erstmals seit Einführung des EMS wurde 2005 ein neuer Untertest in den Test aufgenommen. Dies geschah als Antwort auf die durch die Studienreform geänderten Anforderungen. In Zusammenarbeit mit Vertretern der Universitäten gelangte man 2003 im Rahmen eines Workshops zu der Hypothese, dass die Fähigkeit zu Planung und Selbstorganisation in immer stärkerem Masse den Studienerfolg mitbestimmt und entsprechend Eingang in die Vorhersage des Studienerfolgs erhalten sollte.

Erste Rückmeldungen aus den Testlokalen deuteten darauf hin, dass der Untertest von den Testteilnehmern als „leicht“ empfunden wurde. Aus den Protokollen war ersichtlich, dass fast 50% der Teilnehmer den Untertest bereits 10 Minuten vor Ablauf der Zeit (insgesamt 60 Minuten) als „beendet“ betrachteten. Aufgrund der erreichten Punktzahlen in diesem Untertest kann diese Einschätzung als „zu optimistisch“ eingestuft werden.

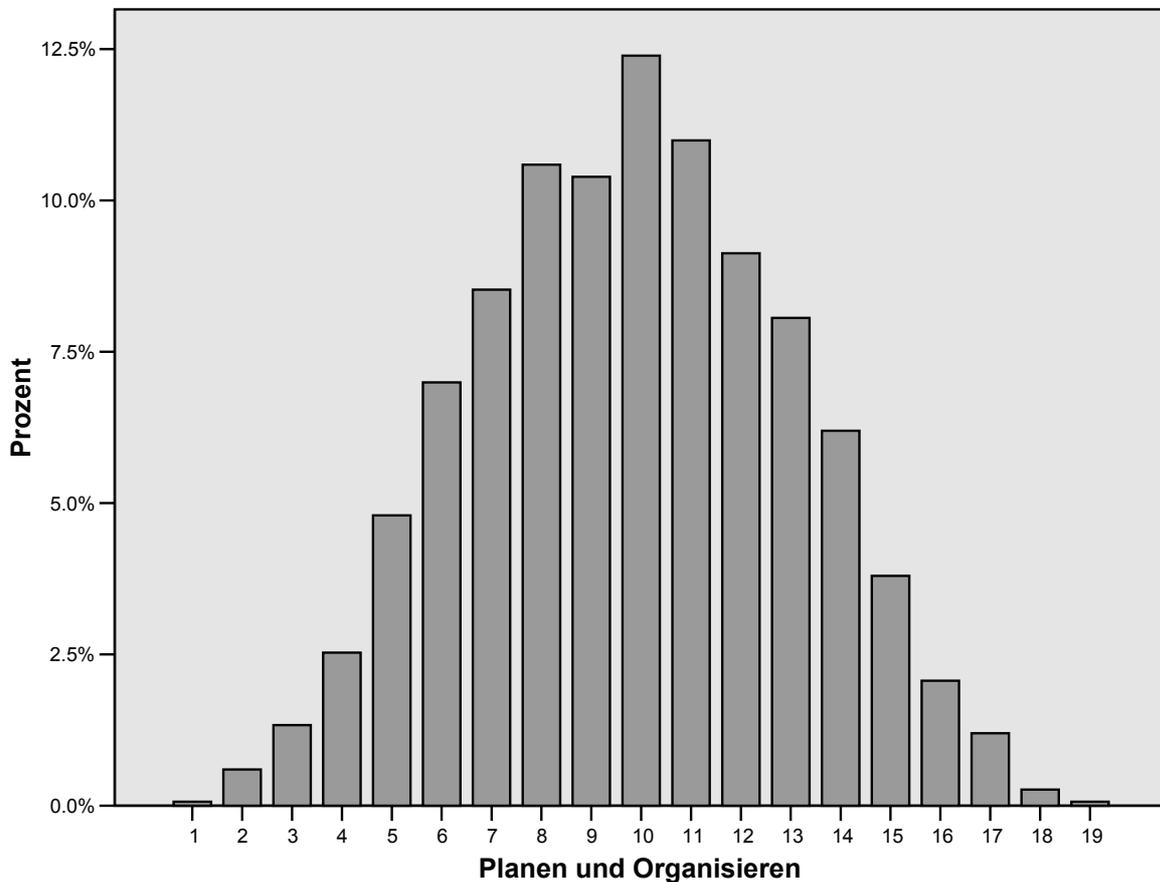


Abbildung 55: Verteilung der Punktzahlen für „Planen und Organisieren“ (Gesamtstichprobe).

Die Schwierigkeiten der Items dieses Untertests sind in Abbildung 56 dargestellt. Die hohe Schwierigkeit des Items 118 ist wohl auch durch dessen Position am Ende des Untertests mitbestimmt, Item 103 als fünftes im Untertest erreicht jedoch eine unerwartete und auffällig hohe Schwierigkeit.

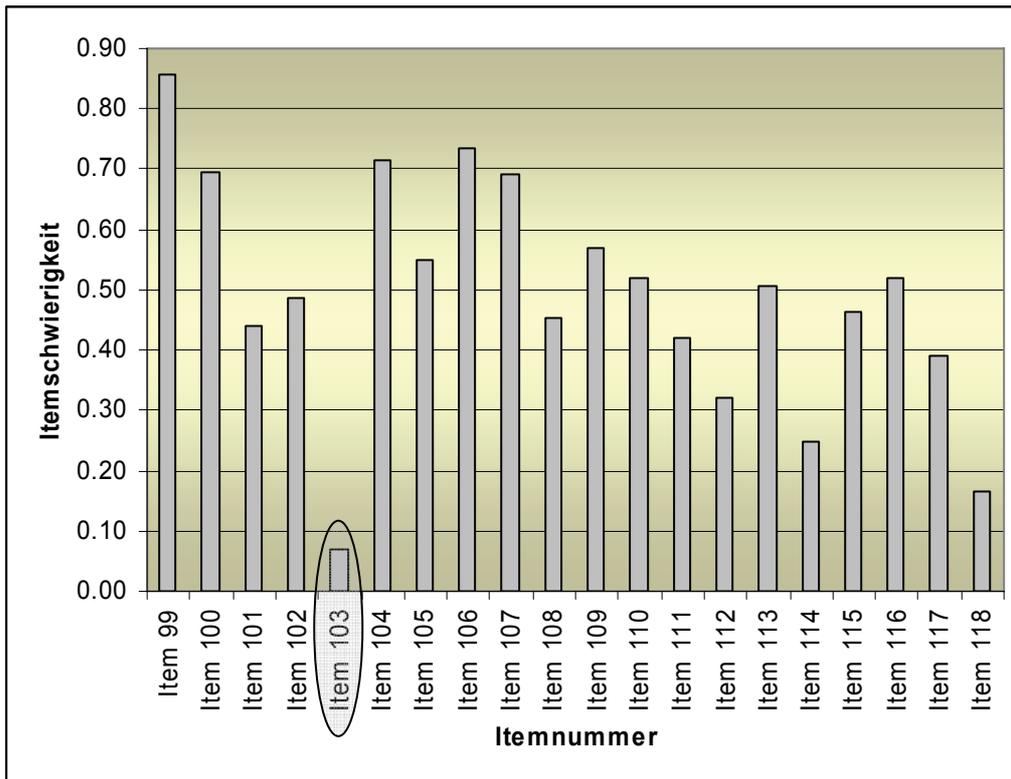


Abbildung 56: Itemschwierigkeiten „Planen und Organisieren“.

Die Trennschärfen der Items sind aus Abbildung 57 ersichtlich. Das fragliche Item 103 schneidet hier mit einer negativen Trennschärfe (die aufgrund der hohen Schwierigkeit zu erwarten war) ungenügend ab.

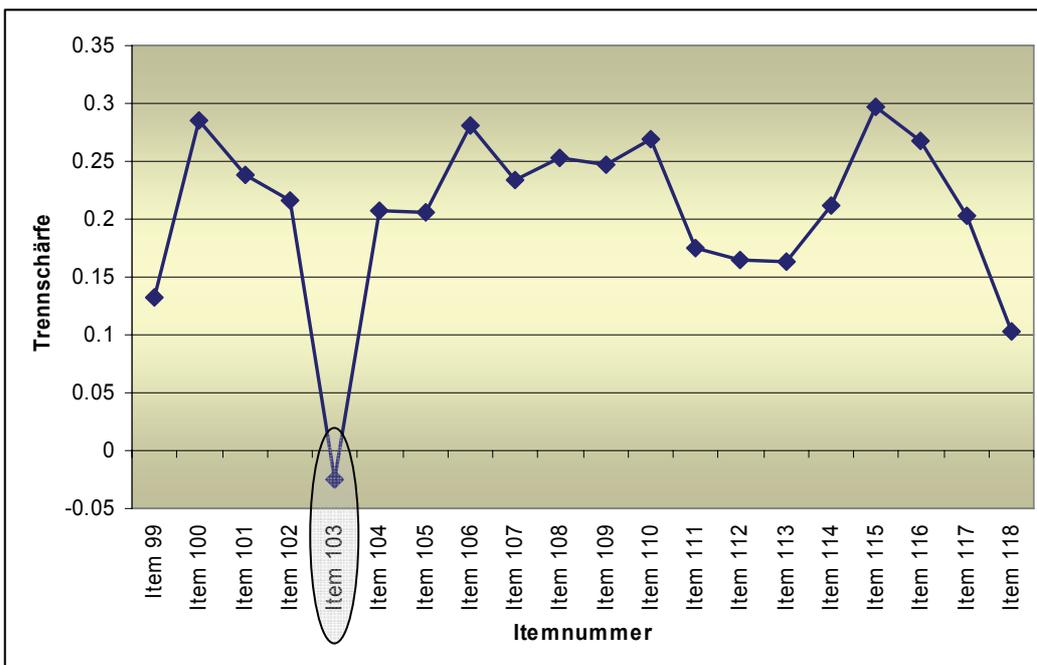


Abbildung 57: Itemtrennschärfen „Planen und Organisieren“.

Abbildung 58 zeigt die Verteilung der Punktwerte im Untertest „Planen und Organisieren“ (ohne Item 103) nach der Wahl der Lösungen bei Item 103. Die Lösung „3“ wäre laut Lösungsschlüssel korrekt. Es zeigt sich, dass die Personen mit den höchsten Ergebnissen in diesem Untertest die Lösungen 1 und 4 bevorzugten.

Aufgrund seiner unzureichenden Kennwerte wird das Item 103 in diesem Untertest nicht gewertet – diese Möglichkeit wurde vorab definiert. Wenn mindestens noch 18 Aufgaben wie bei „Textverständnis“ mit ausreichenden Kennwerten resultieren, besteht keine Einschränkung für die Verwendung des Untertests.

