

EMS

Eignungstest für das Medizinstudium 2007

Bericht 13 über die Durchführung und Ergebnisse 2007

avec un résumé en français

K.-D. Hänsgen und B. Spicher

Bericht 13 (2007)

Hängsen, Klaus-Dieter; Spicher, Benjamin (2007):

EMS Eignungstest für das Medizinstudium 2007;

Berichte des Zentrums für Testentwicklung, Nummer 13, im Auftrag der Schweizerischen Universitätskonferenz (SUK);

Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik, 2007.

Unter Mitarbeit von P. Bergem, M. Bernasconi, A. Debrot und M. Strazzeri.

In Zusammenarbeit mit der Rektorenkonferenz der Schweizer Universitäten (CRUS), welche das Anmelde- und Zulassungsverfahren zum Medizinstudium durchführt und für diesen Bericht statistische Angaben zur Verfügung stellt; K. Wechsler; Chr. Winzenried.

Redaktion: Tanja Läng

Die Mitglieder des Beirates EMS der Schweizerischen Universitätskonferenz 2007:

Dr. N. Ischi (Vorsitzender, SUK Bern); Prof. Dr. P. Eggli (Univ. Bern); Dr. R. Hofer (Univ. Bern);

Dr. C. Holzbaur (Med. Univ. Wien); Prof. Dr. M. Kleinmann (Univ. Zürich); Prof. Dr. J.-P.

Montani (Univ. Freiburg); Prof. Dr. K. Opwis (Univ. Basel); Prof. Dr. M. Perrez (Univ. Freiburg);

Prof. Dr. J.-L. Reymond (Univ. Bern); Prof. Dr. U. M. Spornitz (Univ. Basel); Dr. M. Stauffacher

(CRUS Bern); Prof. Dr. G. Trost (ITB GmbH Bonn Deutschland); K. Wechsler (CRUS Bern);

Prof. Dr. D. Wolfer (Univ. Zürich),

© ZTD Zentrum für Testentwicklung und Diagnostik 2007
am Departement für Psychologie der Universität Freiburg

Direktor: Prof. Dr. K.-D. Hängsen

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 in der Schweiz.....	5
1.3	Testgüte.....	6
1.4	Gleichbehandlung der Sprachgruppen	6
1.5	Differenzierung nach Alter	6
1.6	Differenzierung nach Geschlecht.....	6
1.7	Fazit.....	7
2	RÉSUMÉ	7
2.1	Statistique	7
2.2	Organisation et déroulement en Suisse	8
2.3	Qualité du test.....	8
2.4	Egalité de traitement entre groupes linguistiques	8
2.5	Différences entre classes d'âge	9
2.6	Différences entre femmes et hommes.....	9
2.7	Conclusion	9
3	NEUES ZUM EIGNUNGSTEST UND NUMERUS CLAUSUS	10
4	NUMERUS CLAUSUS (NC) UND MEDIZINSTUDIUM.....	17
5	ANMELDUNG ZUM MEDIZINSTUDIUM UND TEST 2007.....	20
5.1	Anmeldungen, Testantritte und Zulassungen	20
5.2	Grösse der Testlokale	27
5.3	Testorte, Wunschuniversitäten und Anreise	28
5.4	Wunschuniversität und Testort nach Wohnkanton.....	30
5.5	Testabsolvierung nach Alter und Geschlecht	34
5.6	Übernahme des Testergebnisses aus Vorjahren	38
5.7	Erneute Testteilnahme	39
5.8	Sprachgruppen	41
5.9	Alter und Maturitätsjahr	43
6	BESCHREIBUNG DES VERWENDETEN EIGNUNGSTESTS.....	44
6.1	Aufbau des Tests.....	44
6.2	Berechnung der Werte	45
6.3	Mittlerer Rangplatz der Untertests	46
7	TESTANWENDUNG IN DER SCHWEIZ 2007.....	48
7.1	Verteilungsprüfung.....	48
7.2	Vergleich der Testfassungen 1998 bis 2007	51
7.3	Äquivalenz der Sprachversionen	53
7.3.1	Sprachvergleich für die Untertests.....	55
7.3.2	Darstellung des Korrekturverfahrens	57
7.3.3	Effekte der Korrektur.....	66
7.4	Vergleichbarkeit der Testlokale.....	71

7.5	Vergleich für die Geschlechter	73
7.6	Vergleiche für Altersgruppen	77
7.7	Vergleiche nach Wunschuniversitäten	79
8	ERGEBNISSE ZUR TESTGÜTE	80
8.1	Zuverlässigkeit	80
8.2	Binnenstruktur.....	82
8.3	Item-Trennschärfen.....	86
8.4	Konzentriertes und sorgfältiges Arbeiten.....	91
9	TEILNAHMEN AM EMS NACH KANTONEN SEIT 1998.....	96
10	BEISPIELAUFGABEN FÜR DIE UNTERTESTS	97
10.1	Untertest: Quantitative und formale Probleme.....	97
10.2	Untertest: Schlauchfiguren	97
10.3	Untertest: Textverständnis	98
10.4	Untertest: Planen und Organisieren.....	99
10.5	Untertest: Konzentriertes und Sorgfältiges Arbeiten	101
10.6	Untertest: Medizinisch-naturwissenschaftliches Grundverständnis	101
10.7	Untertest: Figuren lernen	102
10.8	Untertest: Fakten lernen	103
10.9	Untertest: Muster zuordnen	103
10.10	Untertest: Diagramme und Tabellen.....	104
11	LITERATUR	105
11.1	Originaltest zur Information und Vorbereitung	106
11.2	Frühere Berichte des ZTD	107

1 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2007 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Human-, Veterinär- und Zahnmedizin ab Herbstsemester 2007. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2006 herangezogen.