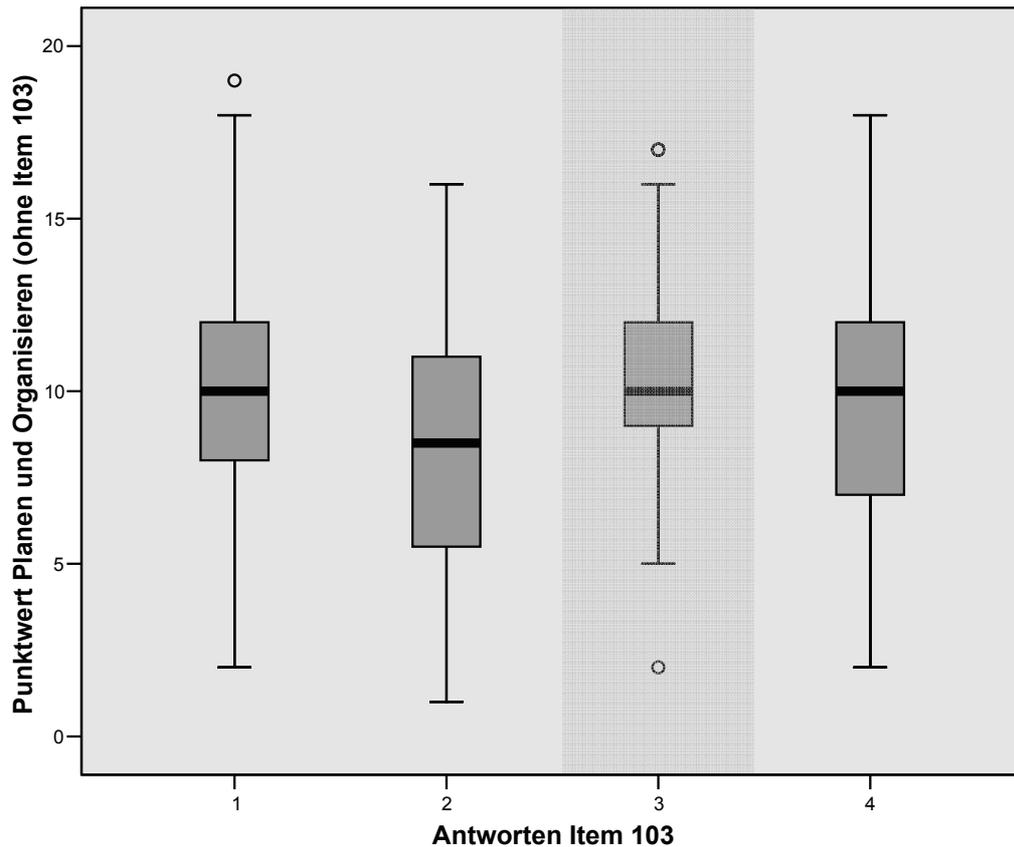


Abbildung 58: Punktwertverteilung „Planen und Organisieren“ nach Antwortstufen Item 103.

8.4 Konzentriertes und sorgfältiges Arbeiten

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

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

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

2005 wurde die nachfolgende Anforderung verwendet: In einer Zeilenfolge von d und b waren diejenigen b zu markieren, denen ein d folgt. Das Testblatt besteht aus 40 Zeilen mit je 40 Zeichen (1600 mögliche zu bearbeitende Zeichen). Davon sind 800 Zeichen b. 400 davon sind durchzustreichen (Target), 400 nicht (Non-Target).

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

Als Parameter ergeben sich:

	Beschreibung	Wertebereich
Richtige	Target b vor d markiert	0...400
Fehler I	Target b vor d nicht markiert, „übersehen“	0...400
Fehler IIa	Nontarget b markiert, „falscher Alarm“	0...400
Fehler IIb	Distraktor d markiert, „falscher Alarm“	0...800

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

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

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

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

Bisherige Version (b mit 2 Strichen) (K = 30)	Neue Version (bd-Test) (K = 20)
Alles richtig	
PW = (600 – 0 – 0) / 30 = 20	PW = (400 – 0 – 0 - 0) / 20 = 20
Nur letztes Zeichen angestrichen - Maximum „Übersehen“	
PW = (1 – 600 – 0)/30 = -20 gesetzt 0	PW = (1 – 400 – 0 – 0)/20 = -20 gesetzt 0
Alle Zeichen angestrichen Fehler Maximum: „Fälschlich angestrichen“	
PW = (600 – 0 – 600)/30 = 0	PW = (400 – 0 – 1200 -0)/20 = -40 ->0
Alle p's angestrichen - Fehler Maximum: Non-Target fälschlich angestrichen“	
Entfällt	PW = (400 – 0 – 400 -0)/20 = 0
Teile werden übersprungen, Beitrag zum Fehler „Übersehen“	
PW = (0 – Übersehen - 0)/30: negativ Nicht „neutral“, sondern mindert PW	P (0 – Übersehen – 0 - 0)/20 = negativ Nicht „neutral“, sondern mindert PW
Leeren Bogen abgegeben	
PW = (0 – 0 - 0)/30 = 0	PW = (0- 0- 0 – 0)/20 = 0

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

	richtig	Fehler	Fehler I	Fehler IIa	Fehler IIb
Mittelwert	229.6649	20.5789	10.1439	5.1046	5.3304
Median	228.0000	11.0000	4.0000	2.0000	1.0000
Standardabweichung	43.68678	39.51086	20.83687	19.10128	15.21149
Schiefe	-.128	8.453	7.634	11.725	12.197
Kurtosis	1.672	102.363	84.739	152.489	224.408

Tabelle 55: Statistiken für Parameter des Untertests „Konzentriertes und sorgfältiges Arbeiten“.

	Testsprache					
	deutsch		französisch		italienisch	
	Mittelwert	Standard-abweichung	Mittelwert	Standard-abweichung	Mittelwert	Standard-abweichung
Richtig	231.25	43.61	214.03	45.51	221.35	32.14
Fehler	21.07	39.88	17.49	42.09	13.72	12.89
Fehler I	10.36	20.96	9.31	22.83	5.85	7.70
Fehler IIa	5.16	19.40	5.31	19.33	2.87	2.69
Fehler IIb	5.55	15.88	2.87	6.75	5.00	8.08

Tabelle 56: Sprachspezifische Statistiken für Parameter des Untertests „Konzentriertes und sorgfältiges Arbeiten“.

Insgesamt scheint die Anforderung etwas schwieriger als beim ursprünglichen Untertest zu sein: Es werden im Mittel nur ca. die Hälfte der Zeichen bearbeitet wie in der Vorgängerversion. Beim bisherigen Konzentrationstest wurden zwischen 4 und 5% Fehler gemacht, jetzt sind

es 8 bis 9%. Die Genauigkeit hat dadurch etwas an Bedeutung für den Punktwert zugenommen. Auch das ist gewünscht, weil Sorgfalt und Arbeitsgeschwindigkeit gleichgewichtig sein sollen.

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

		Fehler	Fehler_I	Fehler_Ila	Fehler_Ilb
Richtig	Korrelation nach Pearson	-.159(**)	-.219(**)	-.030	-.074(**)
	Signifikanz (2-seitig)	.000	.000	.242	.004
Fehler	Korrelation nach Pearson		.827(**)	.618(**)	.688(**)
	Signifikanz (2-seitig)		.000	.000	.000
Fehler_I	Korrelation nach Pearson			.210(**)	.516(**)
	Signifikanz (2-seitig)			.000	.000
Fehler_Ila	Korrelation nach Pearson				.063(*)
	Signifikanz (2-seitig)				.014

Tabelle 57: Korrelationen der Parameter im Konzentrationstest;

**/* Die Korrelation ist auf dem Niveau von 0,01/0,05 (2-seitig) signifikant.

Die Beziehung zwischen Mengen- und Sorgfaltsleistung zeigt keine lineare Abhängigkeit, die Korrelation mit $r = -0.10$ ist angesichts der hohen Personenzahl vergleichsweise gering. Extreme Fehlerzahlen resultieren vor allem durch Übersehen von Zeichen, nicht instruktionsgemässes Vorgehen (alle b durchgestrichen) und nicht regelkonformer Markierung der Zeichen.

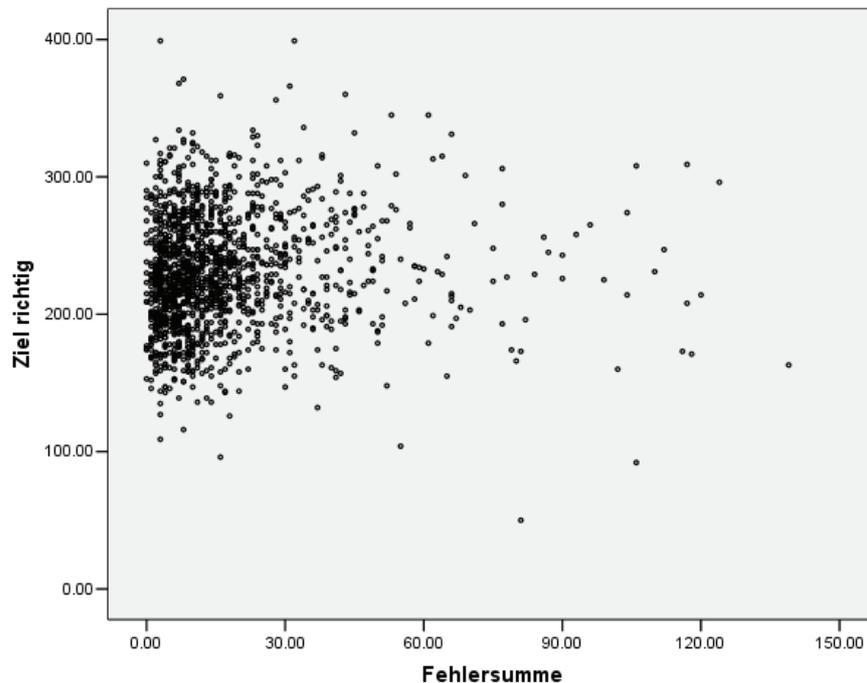


Abbildung 59: Scatterplot für alle Personen – Mengenleistung (Richtige) und Genauigkeit (niedrige Fehlerzahl).

8.5 Item-Trennschärfen

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

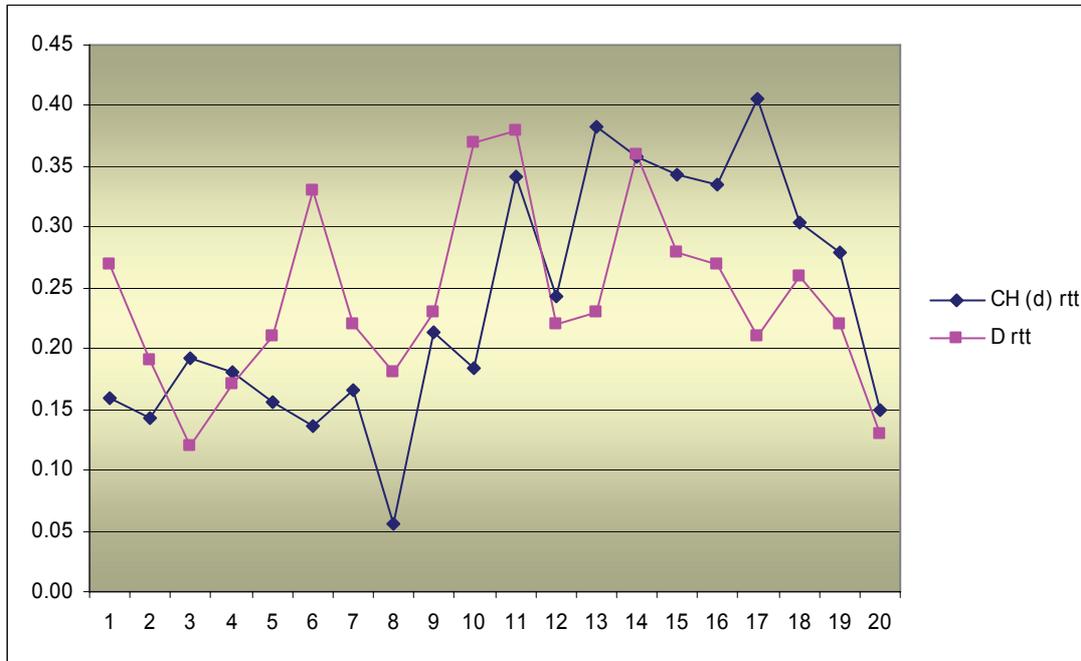


Abbildung 60: Trennschärfen für den Untertest „Muster zuordnen“.

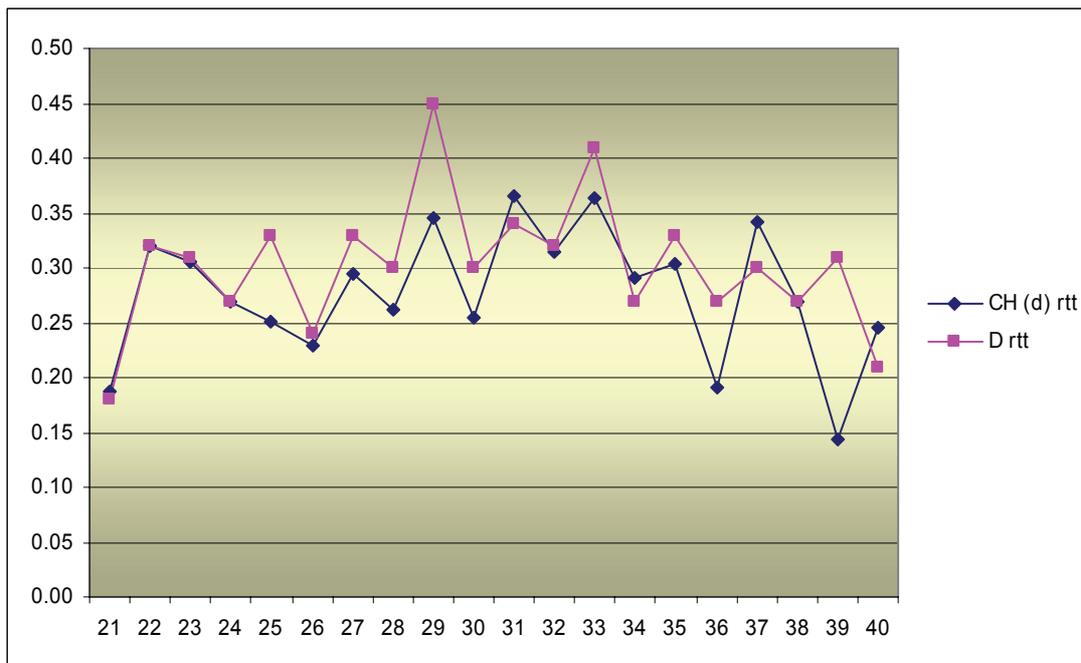


Abbildung 61: Trennschärfen für den Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“.

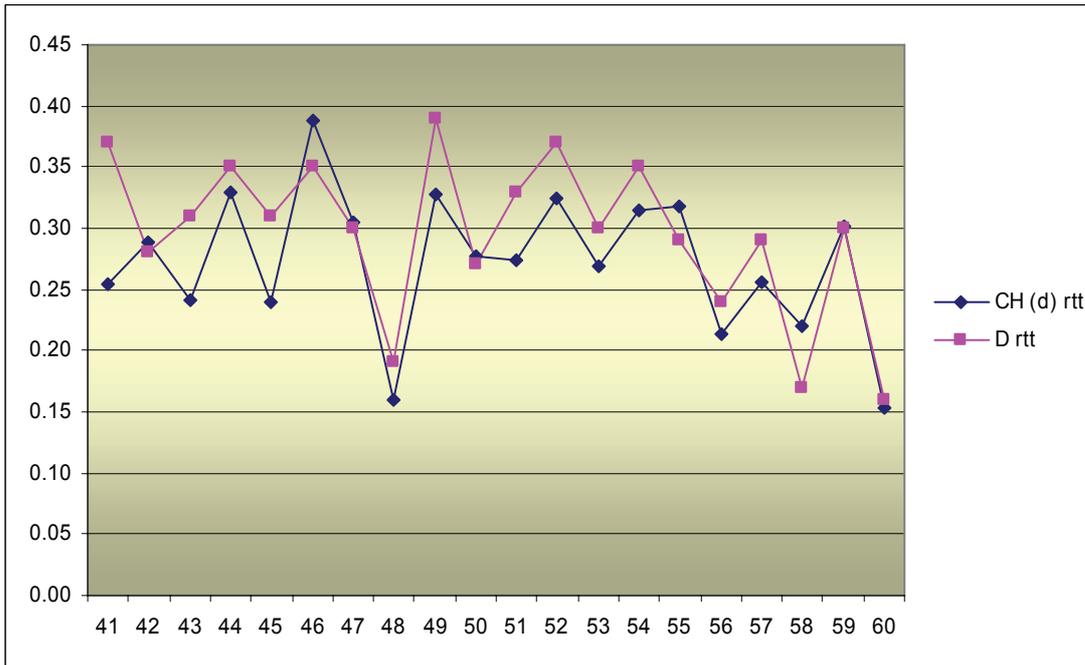


Abbildung 62: Trennschärpen für den Untertest „Schlauchfiguren“.

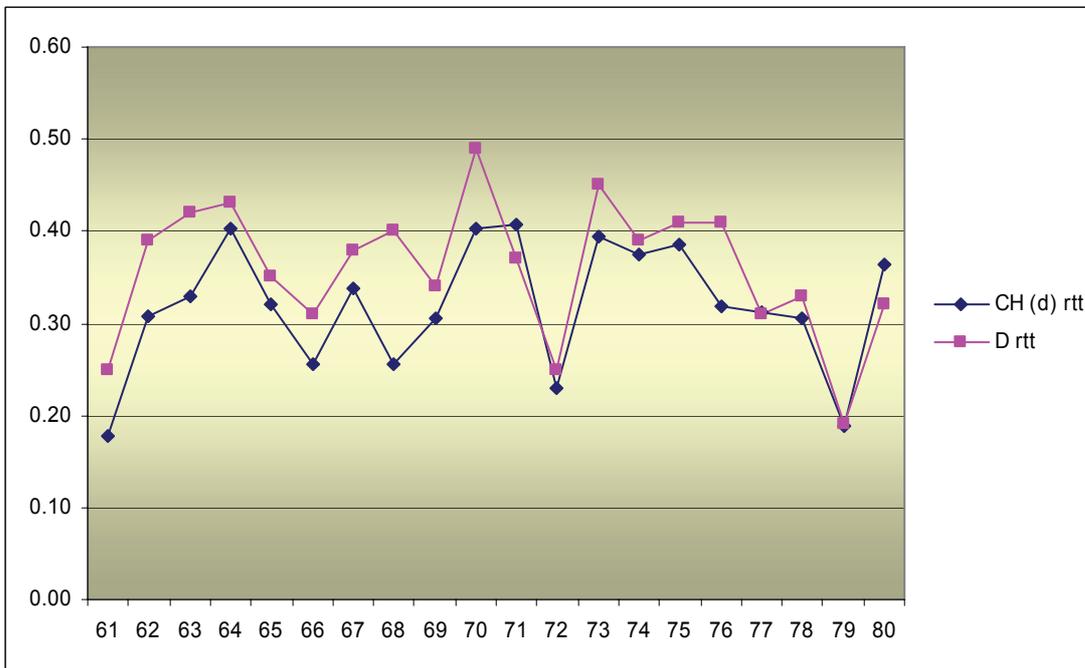


Abbildung 63: Trennschärpen für den Untertest „Quantitative und formale Probleme“.

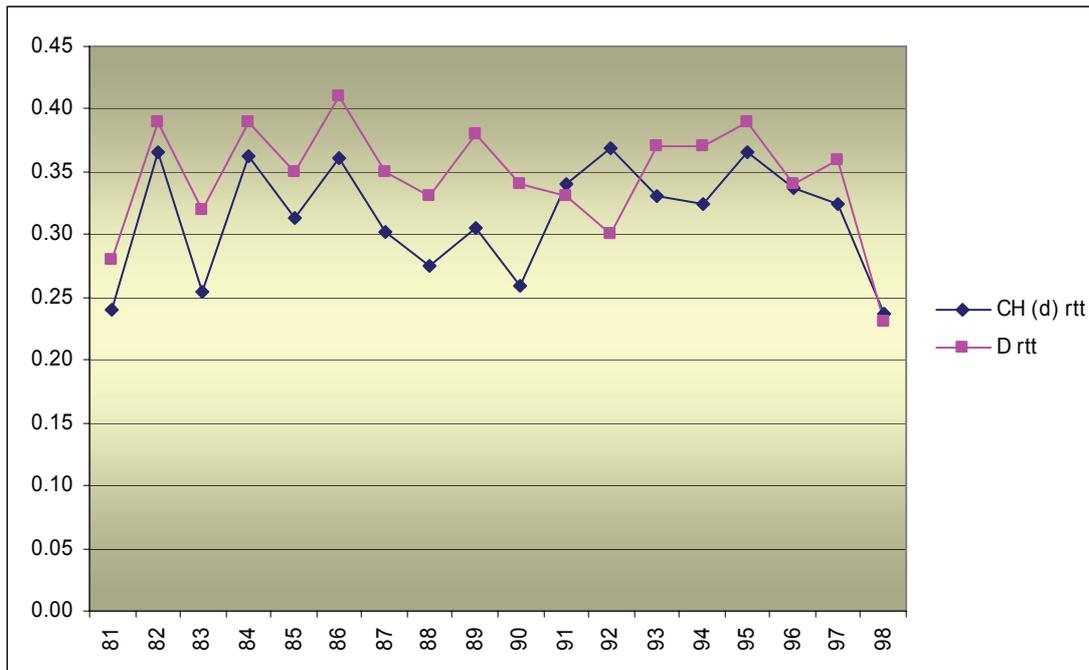


Abbildung 64: Trennschärfen für den Untertest „Textverständnis“.

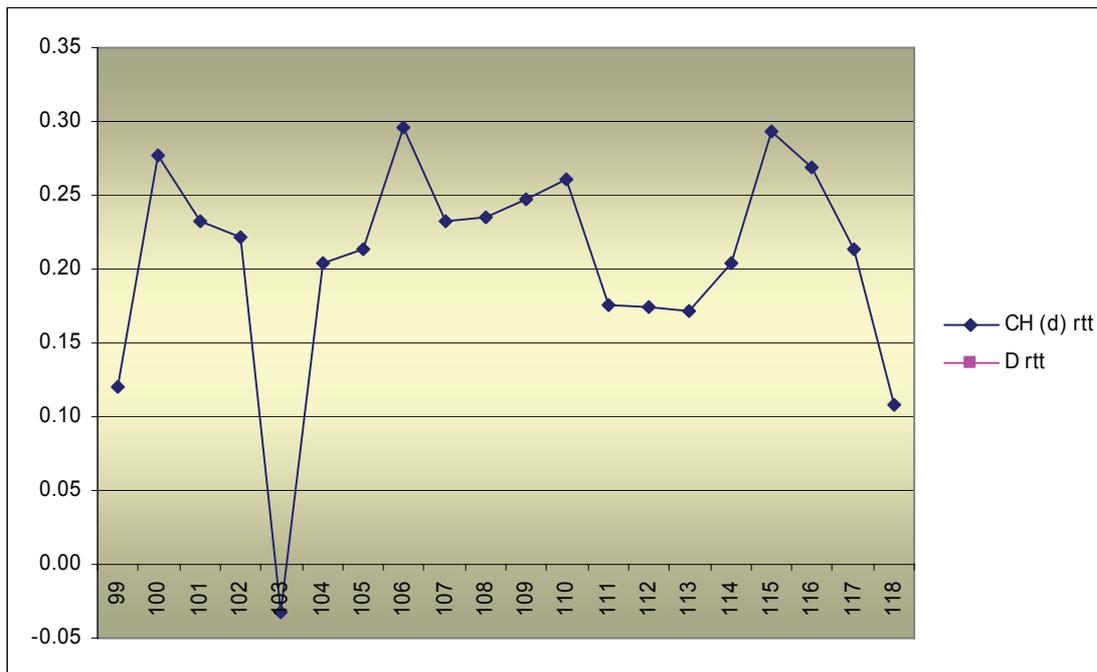


Abbildung 65: Trennschärfen für den Untertest „Planen und Organisieren“ (keine deutschen Vergleichswerte). Item 103 wurde für den Punktwert eliminiert.