1.1 Statistik

In diesem Jahr wurde der EMS in der **Schweiz** zum zehnten Mal durchgeführt. Ein Numerus clausus ist 2007 wiederum für Humanmedizin (seit 1998), Veterinärmedizin (seit 1999) und Zahnmedizin (seit 2004) und dort für Personen notwendig, die sich an den Universitäten Basel, Bern, Freiburg oder Zürich angemeldet haben. Seit 7 Jahren steigt die Zahl der Anmeldungen zum Medizinstudium stetig an, der diesjährige Zuwachs gegenüber 2006 ist der grösste aller bisherigen Jahre (auch bedingt durch die erstmals bestehende Möglichkeit einer Anmeldung im Internet).

Im Februar 2007 waren mit den vorliegenden Anmeldungen die Kapazitäten dieser Universitäten in Humanmedizin zu 372% (Vorjahr 314%), in Veterinärmedizin zu 274% (Vorjahr 232%) und in Zahnmedizin zu 188% (Vorjahr 153%) ausgelastet.

Insgesamt 1932 Personen haben sich bis zur gesetzten Frist im Mai 2007 bei der CRUS zum EMS angemeldet, 80 Personen übernehmen ihr Testergebnis aus dem Vorjahr.

Erfolgreich absolviert haben den EMS 1831 Personen. 2 Personen haben den Test während des Vormittagsteils abgebrochen, 99 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen.

Für **Österreich** gilt: Für beide Universitäten hatten sich insgesamt 6914 Personen angemeldet, wobei sich in Innsbruck 2638 Personen auf 400 Studienplätze und in Wien 4276 Personen auf 740 Studienplätze bewarben. Es war die Bereitstellung der Testlogistik für diese 6914 Personen notwendig. Absolviert haben den Test dann tatsächlich 3940 Personen, 57 Personen übernehmen ihr Testergebnis aus dem Vorjahr.

In diesem Bericht werden die Gütekennwerte dieses Testeinsatzes in Österreich mit berücksichtigt. Weitere inhaltliche Auswertungen und Evaluationen liegen in den Händen der österreichischen Partner und werden in einem eigenen Bericht dargestellt.

1.2 Organisation und Ablauf in der Schweiz

Der Test fand am 6.7.2007 gleichzeitig an acht Testorten in drei Sprachen statt (deutsch: Aarau-Suhr, Basel, Bern, Chur, St. Gallen, Zürich; französisch: Freiburg; italienisch: Bellinzona). Zu beachten ist, dass die Wahl des Testortes unabhängig vom zukünftigen Studienort möglich ist. Insgesamt waren 21 Testlokale eingerichtet. Dank einer engagierten Mitwirkung aller Beteiligten kann die Testabnahme erneut als erfolgreich eingeschätzt werden.

In diesem Jahr erfolgte die Anmeldung zum Studium im Februar erstmals elektronisch – in den Vorjahren bekanntlich durch Einsendung eines Formulars, welches zunächst beschafft werden musste. Für das überdeutliche Ansteigen der Anmeldezahlen 2007 und die bisher grössten Rückzüge bis zum Zeitpunkt der Anmeldung zum EMS muss man die Tatsache berücksichtigen, dass die „Schwelle“ für die elektronische Anmeldung wesentlich niedriger

liegt. Trotz dieser Bedingungen ist die Zahl der Personen mit gültigem Testwert auch 2007 gegenüber den Vorjahren weiter gestiegen. Sollte eine Rückkehr zur nichtelektronischen Anmeldung erfolgen, müssen diese Fakten bei der Bewertung der Anmeldung 2008 berücksichtigt werden. Mit einem Rückgang oder Stagnieren der Februar-Zahlen wäre zu rechnen, was trotzdem zu einem weiteren Ansteigen der gültigen Testteilnahmen führen kann.

1.3 Testgüte

Die Zuverlässigkeitskennwerte liegen 2007 mit 0.80 (Konsistenz) bzw. 0.90 (Testhalbierungsmethode) im gewohnt hohen Bereich und sie erfüllen die Qualitätsanforderungen an einen psychologischen Test. Auch die Faktorenstruktur bleibt über die Jahre hinweg sehr gut vergleichbar, was die Übernahmemöglichkeit der Testergebnisse ins Folgejahr rechtfertigt.

Alle Items konnten 2007 gewertet werden, da es keine Ausschlussgründe wegen nicht ausreichender Itemkennwerte (Schwierigkeit, Trennschärfe) gab.

Man kann aufgrund der zweijährigen Erfahrungen des Testeinsatzes in Österreich nun einschätzen, dass im deutschen Sprachraum keine bedeutsamen Unterschiede bezüglich der Ergebnisse des EMS (bzw. in Deutschland des TMS) bestehen.