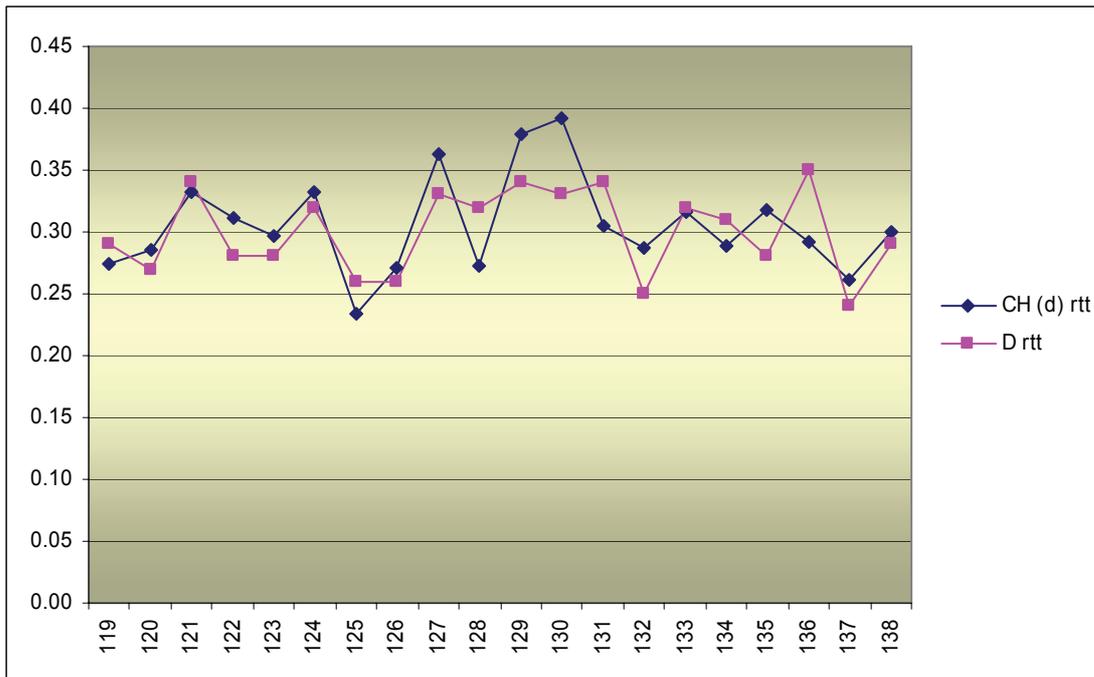


Abbildung 66: Trennschärfen für den Untertest „Figuren lernen“.

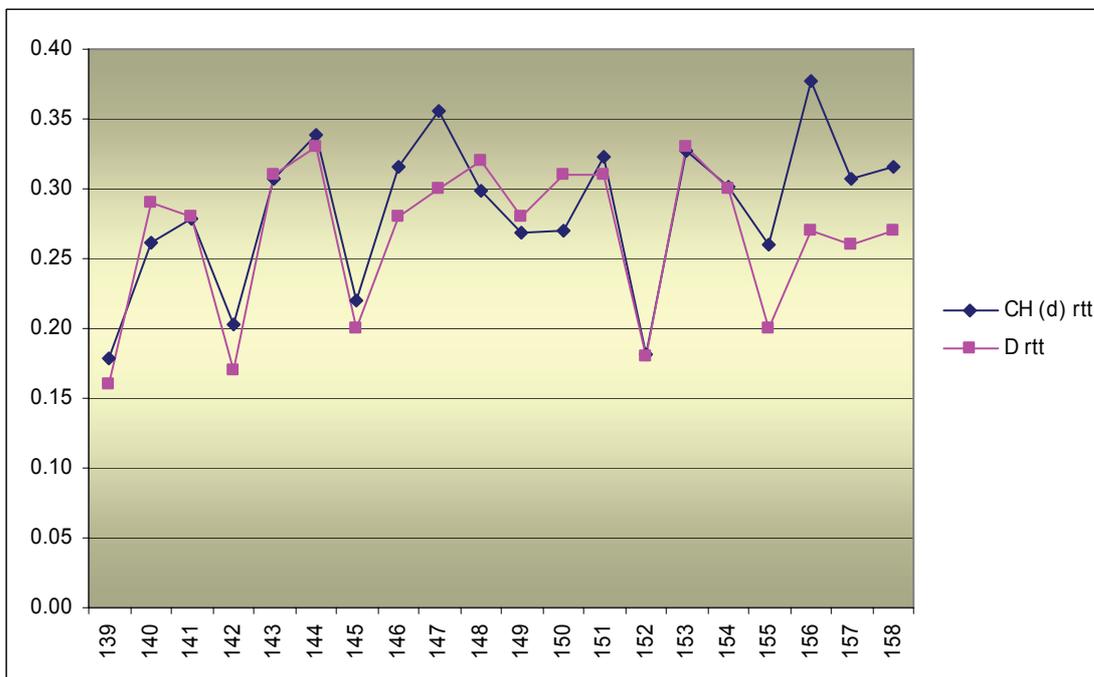


Abbildung 67: Trennschärfen für den Untertest „Fakten lernen“.

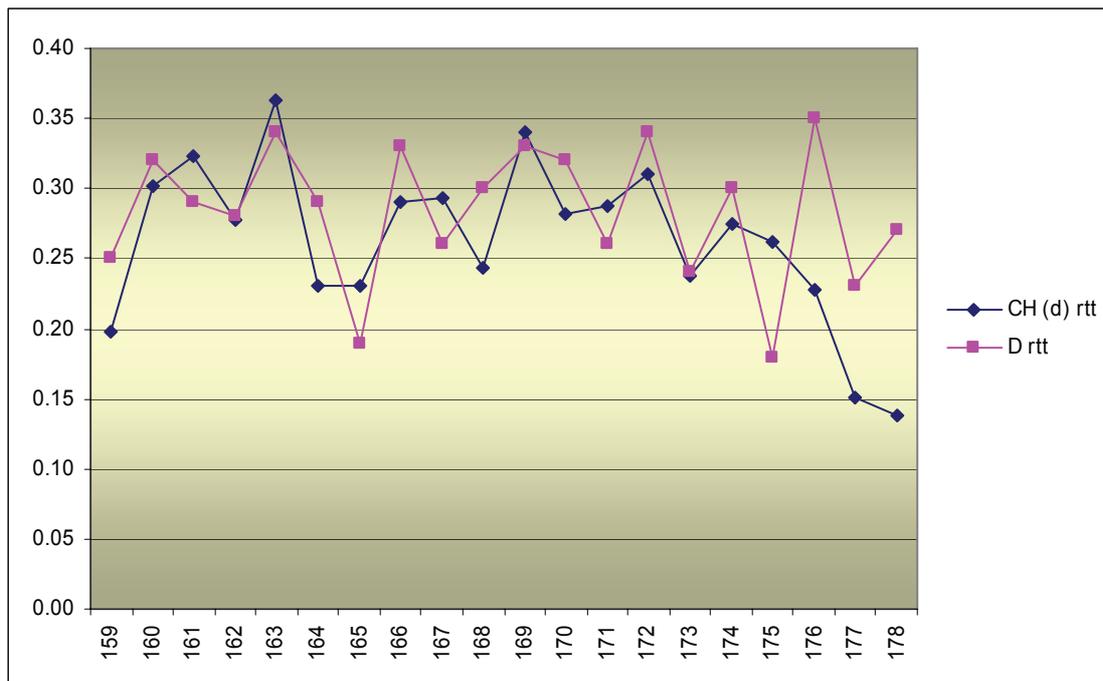


Abbildung 68: Trennschärpen für den Untertest „Diagramme und Tabellen“.

9 Zusammenfassung Kantone 1998 bis 2005

Seit 1998 haben bisher 8221 Personen am EMS teilgenommen. In der nachfolgenden Tabelle sind alle Teilnehmer am EMS 1998 bis 2005 nach Kantonen aufgeführt. Zu beachten ist, dass 1998 nur Humanmedizin, seit 1999 Human- und Veterinärmedizin und erst seit 2004 zusätzlich Zahnmedizin einem NC an den Universitäten Basel, Bern, Freiburg und Zürich unterlagen.

Wohnkanton/ Wohnort	Disziplin			Geschlecht		Sprache			Total
	Humanmedizin	Veterinärmedizin	Zahnmedizin	männlich	weiblich	deutsch	französisch	italienisch	
AG	512	96	24	219	413	632			632
AI	13	5	1	6	13	19			19
AR	46	8	6	23	37	60			60
BE	1050	206	58	452	862	1292	21	1	1314
BL	379	51	21	166	285	449	2		451
BS	319	30	10	131	228	358		1	359
FR	298	55	5	136	222	130	228		358
GE	4	51		13	42	2	53		55
GL	29	8	3	11	29	40			40
GR	226	49	11	101	185	274		12	286
JU	5	13		6	12	1	17		18
LU	380	64	7	153	298	451			451
NE	9	40		9	40	4	45		49
NW	36	7	2	12	33	45			45
OW	36		2	13	25	38			38
SG	441	76	16	189	344	533			533
SH	68	10	8	29	57	86			86
SO	239	46	11	120	176	295	1		296
SZ	127	16	3	45	101	146			146
TG	200	45	15	74	186	260			260
TI	213	53	11	117	160	5		272	277
UR	21	6	2	11	18	29			29
VD	14	121		41	94	6	129		135
VS	92	40	8	50	90	95	45		140
ZG	100	29	2	51	80	129	1	1	131
ZH	1568	232	66	650	1216	1849	6	11	1866
FL	34	4	4	12	30	42			42
Ausland	84	16	5	39	66	95	9	1	105
Total	6543	1377	301	2879	5342	7365	557	299	8221

Tabelle 58: Gesamtzahl der Teilnehmer am EMS 1998 bis 2005. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

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

	Humanmedizin					Veterinärmedizin					Zahnmedizin				
	n	m	s	Min	Max	n	m	s	Min	Max	n	m	s	Min	Max
AG	512	102.6	9.5	74	129	96	100.1	9.4	75	124	24	97.0	9.0	78	111
AI	13	106.1	9.8	88	123	5	100.6	7.1	95	111	1	102.0	.	102	102
AR	46	99.9	9.1	81	127	8	102.5	9.7	91	118	6	89.8	11.7	70	105
BE	1050	100.4	10.2	72	130	206	97.5	9.0	76	128	58	96.4	10.0	70	116
BL	379	100.4	10.4	70	125	51	99.6	9.1	74	118	21	98.8	9.2	80	115
BS	319	97.0	10.0	71	122	30	98.1	9.9	76	117	10	93.7	7.2	85	104
FR	298	96.1	9.3	70	124	55	92.7	8.1	73	108	5	93.4	12.9	78	110
GE	4	100.3	10.8	85	110	51	92.5	8.2	76	117
GL	29	102.1	9.2	90	130	8	94.1	13.1	70	112	3	85.0	3.6	81	88
GR	226	99.6	10.0	76	128	49	97.2	8.0	77	113	11	97.1	10.1	80	115
JU	5	90.0	4.2	85	96	13	93.8	7.5	85	109
LU	380	101.1	9.7	71	127	64	97.9	8.9	76	118	7	98.0	6.9	89	110
NE	9	104.6	9.4	91	116	40	92.1	8.1	76	106
NW	36	101.4	10.8	78	121	7	99.3	10.2	88	113	2	109.5	.7	109	110
OW	36	101.1	8.7	89	126	2	96.0	4.2	93	99
SG	441	101.7	9.9	70	128	76	99.9	7.8	79	118	16	99.4	11.5	83	119
SH	68	100.8	9.3	82	123	10	97.5	10.0	86	114	8	95.3	10.2	77	109
SO	239	100.9	9.6	70	127	46	97.9	8.9	79	119	11	98.6	12.5	78	120
SZ	127	98.5	9.7	72	127	16	99.8	10.2	80	121	3	89.3	13.0	76	102
TG	200	101.6	9.9	77	123	45	101.9	9.1	82	122	15	94.9	8.8	80	107
TI	213	97.4	9.3	73	120	53	93.8	7.1	82	112	11	90.8	11.3	70	109
UR	21	101.4	7.5	87	120	6	100.2	9.0	91	112	2	92.5	10.6	85	100
VD	14	101.7	9.2	87	119	121	93.4	7.9	73	115
VS	92	98.4	10.0	70	126	40	99.3	10.4	75	120	8	93.0	14.0	75	116
ZG	100	102.3	10.3	73	128	29	98.0	7.1	84	109	2	98.5	2.1	97	100
ZH	1568	102.3	10.4	66	131	232	100.1	10.1	76	126	66	98.5	11.2	70	123
FL	34	98.6	9.9	77	120	4	105.0	12.7	92	122	4	99.8	5.9	94	107
Aus-land	84	98.6	9.4	81	119	16	90.9	10.5	75	107	5	97.2	8.6	83	106
Total	6543	100.7	10.1	66	131	1377	97.5	9.4	70	128	301	96.7	10.3	70	123

Tabelle 59: Kennwerte für den Testwert (n = Bewerberzahl, m = Mittelwert, s = Standardabweichung) aller Teilnehmer am EMS nach Kantonen und Disziplinen. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Nachfolgend werden wiederum die mittleren Punktwerte der einzelnen Untertests für alle Teilnehmenden am EMS aufgeführt. Für „Planen und Organisieren“ reicht die Datenbasis für eine kantonale Analyse noch nicht aus. Diese Punktwerte sind nur innerhalb eines Jahres numerisch vergleichbar – durch die Mittelung über alle Jahre sind dennoch Aussagen möglich, aus welchen Teilergebnissen die Testwertunterschiede pro Kanton resultieren.

WK	Muster zuordnen		Med.-nat. Grundverst.		Schlauchfiguren		Quant. formale Probl.		Textverständnis		Figuren lernen		Fakten lernen		Diagr. u. Tabellen		Konz. u. sorgfält. Arbeit.	
	m	s	m	s	m	s	m	s	m	s	m	s	m	s	m	s	m	s
AG	11.4	3.0	11.4	3.5	12.8	3.5	11.7	3.8	9.8	3.7	12.9	3.9	11.9	3.7	11.1	3.3	12.7	3.6
AI	11.9	3.5	11.3	2.7	12.3	3.6	13.3	2.7	8.5	2.8	14.4	4.4	12.3	3.7	12.1	2.9	13.3	3.5
AR	10.9	3.0	10.5	3.9	12.9	4.0	10.9	3.9	8.9	3.5	12.2	4.0	10.0	3.9	10.8	3.4	12.6	4.1
BE	11.1	3.1	10.8	3.6	12.6	3.7	10.5	3.9	9.2	3.5	12.8	3.9	11.3	3.8	10.6	3.4	12.0	3.8
BL	11.5	3.2	10.8	3.5	12.8	3.7	10.3	3.9	9.1	3.4	13.0	4.0	11.6	4.1	10.7	3.3	12.6	3.9
BS	11.0	3.1	9.9	3.7	12.4	3.5	9.4	3.7	8.1	3.6	12.4	4.1	10.9	3.7	9.9	3.4	12.1	4.0
FR	10.4	3.2	9.3	3.3	11.1	3.4	10.5	3.9	8.2	3.3	11.4	3.9	10.5	3.5	9.5	3.1	10.8	3.5
GE	9.7	2.9	8.1	3.3	11.7	3.4	9.7	3.7	6.8	3.1	11.1	4.0	9.0	3.2	8.5	3.1	10.7	3.8
GL	11.2	2.9	11.0	3.4	12.3	4.1	10.4	3.7	9.5	3.8	12.2	3.8	10.9	3.3	10.9	3.6	12.6	4.5
GR	11.0	3.1	10.3	3.5	12.1	3.5	10.7	3.6	8.7	3.6	12.7	3.9	11.4	3.7	10.4	3.2	12.7	3.9
JU	9.7	2.7	7.8	3.7	9.4	3.1	9.5	3.2	6.7	2.2	12.1	2.9	10.4	2.8	9.2	3.0	10.6	3.5
LU	11.3	3.3	11.0	3.3	12.4	3.4	11.3	3.9	9.3	3.6	13.0	3.9	11.4	3.6	11.0	3.3	12.4	3.6
NE	11.1	3.5	9.1	3.6	11.0	3.9	9.2	4.1	6.8	3.3	11.8	4.2	10.8	3.5	8.7	3.1	10.4	3.5
NW	11.6	3.4	11.2	3.3	12.2	4.2	11.4	3.6	9.4	3.3	14.1	3.9	12.7	3.4	10.9	3.5	12.1	4.1
OW	11.8	3.0	10.7	3.4	12.6	2.9	10.5	4.2	9.0	3.4	12.8	3.5	11.6	3.5	10.5	3.4	13.0	3.5
SG	11.7	3.1	10.8	3.3	12.9	3.7	11.3	3.7	9.2	3.5	13.0	4.0	11.7	3.7	11.0	3.3	12.9	3.8
SH	11.2	2.8	10.3	3.3	12.2	3.5	11.0	3.6	9.3	3.6	12.0	3.9	11.2	3.9	10.7	3.4	13.1	3.8
SO	11.5	3.1	10.6	3.3	12.8	3.7	10.6	3.6	8.8	3.5	13.3	3.8	11.2	3.8	10.7	3.2	12.9	4.1
SZ	11.4	3.4	9.9	3.8	12.0	3.8	10.5	3.8	8.8	3.8	12.7	3.8	10.8	3.6	10.2	3.2	12.6	3.8
TG	11.6	3.1	10.7	3.6	13.0	3.7	11.0	3.9	9.2	3.5	13.3	4.0	11.7	3.6	10.8	3.3	13.0	4.0
TI	10.6	2.9	9.5	3.4	10.9	3.2	10.7	3.9	8.4	3.4	11.3	3.6	10.6	3.6	9.2	3.2	11.5	3.1
UR	11.2	2.5	10.4	3.9	13.5	3.1	10.8	3.9	8.7	3.6	12.6	4.1	12.1	3.7	11.3	3.0	12.9	3.2
VD	10.4	2.9	9.0	2.9	10.7	3.2	8.9	4.0	8.0	2.9	11.1	3.5	10.3	3.5	9.1	3.1	10.7	3.1
VS	10.8	2.9	10.0	3.5	11.8	4.0	11.0	3.9	9.0	3.3	12.1	4.1	11.1	3.9	10.2	3.2	11.7	3.5
ZG	11.7	3.1	10.7	3.4	13.1	3.6	10.9	3.7	9.2	3.7	13.1	4.0	11.2	3.6	10.9	3.6	12.7	4.0
ZH	11.6	3.1	11.1	3.5	13.0	3.8	11.3	3.9	9.6	3.7	13.0	4.1	12.0	3.9	10.9	3.4	12.8	4.2
FL	11.8	3.2	10.3	3.1	12.1	3.7	9.6	3.4	8.5	3.5	12.3	4.1	11.3	3.8	10.3	3.5	13.6	3.9
Ausland	10.7	3.3	10.0	3.9	12.2	3.6	9.5	4.1	8.5	3.5	12.0	4.1	11.3	3.2	9.9	3.6	12.1	4.0
Total	11.3	3.1	10.6	3.5	12.5	3.7	10.8	3.9	9.1	3.6	12.7	4.0	11.4	3.8	10.6	3.4	12.4	3.9

Tabelle 60: Mittelwerte (m) und Standardabweichungen (s) für die Punktwerte der Untertests nach Wohnkantonen (WK) für alle EMS-Teilnehmenden.

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

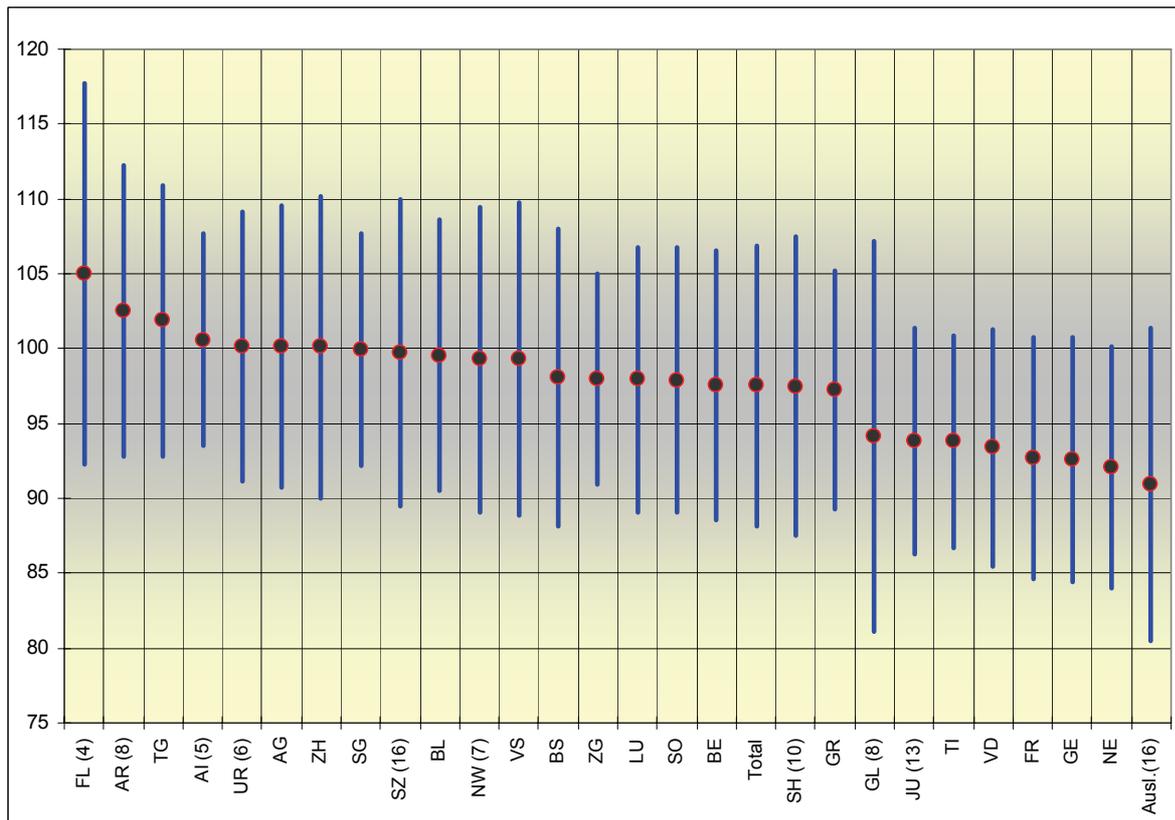


Abbildung 69: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1998 bis 2005 zusammengefasst; **Humanmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