1.4 Gleichbehandlung der Sprachgruppen

Auch 2007 wurden durch das sogenannte DIF-Verfahren diejenigen Items von sprachabhängigen Untertests identifiziert, die in der französischen und italienischen Sprachgruppe vom allgemeinen Trend der Ergebnisse in allen Aufgaben abweichen. Weil zumindest ein Risiko besteht, dass testbedingte Ursachen dafür mitverantwortlich sein können, erfolgte eine Korrektur dieses Unterschiedes. Durchschnittlich werden den französischsprachigen Teilnehmern 0.73 Punkte, den italienischsprachigen Teilnehmern 1.57 Punkte gutgeschrieben. Der Bonus beträgt 0 oder 1 (französischsprachige Teilnehmer), bzw. 0 bis 3 Punkt(e) (italienischsprachige Teilnehmer). Der Einfluss dieser Korrektur ist damit wie in den Vorjahren insgesamt gering. Es kann erneut nachgewiesen werden, dass die Unterschiede in den nicht sprachabhängigen Untertests grösser sind als diejenigen in den sprachabhängigen Tests. Eine sprachbedingte Benachteiligung kann daher auch in diesem Jahr ausgeschlossen werden. Die vorhandenen Unterschiede sind wohl vor allem auf eine unterschiedliche Zusammensetzung der drei Sprachgruppen zurückzuführen, die französische und die italienische Sprachgruppe wählen beispielsweise auch andere Studienorte ohne Numerus clausus.

1.5 Differenzierung nach Alter

Die bekannten Ergebnisse werden erneut bestätigt: Die Gruppe der Ältesten mit später Maturitätsprüfung (vor allem zweiter Bildungsweg) erreichen signifikant geringere Testwerte. Bekanntlich sagt hier der Test korrekt eine etwas geringere Wahrscheinlichkeit für Studienerfolg voraus, die sich für diese Gruppe in den Evaluationsstudien nachweisen liess.

Die Älteren mit früher Maturitätsprüfung unterscheiden sich wie in den Vorjahren nicht signifikant von den anderen beiden untersuchten Altersgruppen.

1.6 Differenzierung nach Geschlecht

Erneut bewerben sich mehr Frauen (64%) um ein Medizinstudium. Besonders deutlich fällt der Unterschied im Fach Veterinärmedizin aus, für welches 82% Bewerberinnen sind.

Chancengleichheit definiert sich auch hier nicht über gleiche Mittelwerte beim Testwert, sondern über die korrekte Vorhersage des Studienerfolges für beide Gruppen. Dies wurde in den früheren Evaluationsstudien nachgewiesen. Die Quotenunterschiede weisen auf unterschiedliche Repräsentativität der Bewerberinnen und Bewerber für ihre jeweilige Geschlechterkohorte hin und allein deshalb wären identische Testwerte nicht zu erwarten. Von 1998 bis 2006 betrug der Testwert-Unterschied zwischen den Geschlechtern immer zwischen 1.4 und 2.2 Punkten zugunsten der Männer. Der Unterschied von 2 Punkten im Jahr 2007 liegt im Rahmen der Vorjahre; es kann daher auch für 2007 von einer Chancengleichheit ausgegangen werden.

1.7 Fazit

Auch die zehnte Testsession des EMS verlief in der Schweiz erfolgreich – die Ergebnisse erfüllen alle Anforderungen an die Güte eines psychologischen Testverfahrens. Mit dem EMS steht für den Numerus clausus ein bewährtes Selektionsinstrument zur Verfügung, welches eine kapazitätsgenaue Zulassung aufgrund der Studieneignung ermöglicht und auch die notwendigen Umleitungen zwischen den Universitäten nach einem objektiven Kriterium regelt.

Die Entscheidung über die Notwendigkeit des Numerus clausus wird weiterhin jährlich getroffen. Aufgrund der statistischen Prognosen für das Interesse an einem Medizinstudium muss auch künftig ein qualitativ hochstehender Test als Zulassungskriterium zur Verfügung stehen. Selbst bei einer Kapazitätserhöhung um 20% stünden 2007 für 2000 Angemeldete in Humanmedizin nur 700 Plätze zur Verfügung.

2 Résumé

Le présent rapport expose les résultats concernant le test d'aptitudes aux études de médecine en Suisse (AMS) ainsi que les admissions aux études de médecine humaine, vétérinaire et dentaire au semestre d'automne 2007. Il fait également intervenir des données comparatives des éditions antérieures du test (1998-2006).

2.1 Statistique

En 2007, l'AMS a été conduit pour la dixième fois en **Suisse**. Un Numerus clausus a été à nouveau nécessaire pour la médecine humaine (NC depuis 1998), la médecine vétérinaire (NC depuis 1999) et la médecine dentaire (NC depuis 2004) pour les personnes inscrites aux Universités de Bâle, Berne, Fribourg ou Zurich. Le nombre d'inscriptions augmente continuellement depuis 7 ans, l'accroissement de cette année par rapport à l'année précédente a été le plus important depuis l'introduction de l'AMS (aussi conditionné par le fait qu'on a pu s'inscrire pour la première fois électroniquement).

En février 2007, les préinscriptions dépassaient les capacités de ces Universités, que ce soit en médecine humaine à raison de 372% (année précédente : 314%), en médecine vétérinaire à raison de 274% (année précédente : 232%) ou en médecine dentaire à raison de 188% (année précédente : 153%).

Au délai fixé en mai 2007, un total de 1932 personnes s'était annoncé à la CRUS pour le test AMS ; 80 personnes ont repris leur résultat de l'année précédente.

1831 personnes ont achevé l'AMS, 2 personnes l'ont interrompu en cours de matinée ; 99 personnes ont retiré leur inscription ou ne se sont pas présentées au test.

En **Autriche**, 2638 se sont inscrites pour les 400 places de l'Université d'Innsbruck et 4276 pour les 740 places de celle de Vienne. La logistique du test ainsi a dû être organisée pour 6914 personnes. 3940 personnes ont effectivement réalisé le test, 57 ont repris leur résultat de l'année précédente.

Dans ce rapport, les données sur la mise en place de ce test en Autriche sont aussi prises en compte. L'organisation de la passation du test ainsi que d'autres évaluations sont entre les mains des partenaires autrichiens et seront présentées dans leur propre rapport.

2.2 Organisation et déroulement en Suisse

Le test s'est déroulé le 6 juillet 2007 simultanément en huit sites différents et en trois langues (français : Fribourg ; allemand : Aarau-Suhr, Bâle, Berne, Coire, St-Gall, Zurich ; italien : Bellinzone). Il est à noter qu'il est possible de choisir le lieu du test indépendamment du futur lieu d'études désiré. Grâce à la collaboration engagée de toutes les personnes impliquées, l'édition 2007 peut à nouveau être considérée comme réussie.

Cette année, l'inscription aux études en février s'est faite pour la première fois électroniquement – les années précédentes, il fallait renvoyer un formulaire que l'on devait se procurer auparavant. Pour expliquer l'augmentation drastique du nombre d'inscriptions et la plus grande part de retrait jamais connue jusqu'à l'inscription au test, il faut tenir en compte le fait que le « seuil » pour une inscription électronique se situe clairement plus bas. Malgré ces conditions, le nombre de personnes avec un résultat au test valable a, en 2007, aussi augmenté par rapport à l'année précédente. Si l'on décide de retourner à une inscription non électronique, ces faits doivent être pris en compte dans l'évaluation des inscriptions 2008. On peut s'attendre à ce que le nombre d'inscriptions en février recule ou stagne, ce qui peut tout de même conduire à une augmentation du nombre de participations au test valables.

2.3 Qualité du test

Les indices de fiabilité sont, en 2007 comme pour les années précédentes, élevés, avec 0.80 (consistance) et 0.90 (méthode de partition pairs-impairs) et ils remplissent les exigences de qualité des tests psychologiques. La structure factorielle reste également comparable d'année en année, ce qui légitime la possibilité de reprendre son résultat pour l'année suivante.

Tous les items ont pu être évalués cette année ; il n'y avait pas de raison d'en supprimer pour cause de consistance insuffisante des items (difficulté, sélectivité).

Grâce à l'expérience d'une deuxième année d'application du test en Autriche, l'on peut maintenant estimer que, dans l'espace germanophone, il n'y a pas de différence significative quant aux résultats du TMS/EMS.

2.4 Egalité de traitement entre groupes linguistiques

Comme chaque année depuis 1999, grâce à la procédure appelée DIF, on a pu identifier les items des sous-tests à forte composante linguistique qui s'écartent des tendances générales des résultats dans tous les exercices dans les groupes francophones ou italophones. Comme il existe un risque que des facteurs liés au test soient responsables de ces différences, elles sont corrigées. En moyenne, les participants francophones sont crédités de 0.73 points, les participants italophones de 1.57 points. Le bonus se situe entre 0 et 1 point pour les

participants francophones et entre 0 et 3 points pour les participants italophones. Au total, l'influence de ces corrections est faible, comme pour les années précédentes. On peut aussi à nouveau prouver que les différences sont plus grandes dans les sous-tests à faible composante linguistique que dans ceux à forte composante linguistique. Cette année aussi, un préjudice lié à la version linguistique du test peut donc être exclu. Les différences existantes doivent plutôt être attribuées à une composition différente des trois groupes linguistiques : les francophones et les italophones peuvent par exemple choisir d'autres lieux d'études sans Numerus clausus.

2.5 Différences entre classes d'âge

Les résultats des années précédentes peuvent à nouveau être confirmés : le groupe des plus âgés ayant obtenu plus tard leur maturité (surtout lors d'une deuxième formation) ont des résultats significativement plus faibles. Il avait déjà été précédemment montré que le test prédit correctement une probabilité quelque peu moins élevée de succès aux études pour ce groupe, fait pouvant être confirmé dans des études d'évaluation.

Les participants plus âgés ayant obtenu tôt leur maturité ne se différencient pas des deux autres groupes d'âge étudiés, tout comme les années précédentes.

2.6 Différences entre femmes et hommes

Il y a toujours plus de femmes (64%) qui se présentent aux études de médecine. Cette différence est particulièrement marquée dans la branche de médecine vétérinaire, où 82% sont des participantes. L'égalité des chances ne se définit ici pas non plus par des moyennes égales au test, mais par la prédiction correcte du succès aux études pour les deux groupes. Cela a été confirmé dans des études d'évaluation antérieures. Les différences dans les quotas indiquent que la représentativité des candidats masculins et féminins est différente ; rien que pour cette raison, on ne devrait pas s'attendre à des scores au test identiques. De 1998 à 2006, cette différence s'est toujours élevée à des valeurs comprises entre 1.4 et 2.2 points en faveur des hommes. La différence de 2 points de 2007 se situe dans la même tranche que les années précédentes ; l'on peut donc présumer que l'égalité des chances est donnée cette année aussi.

2.7 Conclusion

La dixième session de l'AMS s'est aussi déroulée avec succès en Suisse – les résultats remplissent toutes les conditions requises pour un test psychologique de qualité. L'AMS représente un instrument de sélection éprouvé pour le Numerus clausus, qui permet une admission correspondant aux capacités en fonction des aptitudes aux études et qui règle l'attribution des lieux d'études selon un critère objectif.

La décision de la nécessité de l'AMS continue à être prise annuellement. En raison des pronostics statistiques sur le dépassement des capacités des universités, un test de haute qualité doit être disponible pour les années à venir.