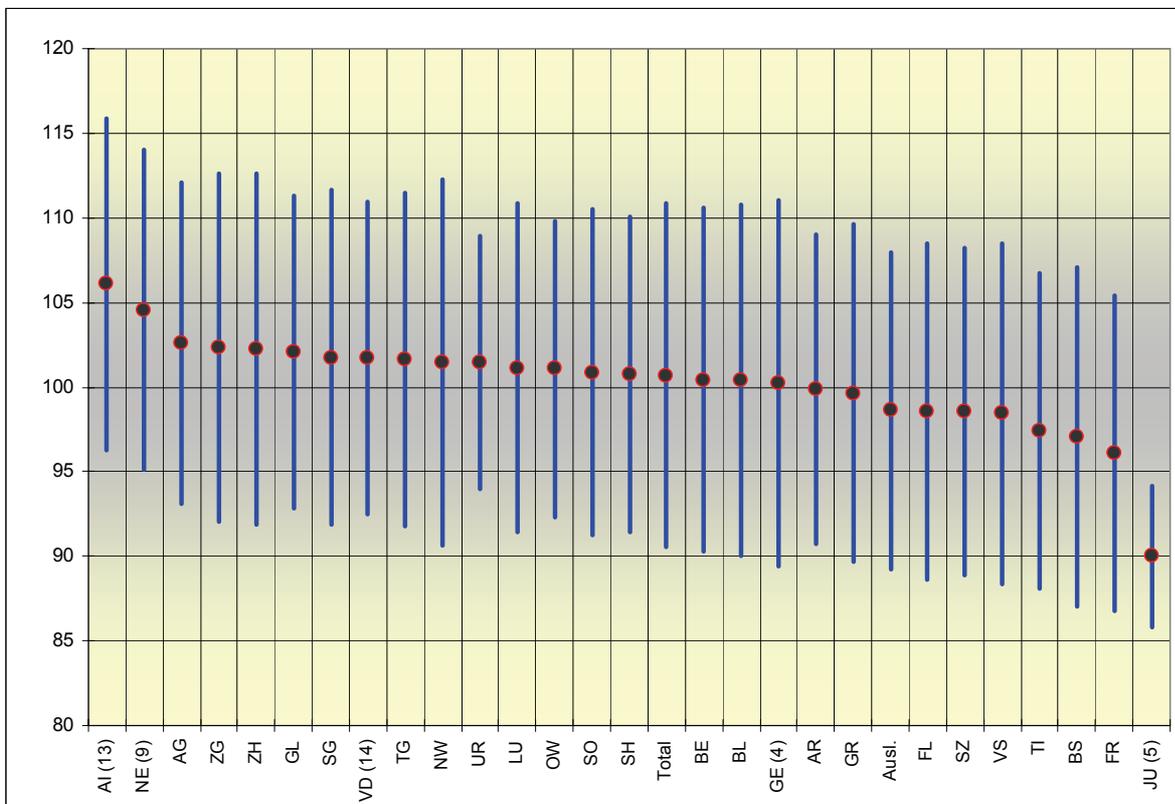


Abbildung 70: Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1999 bis 2005 zusammengefasst; **Veterinärmedizin**. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

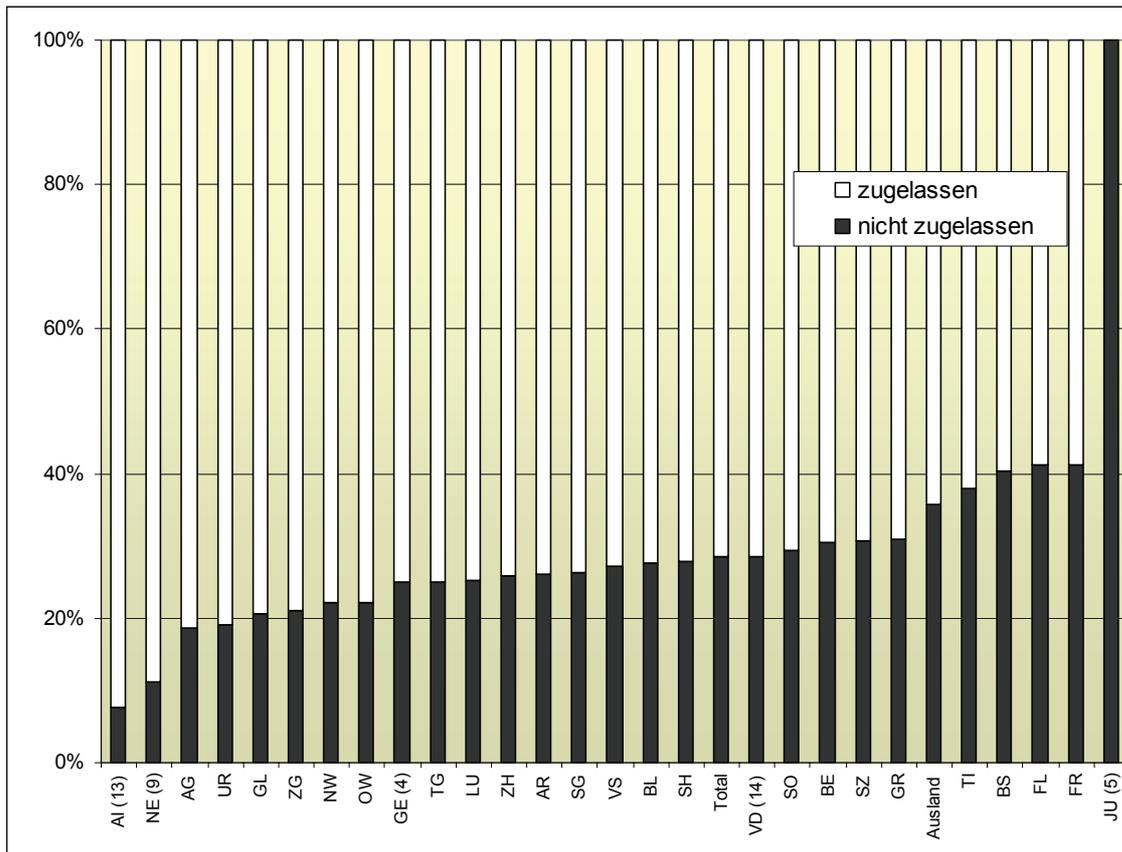


Abbildung 71: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1998 bis 2005; **Humanmedizin**.