3 Neues zum Eignungstest und Numerus clausus

In diesem Jahr erfolgte die mittlerweile zehnte Testdurchführung des Eignungstests für das Medizinstudium (EMS) in der Schweiz. Am selben Tag wurde die gleiche Version des EMS in Innsbruck und Wien eingesetzt, wo der Test zum zweiten Mal durchgeführt wurde. Schliesslich wurde durch das Institut für Test- und Begabungsforschung (ITB GmbH) Bonn, unserem langjährigen Kooperationspartner, im Mai dieses Jahres der „Test für Medizinische Studiengänge“ (TMS) auch in Deutschland für vier Universitäten in Baden Württemberg (Freiburg, Heidelberg, Tübingen und Ulm) wieder durchgeführt. Grundsätzlich ergeben sich dadurch neue Chancen, das Zulassungsinstrument durch gemeinsame Anstrengungen dreier Länder weiterzuentwickeln.

Die Diskussion zur Entwicklung von Studienfach und Berufsfeld „Medizin“ in der Schweiz ist nicht abgeschlossen, viele – auch widersprüchliche – Interessen gilt es gegeneinander abzuwägen (Ärztebedarf, strukturelle Ausgeglichenheit zwischen den Fachgebieten, Gesundheitskosten, Ausbildungskapazitäten und den Kantonen zur Verfügung stehende Mittel sowie Nachfrage nach einem Medizinstudium). Jede Lösung wird sich um Abwägung von Interessen bemühen müssen. Bestimmte historische Fakten dürfen dabei nicht vergessen werden, um die Zulassung zum Studium als „erste Entscheidung“ für einen Beruf de Mediziner weiterhin optimal zu gestalten.

Hin und wieder gibt es auch Diskussionen darüber¹, ob das gegenwärtige Auswahlverfahren angemessen ist und ob sich zu viele in erster Linie naturwissenschaftlich interessierte Studierende zum Medizinstudium anmelden und andere „abgeschreckt“ werden. Können die Auswahl diejenigen benachteiligen, die beispielsweise Hausarzt oder Psychiater werden wollen? Was kann man dazu anmerken?

Zuerst: Weder Test noch Studienanforderungen sind etwas Statisches. Es gibt zwei wichtige Forderungen, die bei jeder Studierendenauswahl, jedem Numerus clausus (NC) zwingend zu beachten sind:

- Die **gesetzlichen** Grundlagen² des Numerus clausus sind kantonal geregelt und in allen Gesetzen steht sinngemäss gleichlautend, dass die **Eignung** zum **Studium** bei der Zulassung zu berücksichtigen ist.
- Die **Bundesgerichtsurteile** zum NC³, welches diesen guthies (aber gesetzliche Grundlagen forderte), formulierten zwei wichtige Bedingungen: **Rechtsgleichheit und Willkürfreiheit**

Wenn aktuell nur noch ca. 50% der Personen, die sich für Humanmedizin an einer Universität mit NC anmelden, einen Studienplatz angeboten bekommen können, steht das zwar für eine „Verschärfung“ des NC in der Schweiz – ist im internationalen Massstab aber immer noch moderat. 30% oder gar 10% sind in einigen Ländern nicht ungewöhnlich. Medizin ist ein begehrtes und offenbar zukunftssicher eingeschätztes Studium, was zunehmende Nachfrage findet.

Man sollte sich darüber klar sein: Das „**Image**“ des **Medizinstudiums** ist bisher, dass es ein anspruchsvolles Studium ist und auch gut entwickelte Fähigkeiten zur Wissensaneignung erfordert. Würde dieses Image **grundsätzlich** verändert, dann verändert sich wahrscheinlich auch die Bewerberzusammensetzung und die Nachfrage könnte weiter steigen.

Der NC in der Schweiz ist eine **politische** Entscheidung, eine Interessenabwägung zwischen einer bezahlbaren Ausbildungsqualität und der Studiennachfrage – primär nicht gedacht als berufspolitisches Regelungsinstrument. Je „schärfer“ der NC wird, umso mehr gilt, dass

¹ z.B. „*Bilden wir die falschen Studenten zu Ärzten aus?*“ von Urs Stoffel NZZ vom 4. Oktober 2007

² www.crus.ch/information-programme/anmeldung-zum-medizinstudium.html

³ Urteile des Schweizerischen Bundesgerichtes 2P.171/1998/mks vom 3. März 1999 sowie 2P.228/1998 – 2P.332/1998/luc vom 20. August 1999.

einige der Abgelehnten das Medizinstudium sicherlich auch bewältigt hätten. Es muss daher gewährleistet sein, dass die tatsächlich Zugelassenen es **besser** bewältigt haben und z.B. den Abschluss schneller und mit besseren Ergebnissen erreicht haben.

Für eine „**Elitenbildung**“ **bereits bei der Anmeldung** zum Studium gibt es keine Belege. Seit der Einführung des EMS 1998 haben die Anmeldungen an den Universitäten mit NC für Humanmedizin um 112% zugenommen. Das „Leistungsspektrum“ der am EMS teilnehmenden Personen wurde aber nicht besser, es weist zwischen den Jahren gleiche Verläufe und Variabilitäten auf (siehe z.B. Kapitel 7.2 auf Seite 51). Die von einigen Seiten auch erhoffte „Dissuasion“ durch den Test hat es – wenn überhaupt – nur in den ersten Jahren gegeben.

Die **Gründe für die Einführung des Numerus clausus** dürfen nicht vergessen werden. Die Universitäten waren im ersten Semester stark überlastet, die Studienbedingungen katastrophal und mit Sicherheit nicht im gewünschten Sinne persönlichkeitsbildend (auch im Sinne eines „Kampfes“ um Ressourcen, um alle Bücher, Einschreibungen u.a. rechtzeitig zu erhalten).

Es gab hohe Abbrecherquoten und viele Personen wurden auch durch Prüfungen „selegiert“, um die nachfolgende klinische Ausbildung noch sinnvoll durchführen zu können. Deren Kapazität kann aber durch die Patientenzahlen und Betreuungskapazitäten der Universitätskliniken damals wie heute nicht beliebig erhöht werden.

Seitdem hat die Zahl der ihr Studium erfolgreich abschliessenden Personen zugenommen. Die begrenzte Studienplatzkapazität wird besser „bewirtschaftet“, würde ein Bildungsökonom sagen¹. **Ein NC erhöht bei konstanter Kapazität die Absolventenzahlen.** Einige Universitäten konnten überhaupt erst nach Einführung des Numerus clausus durch die Verbesserung der Betreuungsverhältnisse Studienreformen (problemorientierten Unterricht, früheren Patientenkontakt usw.) einführen und dadurch die Persönlichkeitsentwicklung günstiger gestalten. Man muss den Preis dieser heute als normal empfundenen Verbesserungen nicht vergessen, weil ein NC natürlich eine auch Interessenabwägung zwischen freiem Hochschulzugang und optimalen Studienbedingungen ist.

Ausbildungskapazitäten und deren Erhöhung sind ein anderes, ein politisches Thema. Ärztebedarf und Studienplatzangebot müssen so in Übereinstimmung gebracht werden, dass es für die Kantone auch bezahlbar bleibt. Die Ärztedichte hat sich ja nicht vermindert, sondern es scheint ein „Strukturproblem“ zu geben, indem einzelne Versorgungsbereiche für die Absolventen weniger attraktiv sind. Allein zwischen 2002 und 2006 hat die Zahl der berufstätigen Ärzte um 11.2% zugenommen².

Wenn die Nachfrage nach einem Medizinstudium wie bisher bleibt oder weiter steigt, werden ziemlich sicher auch im Falle der **Anhebung der Ausbildungskapazitäten** weiterhin ein Numerus clausus und ein funktionierendes **Zulassungskriterium** notwendig sein. Beim Vorschlag des Wissenschaftsrates, die Kapazität um 20% anzuheben², stünden für 2007 dann 2 000 Anmeldungen immer noch nur 700 Plätze in Humanmedizin gegenüber. Eine neue Mobilität innerhalb Europas bei der Wahl des Studienortes stellt hier neue Anforderungen für eine Sicherung der Ausbildungskapazitäten gegen Überlastung (was sich z.B. in Österreich als Problem zeigt, weil die Möglichkeit des Hochschulzuganges innerhalb der EU unabhängig vom Wohnland gefordert wird). Auch die Schweiz wird hier vermutlich langfristig mehr Nachfrage aus den EU-Staaten erhalten.

Weshalb wird der EMS mit Schwerpunkt auf **naturwissenschaftlichen** Fähigkeiten durchgeführt? Die Gesetze der Kantone schreiben vor, dass die **Studieneignung** bei der Auswahl zu berücksichtigen ist. Die hohen Abbruchquoten des Studiums waren vor allem auf fehlende Kompetenzen auf diesem Gebiet zurückzuführen. In den ersten Jahren des NC wurden nur wenige Personen abgewiesen und im Nachrückverfahren doch zugelassen. Diese Personen haben in der ersten und zweiten Vorprüfung deutlich schlechtere Leistungen

¹ Siehe auch Bericht 12 des ZTD S. 11, www.unifr.ch/ztd/ems

² „Ärztedemografie und Reform der ärztlichen Berufsbildung“ Bericht des SWTR vom 6.9.2007

erreicht bzw. sind gar nicht angetreten – der Test hat diese geringen Erfolgsaussichten vorhergesagt (siehe Bericht 7 des ZTD). Es sollen diejenigen bevorzugt einen Studienplatz erhalten, die die grösste Chance haben, das Studium auch zu beenden. Dies spart nicht nur Kosten für Studienplätze, die Vergeudung von individuellen Ressourcen ist ebenso bedeutsam, wenn Studierende Lebenszeit erfolglos auf diesem Weg verbringen.

Der EMS ist kein Wissenstest, sondern ein **Fähigkeitstest**. Es werden keine Fakten abgefragt, sondern Problemstellungen mit den notwendigen Informationen vorgegeben, die zu lösen sind. Intendiert ist, dass es sich um eine **Arbeitsprobe** für die Studierfähigkeit handelt. Die seinerzeit durchgeführten Anforderungsanalysen für ein Studium der Medizin wurden sehr aufwändig und von verschiedenen Seiten durchgeführt (siehe Trost, 1989). Auf der Grundlage dieser Analysen wurden ca. 50 Bereiche und Aufgabentypen geprüft und die unter vielen Aspekten am besten geeignet erscheinenden 13 für die Erprobungen im Übergangsverfahren des TMS ausgewählt und empirisch erprobt. Von diesen haben vor allem aufgrund der Ergebnisse zur Validität dann 9 Aufgabengruppen in den TMS Eingang gefunden.

Der EMS sagt diese Studieneignung gemessen an Massstäben vergleichbarer Tests sehr gut voraus (siehe Evaluationsberichte zum EMS¹), je höher das Testergebnis, umso wahrscheinlicher ist das Bestehen der Prüfungen. Die Zulassung aufgrund des Testwertes ist also ausreichend begründet. Für das Auswahlverfahren entsteht erst dann neuer Handlungsbedarf, wenn sich die Studienanforderungen ändern. Dies wird kontinuierlich beobachtet. 2003 fand dazu z.B. ein Workshop mit den beteiligten Fakultäten statt und ein Ergebnis ist, dass die Planungskompetenz durch einen neuen Untertest „Planen und Organisieren“ seit 2005 erfasst wird. Eine nächste Runde wird aktuell geplant.