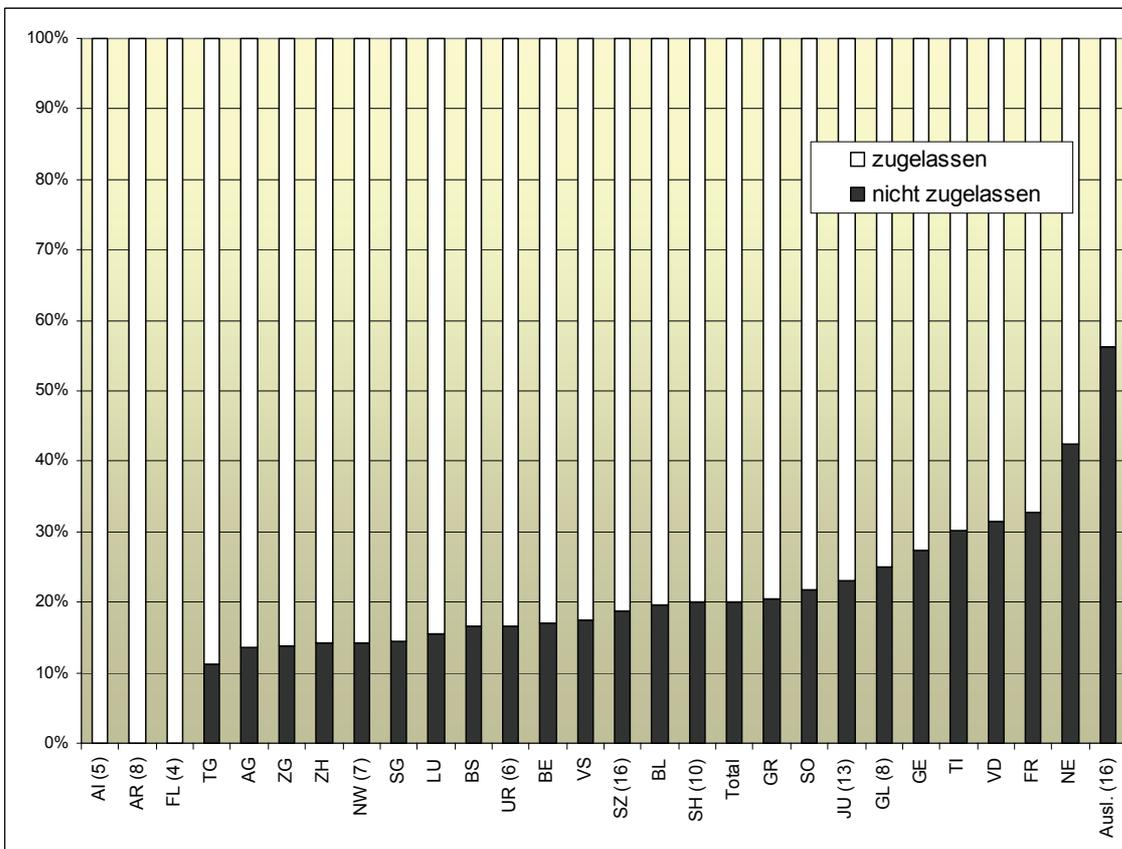


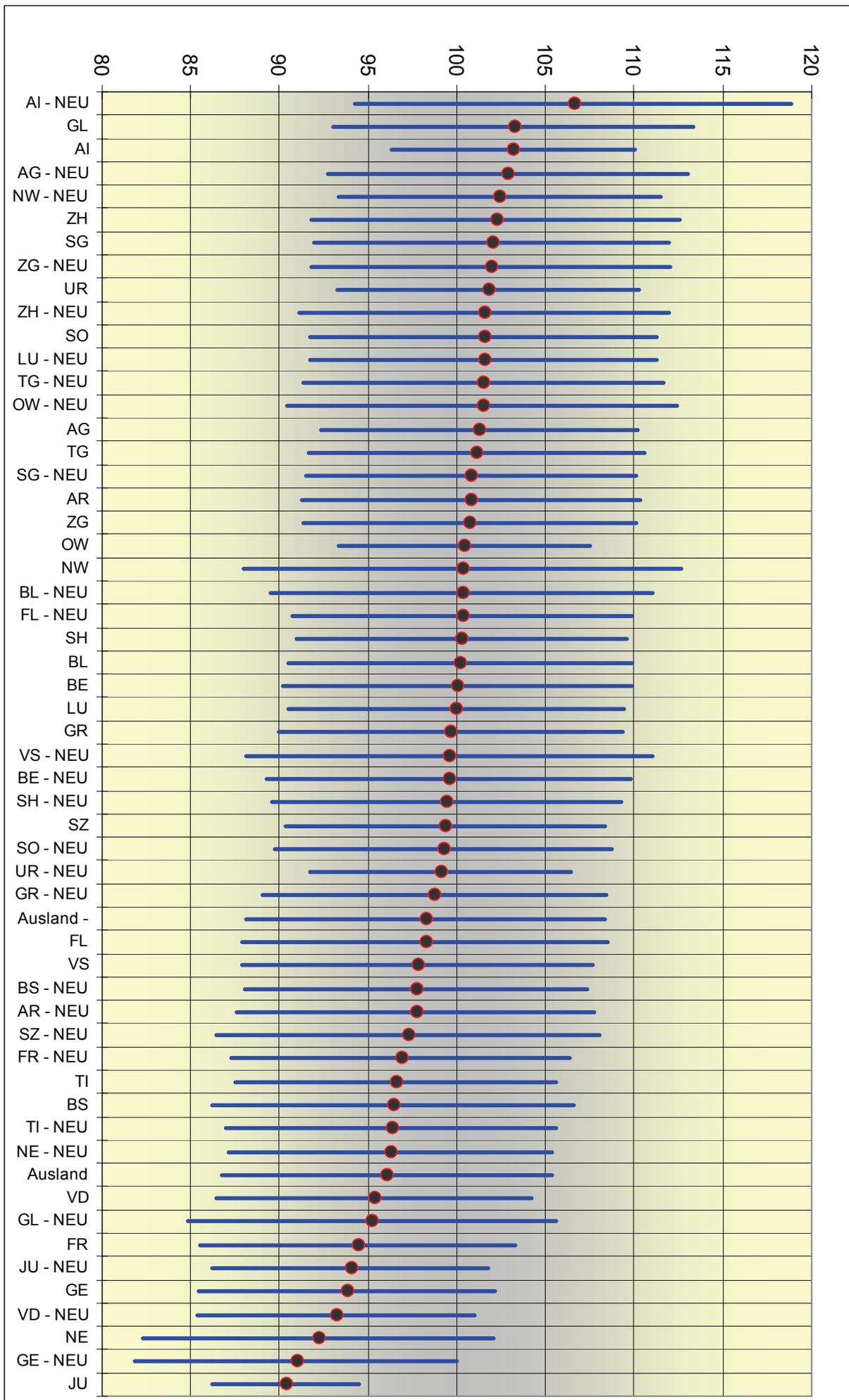
Abbildung 72: Mehrjährige theoretische Zulassungsquoten auf der Basis der Studienplatzvergabe nach dem Testtermin (ohne Nachrücker) 1999 bis 2005; **Veterinärmedizin**.

Mittlerweile ist es möglich, bezüglich der Gesamt-Teilnehmerzahl am EMS auch zeitliche Trends miteinander zu vergleichen. Dazu wurden die Ergebnisse 1998 bis 2002 denen von 2003 bis 2005 gegenübergestellt. 4259 Teilnehmer im ersten Zeitraum werden 3962 Teilnehmern im späteren Zeitraum gegenübergestellt. Bei allen französischsprachigen oder zweisprachigen Kantonen ist hierbei erneut zu beachten, dass den Kandidaten ein Studium in Lausanne oder Genf möglich ist – die Entscheidungen für eine Bewerbung mit NC vielfältig sein können.

	1998 bis 2002			2003 bis 2005			Total		
	n	m	s	n	m	s	n	m	s
AG	331	101.3	8.9	301	102.9	10.2	632	102.0	9.6
AI	12	103.2	6.8	7	106.6	12.3	19	104.4	9.0
AR	30	100.8	9.5	30	97.7	10.1	60	99.2	9.9
BE	684	100.0	9.9	630	99.5	10.3	1314	99.8	10.1
BL	238	100.2	9.7	213	100.3	10.8	451	100.2	10.2
BS	184	96.4	10.2	175	97.7	9.7	359	97.0	9.9
FR	190	94.4	8.9	168	96.8	9.5	358	95.6	9.3
GE	42	93.8	8.3	13	90.9	9.1	55	93.1	8.5
GL	20	103.2	10.1	20	95.2	10.4	40	99.2	10.9
GR	125	99.6	9.7	161	98.7	9.7	286	99.1	9.7
JU	6	90.3	4.1	12	94.0	7.8	18	92.8	6.9
LU	250	99.9	9.5	201	101.5	9.8	451	100.6	9.6
NE	23	92.2	9.9	26	96.2	9.1	49	94.3	9.6
NW	20	100.3	12.3	25	102.4	9.1	45	101.5	10.6
OW	24	100.4	7.1	14	101.4	11.0	38	100.8	8.6
SG	287	102.0	10.0	246	100.8	9.3	533	101.4	9.7
SH	47	100.2	9.3	39	99.4	9.9	86	99.9	9.5
SO	142	101.5	9.8	154	99.2	9.5	296	100.3	9.7
SZ	87	99.3	9.0	59	97.2	10.8	146	98.5	9.8
TG	137	101.1	9.5	123	101.5	10.1	260	101.3	9.8
TI	151	96.5	9.1	126	96.3	9.3	277	96.4	9.2
UR	16	101.8	8.5	13	99.1	7.3	29	100.6	8.0
VD	68	95.3	8.9	67	93.2	7.9	135	94.3	8.4
VS	94	97.8	9.9	46	99.6	11.4	140	98.4	10.4
ZG	67	100.7	9.4	64	101.9	10.1	131	101.3	9.7
ZH	923	102.2	10.4	943	101.5	10.5	1866	101.9	10.4
FL	20	98.2	10.3	22	100.3	9.6	42	99.3	9.9
Ausl.	41	96.0	9.3	64	98.2	10.1	105	97.4	9.8

Tabelle 61: Mittelwerte (m) und Standardabweichungen (s) für die Testwerte nach Kantonen zusammengefasst für die Jahre bis 2002 und ab 2003; n ist die Personenzahl.

Abbildung 73 (folgende Seite): Mittelwerte und Streubereiche ($m \pm s$) der Testwerte für Kantone 1998 bis 2003 und 2004 bis 2005 zusammengefasst; die Ergebnisse 2004/05 sind mit neu gekennzeichnet. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.



10 Beispielaufgaben für die Untertests

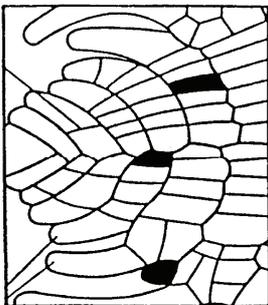
Nachfolgend wird pro Untertest eine Beispielaufgabe dargestellt. So können lediglich die Prinzipien der Aufgabenstruktur verdeutlicht werden – die Aufgaben unterscheiden sich innerhalb jedes Untertests bezüglich des Schwierigkeitsgrades und der Anforderung.

10.1 Untertest: Muster zuordnen

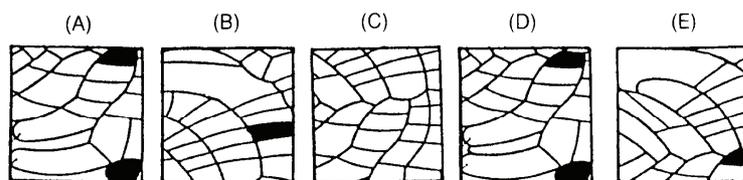
In diesem Untertest wird die Fähigkeit geprüft, Ausschnitte in einem komplexen Bild wieder zu erkennen. Dazu werden pro Aufgabe ein Muster und je fünf Musterausschnitte (A) bis (E) vorgegeben. Die Testteilnehmerin oder der Testteilnehmer soll herausfinden, welcher dieser fünf Musterausschnitte an irgendeiner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

10.2 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 kor-

respondierende 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.

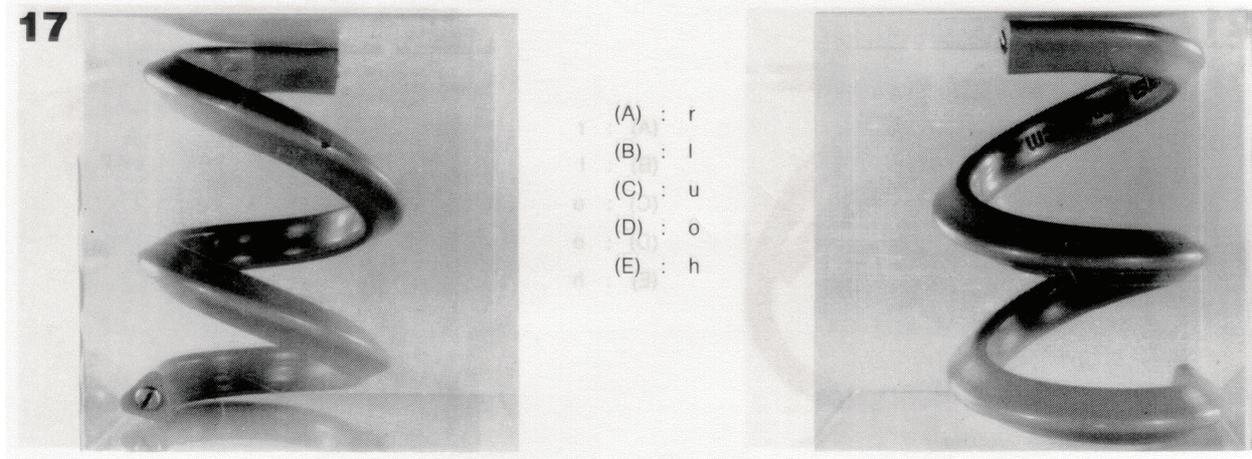
Nach dieser Analyse des Textes ist es offensichtlich, dass die Antwort (C) richtig ist.

10.3 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!)

10.4 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.

10.5 Untertest: Textverständnis

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

Ein Beispiel:

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

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

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

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

Diese Freisetzung von T_3 und T_4 ins Blut (Sekretion) wird über die Hirnanhangsdrüse und den Hypothalamus, einen Teil des Zwischenhirns, gesteuert: Das erwähnte Hormon TSH stimuliert ausser der Bildung auch die Sekretion von T_3 und T_4 ; es ist hinsichtlich seiner eigenen Sekretionsrate jedoch abhängig von der Stimulation durch das hypothalamische Hormon TRH. Die TRH-Sekretion wiederum wird z.B. durch Kälte stimuliert, während Wärme hemmend wirken kann. Neben diesen übergeordneten Steuerungsmechanismen existiert noch ein sogenannter Rückkoppelungsmechanismus: 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 T₃ 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 T₃ führen.

10.6 Untertest: Konzentriertes und Sorgfältiges Arbeiten

Bei diesem Untertest soll die Fähigkeit, rasch, sorgfältig und konzentriert zu arbeiten, gemessen werden. Es werden Zeichen vorgegeben und bestimmte Zeichen sind zu markieren. Dies kann ein Merkmal eines einzelnen Zeichens sein (ein b mit 2 Strichen) oder eine Zeichenfolge (wenn ein p auf ein q folgt). Bei diesem Test hatte sich gezeigt, dass exzessives Üben zu verbesserten Leistungen führt. Da ein „Fleisstest“ nicht intendiert ist, werden seit 2004 die Zeichen und die Regel vor dem Test nicht mehr bekannt gegeben. Übungseffekte haben dadurch nachweislich einen geringen Einfluss.