Soziale Kompetenzen sollen grundsätzlich auch erfasst werden – hier gibt es aber zwei Probleme:

- Sie gehen nicht bis kaum in die **Bewertungen** (Notenvergabe) beim Studium ein. Wenn man „Studieneignung“ streng definiert, dürfte man soziale Kompetenzen nicht als Zulassungskriterium verwenden. Man kann nur indirekt darauf schliessen, dass soziale Kompetenzen (wie die Planungskompetenz auch) Anforderungen für eine erfolgreiche Gestaltung des Studiums insgesamt sind. Sie werden also durch Analyse der Anforderungen für die Tätigkeit „Studieren“ begründet – müssen sich in zukünftigen Evaluationen dennoch als Prognoseinstrumente für Studieneignung bewähren.
- Alle bisher betrachteten Tests zur „Sozialen Kompetenz“ sind trainierbar und damit verfälschbar. Man kann sich relativ leicht in Fragebögen als sozial kompetent beschreiben, wenn man die Konstruktionsprinzipien ungefähr kennt. Sie sind daher nicht für ein objektives (willkürfreies) Zulassungsverfahren geeignet. Auch Interviews können nicht für 2 000 Personen vergleichbar objektiv durchgeführt werden. Nur ganz wenige Testkonzepte, die mehr Wert auf die Erfassung sozialer **Fähigkeiten** legen sind eventuell geeignet. Diese sind vergleichsweise anspruchsvoll (z.B. Videodarbietung mit gleichen Bedingungen an allen Testorten) und der zusätzliche Aufwand steht bisher nicht im Verhältnis zum Gewinn bei der Vorhersage von Studieneignung.

Die Erfassung der sozialen Kompetenz bleibt eine methodische Herausforderung für die Zukunft.

Man kommt hier aber weiter, wenn man **Eignung und Neigung** unterscheidet. Neigung wäre alles, was mit Interesse und Motivation für ein Studienfach umschrieben werden kann. Bei fehlender Neigung wird ein Studium sicher gar nicht aufgenommen. Problematischer sind die Fälle, wo man während des Studiums entdeckt, dass ein Fach nicht den eigenen Neigungen entspricht. Dies kann an fehlenden oder falschen Informationen über Studium und Beruf liegen. Die Erfassung von Neigungen und deren Verwendung als Zulassungskriterium ist

¹ www.unifr.ch/ztd/ems/berichte/ Vorhersage des Studienerfolges, Studie 1 und 2

wegen der hohen Anfälligkeit für Trainierbarkeit und Verfälschung („sich im besten Licht darstellen“) ebenfalls praktisch unmöglich.

2006 und 2007 wurden die EMS-Teilnehmer befragt und nur 14% erlebten die **Informationsmöglichkeiten** zum **Wunschberuf** in beiden Jahren als zu gering oder viel zu gering, der Rest als ausreichend bis sehr gut. Die Informationsmöglichkeiten zu den **Studienanforderungen** erlebten 17% (2006) bzw. 16% (2007) als zu gering oder viel zu gering. Entweder es gibt wirklich genügend Informationen – oder das **vorhandene** Informationsbedürfnis vor dem Studium wurde im Wesentlichen gestillt, auch wenn es zu niedrig wäre. Die Studierenden müssen die Anforderungen des Studiums und des Berufes vor einer Bewerbung ohne Zweifel besser kennenlernen und sich z.B. auch damit auseinandersetzen, was ein Hausarzt tut und ob man sich dazu berufen fühlt. Eine Fachrichtung wird nur gewählt, wenn sie attraktiv ist und den eigenen Neigungen entspricht. Für die hausärztliche Spezialisierung muss daher gezielt geworben werden, man muss auch diejenigen besonders motivieren, die günstige Voraussetzungen mitbringen.

Wie könnte man dafür sorgen, dass die Übereinstimmung eines Studiums mit den eigenen Neigungen besser berücksichtigt würde? Eine in anderen Ländern erfolgreich eingesetzte Methode sind **Studienberatungsinstrumente**, die zur „Selbstberatung“ angeboten werden und meist im Internet realisiert sind. Es werden detaillierte Informationen über die verschiedensten Anforderungen von Studium und späterem Beruf gegeben und Checklisten sowie Tests angeboten, um ohne Druck eigene Fähigkeiten und Neigungen zu erkunden. Da hier das Interesse der Personen an realistischen Aussagen überwiegt (weil es nicht als Zulassungskriterium verwendet wird), ist es auch nicht sinnvoll, sich in einem „besseren Licht“ darzustellen. Es gibt mittlerweile Universitäten, welche die **Absolvierung** eines solchen Studienberatungstestes zur Pflicht machen (ohne die Ergebnisse selbst für die Zulassung zu verwenden). Es wäre eine Herausforderung mit grossem erwarteten Nutzen, für Medizin ein solches Selbstberatungsinstrument zu entwickeln – die meisten uns bekannten Instrumente beziehen sich bisher auf technische Berufe.