Zu Vorbereitungszwecken wird der ursprüngliche Untertest empfohlen, um ein Gefühl für die Art der Anforderung und die Zeitdauer zu erhalten. Dabei sollen möglichst alle b, die mit zwei Querstreichen versehen sind, die entweder beide unten, beide oben oder je einer unten und oben angebracht sind, markiert werden. Die Lösungsmenge ist ebenso wichtig wie die Fehlerfreiheit der Bearbeitung. Dieser Test ist trainierbar – in der Test-Info wird darauf hingewiesen, diesen Untertest vor der Testabnahme mehrfach zu üben.

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

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

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

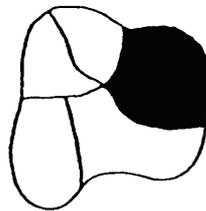
10.7 Untertest: Figuren lernen

Für beide nachfolgenden 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.

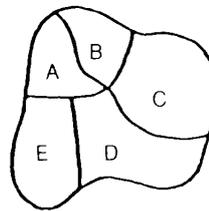
Der Untertest „Figuren lernen“ prüft, wie gut man sich Einzelheiten von Gegenständen einprägen und merken kann.

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.

10.8 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

10.9 Untertest: Planen und Organisieren

Der Untertest „Planen und Organisieren“ ist erstmals Bestandteil des EMS und prüft Fähigkeiten, die für eine effiziente Selbstorganisation im Studium wichtig sind. Er umfasst verschiedene von einander unabhängige Szenarien, auf die sich jeweils mehrere Aufgaben beziehen.

Auch bei diesem Untertest wird kein Wissen vorausgesetzt. Es werden Aufgabenstellungen in Form so genannter Szenarien vorgegeben. Zu diesen Szenarien werden dann genau wie bei anderen Untertests einige Fragen gestellt. Bitte beachten Sie, dass im späteren „richtigen“ EMS sehr vielfältige Szenarien zur Anwendung kommen, die auch ganz anderen Schemen folgen können. Das nachfolgende Szenario soll das Prinzip verdeutlichen.

Szenario: Semesterarbeit

Sie haben die Aufgabe, im kommenden Wintersemester eine Semesterarbeit zu einem vorgegebenen Thema zu schreiben. Nach gründlichem Studium der einschlägigen Literatur sollen Sie in Ihrer Arbeit zunächst einen Überblick über die wichtigsten vorliegenden Ergebnisse zum Thema geben und dann einen eigenen Lösungsansatz entwickeln.

Die folgenden Termine liegen bereits fest:

- 11.10., 17.00 Uhr: Sie erhalten das Thema Ihrer Semesterarbeit, Literaturhinweise und weitere Instruktionen.
- 17.10.: Beginn der Vorlesungszeit
- 19.12. bis 1.1.: Weihnachtsferien (vorlesungsfreie Zeit)
- 11.2.: Ende der Vorlesungszeit
- 7.3., 8.00 Uhr: Abgabe der Semesterarbeit
- Während der Vorlesungswochen: Jede Woche von Montag bis Donnerstag Lehrveranstaltungen; an diesen Tagen können Sie sich nicht der Semesterarbeit widmen.
- In der ersten Hälfte der Weihnachtsferien (16. bis 24.12.): Urlaubsreise
- Zwei volle Wochen im unmittelbaren Anschluss an die Vorlesungszeit: Prüfungsvorbereitung (1 Woche) und Teilnahme an mündlichen Prüfungen (1 Woche)

Die folgenden Arbeitsschritte bzw. -phasen müssen Sie einplanen (sie sind hier nicht in chronologischer Abfolge aufgeführt):

- Zwei Treffen mit dem Betreuer Ihrer Arbeit: Eines unmittelbar vor dem Schreiben der Semesterarbeit, eines vor der abschliessenden Überarbeitung. Dauer: Je ein voller Arbeitstag. (Der Betreuer ist nur dienstags und freitags verfügbar, und zwar von der dritten bis zur neunten Vorlesungswoche sowie nach dem Ende der Vorlesungszeit.)
- Literatursuche und -beschaffung (Dauer: 5 Tage)
- Erstellung des Konzepts der Semesterarbeit; die Literatur muss hier bereits verarbeitet sein (Dauer: 3 Tage)
- Korrekturlesen durch einen Freund (Überprüfung auf Tippfehler, unklare Formulierungen etc.) (Dauer: 3 Tage)
- Lesen und Zusammenfassen der Literatur (Dauer: 10 Tage)
- Schreiben der Semesterarbeit (Dauer: 18 Tage)
- Abschliessendes Überarbeiten der Semesterarbeit (Dauer: 4 Tage)
- Kopieren der Arbeit (unmittelbar vor der Abgabe) (Dauer: 1 Tag)
- Ausführen der Korrekturvorschläge Ihres Freundes, bevor Sie die Arbeit mit dem Betreuer durchsprechen (Dauer: 1 Tag)
- Reserve für Unvorhergesehenes (Dauer: 1 Tag)

Eine Arbeitsphase muss stets beendet sein, bevor die nächste beginnt (Ausnahme: Korrekturlesen).

Ferner müssen Sie einplanen:

- An den Wochenenden arbeiten Sie – wie auch Ihr Freund – nur samstags (im Urlaub natürlich gar nicht).
- Am zweiten Weihnachtstag arbeiten Sie nicht.

Welche der folgenden Aussagen über die Semesterarbeit trifft bzw. treffen zu?

- I. Noch vor den Weihnachtsferien können Sie mit dem Schreiben der Semesterarbeit beginnen.
- II. Insgesamt werden Sie sich in 18 Wochen jeweils mindestens einen Tag lang mit der Semesterarbeit beschäftigen.

- (A) Nur Aussage I trifft zu.
(B) Nur Aussage II trifft zu.
(C) Beide Aussagen treffen zu.
(D) Keine der beiden Aussagen trifft zu.

Welche der folgenden Aussagen über die Einbeziehung Ihres Freundes trifft bzw. treffen zu?

- I. Der günstigste Zeitraum, in dem Ihr Freund Ihre Arbeit Korrektur lesen kann, ist die erste Hälfte der letzten Vorlesungswoche.
- II. Angenommen, Ihr Freund teilt Ihnen mit, er habe leider nur in der Zeit vom 25. bis zum 27. Februar Gelegenheit zum Korrekturlesen; dann nützt Ihnen sein Angebot für den ursprünglich geplanten Zweck dieses Korrekturvorgangs nicht mehr.

- (A) Nur Aussage I trifft zu.
(B) Nur Aussage II trifft zu.
(C) Beide Aussagen treffen zu.
(D) Keine der beiden Aussagen trifft zu.

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

11 Literatur

- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In Holland, P.W., Wainer, H. (Eds.), Differential Item Functioning. Hillsdale, NJ: Erlbaum.
- Beller, M., Gafni, N. (1995). Translated Versions of Israel's interuniversity Psychometric Entrance Test (PET). In T. Oakland & R.K. Hambleton (Eds.), International Perspectives of Academic Assessment, S.207-218. Boston: Kluwer.
- Beller, M. (1996). Translating, equating and validating Scholastic Aptitude Tests: The Israeli Case. In Hänsgen, K.-D., Hofer, R. & Ruefli, D. (Hrsg). Eignungsdiagnostik und Medizinstudium, (S. 14-29), Bericht 2. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.
- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (Second Edition). Hillsdale, NJ: Erlbaum.
- Cook, L.L. (1998). Can Scores Obtained on Test Given in Different Languages to Examinees of Different Cultures be Equally Valid? ICAP San Francisco.
- Deidesheimer Kreis (1997). Hochschulzulassung und Studieneignungstests: studienfeldbezogene Verfahren zur Feststellung der Eignung für Numerus Clausus und andere Studiengänge. Göttingen, Zürich: Vandenhoeck und Ruprecht.
- Ebach, J., Trost, G. (1997). Admission to Medical Schools in Europe. Lengerich: Pabst.
- Hänsgen K-D, Spicher B. (2002). Numerus Clausus: Finden wir mit dem «Eignungstest für das Medizinstudium» die Geeigneten? Schweizerische Ärztezeitung / Bulletin des médecins suisses / Bollettino dei medici svizzeri - 2002; 83(31):1653-1660.
<http://www.saez.ch/pdf/2002/2002-31/2002-31-842.PDF>
- Hänsgen K-D, Spicher B. (2002). Numerus Clausus: Numerus Clausus: le « test d'aptitudes pour les études de médecine » (AMS) permet-il de trouver les personnes les plus aptes? Schweizerische Ärztezeitung / Bulletin des médecins suisses / Bollettino dei medici svizzeri - 2002; 83 (47) 2562 – 2569. <http://www.saez.ch/pdf/2002/2002-47/2002-47-1144.PDF>
- Hänsgen, K.-D., Spicher, B. (2000). Zwei Jahre Numerus Clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 1: Erfahrungen mit dem EMS als Zulassungskriterium. Schweizerische Ärztezeitung Heft 12 S. 666 – 672.
- Hänsgen, K.-D., Spicher, B. (2000). Zwei Jahre Numerus Clausus und Eignungstest für das Medizinstudium in der Schweiz (EMS). Teil 2: EMS und Chancengleichheit. Schweizerische Ärztezeitung Heft 13 S. 723-730.
- Hänsgen, K.-D., Hofer, R. & Ruefli, D. (1995a). Un test d'aptitudes aux études de médecine est-il faisable en Suisse? Bulletin des médecins suisses, 7, S. 267 - 274.

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

11.1 Originaltest zur Information und Vorbereitung

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

11.2 Frühere Berichte des ZTD

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

Bericht 10 (2004)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2004 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 9 (2003)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2003 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 8 (2002)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2002 <i>Bericht über die Durchführung und Ergebnisse, Zusammenfassender Vergleich 1998 bis 2002 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 7 (2001)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2001 <i>Bericht über die Durchführung und Ergebnisse, Vorhersage von Prüfungserfolg Zusammenfassender Vergleich 1998 bis 2001 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 6 (2000)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 2000 <i>Bericht über die Durchführung und Ergebnisse, Evaluation der Vorhersage von Prüfungserfolg Zusammenfassender Vergleich 1998 bis 2000 - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 5 (1999)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 1999 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>
Bericht 4 (1998)
<p>EMS Eignungstest für das Medizinstudium in der Schweiz 1998 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsen und B. Spicher</p>

Bericht 3 (1997) - Rapport 3 (1997)	
<p>Eignungsprüfung für das Medizinstudium - Kriterien und Testverfahren <i>Bericht über das Internationale Symposium in Bern am 8. November 1996</i> Hrsg. von K.-D. Hänsgen und N. Ischi</p>	<p>L'examen d'aptitude aux études de médecine en Suisse - Critères et procédés d'application du test <i>Rapport sur le Symposium international à Berne</i> Editeurs: K.-D. Hänsgen; N. Ischi</p>
<p>Günter Trost Eignungskriterien bei der Zulassung zum Medizinstudium in Europa: Ergebnisse einer Erhebung</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>

Anhang: Berufswünsche der Kandidaten für Humanmedizin

Beeinflussen die unterschiedlichen Zukunftsvorstellungen der Teilnehmer die Art der Vorbereitung und das Testergebnis?

B. Weickhardt, B. Spicher, K.-D. Hänsgen

*Dieser Teil ist separat im Internet erhältlich
(siehe Berichte)*