Wer interessiert sich für das Fachgebiet „Hausarzt“? Die Berufsinteressen der Teilnehmenden am EMS werden seit 2005 jährlich nach dem Test in einem freiwillig auszufüllenden Fragebogen erfragt. Es sind „nur“ jeweils um 10% **aller** Testteilnehmer in der Schweiz, die Hausarzt werden wollen. Die Anteile decken sich mit den Ergebnissen der drei Quoten in Österreich 2006, ein grosser Teil ist noch unentschieden. Das ist offenbar eine sehr stabile Ausgangslage, wobei sich zumindest andeutet, dass die „Hausarztdebatte“ in der Schweiz einen minimalen Effekt zeigt. Diese Ausgangssituation kann man nur ändern, wenn man weiter gezielt für die Orientierung „Hausarzt“ wirbt – und im Auge behält, dass die Entscheidung nach dem Studium eine ganz andere sein kann, wenn man die Fächer genauer kennt.

Was ist über die Interessenten für „Hausarzt“ bekannt? Es zeigen sich über die letzten drei Jahre sehr konsistente Ergebnisse:

2005 lag der mittlere Testwert bei 100.4 (30 Männer: 101.6, 68 Frauen: 99.9), 2006 betrug der mittlere Testwert 100.6 (68 Männer: 101.1, 101 Frauen: 100.3). 2007 war der mittlere Testwert 101.0 (56 Männer: 101.8, 95 Frauen: 100.5).

- Die Bewerber für „Hausarzt“ erreichen keine schlechteren Ergebnisse als andere (der Testwert hat für alle Kandidaten bekanntlich den Mittelwert 100).
- Das Geschlechterverhältnis weicht nicht wesentlich vom Verhältnis in der Gesamtgruppe (40% Männer, 60% Frauen) ab.
- Auch der Testwert-Unterschied Männer zu Frauen liegt im Rahmen der Unterschiede der Gesamtgruppe.

Interessenten für „Hausarzt“ sind also ganz „typische“ Bewerber für ein Medizinstudium und keinesfalls schlechter im naturwissenschaftlichen Denken. **Sie werden bei der Zulassung nicht benachteiligt.** Insofern ist auch die Annahme, dass Interessenten für den Hausarztberuf naturwissenschaftlich „minderbemittelter“ wären, unbegründet.

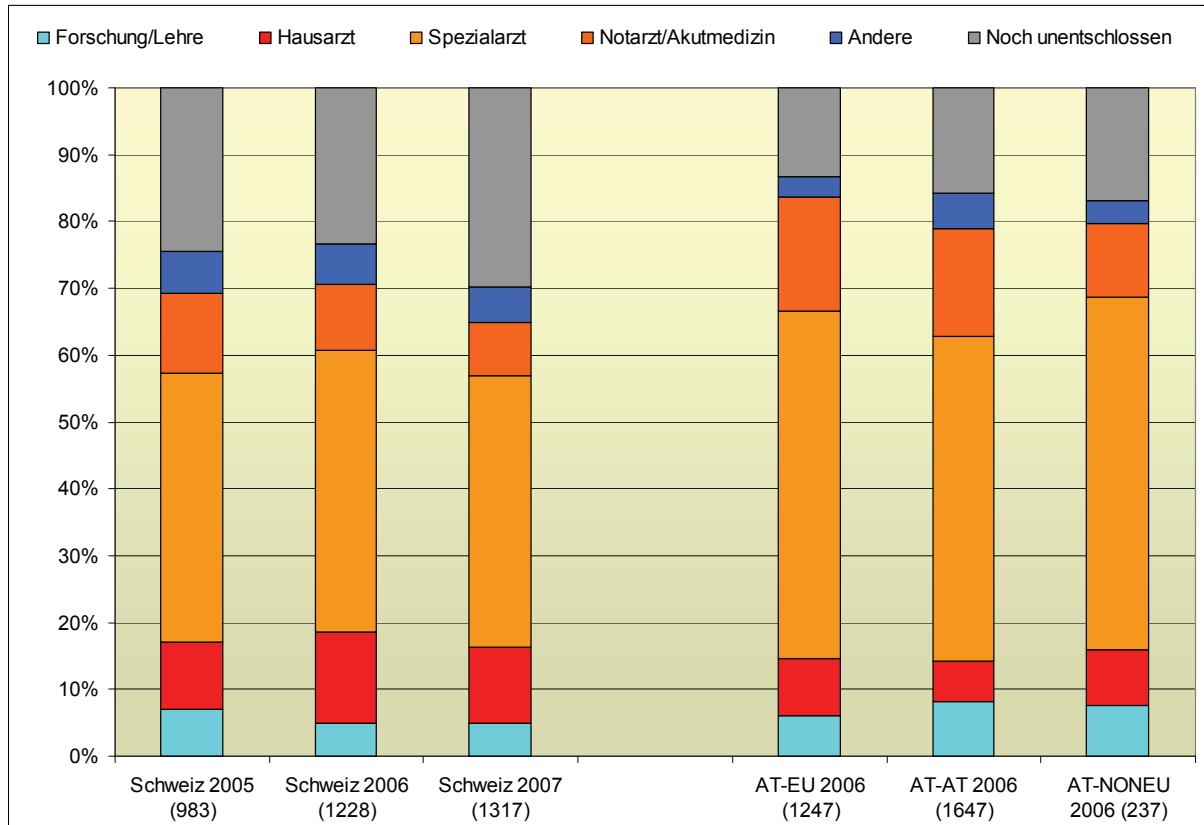


Abbildung 1: Ergebnisse der Befragungen zum zukünftig gewünschtem Fachgebiet (nur Bewerbungen Humanmedizin). Schweiz 2005 bis 2007 und Vergleichsdaten Österreich 2006 für die Quoten „EU-Bürger“ (AT-EU), „Bürger mit österreichischer Maturität“ (AT-AT) und „Nicht-EU-Bürger“ (AT-NONEU), in Klammern die Personenzahl.

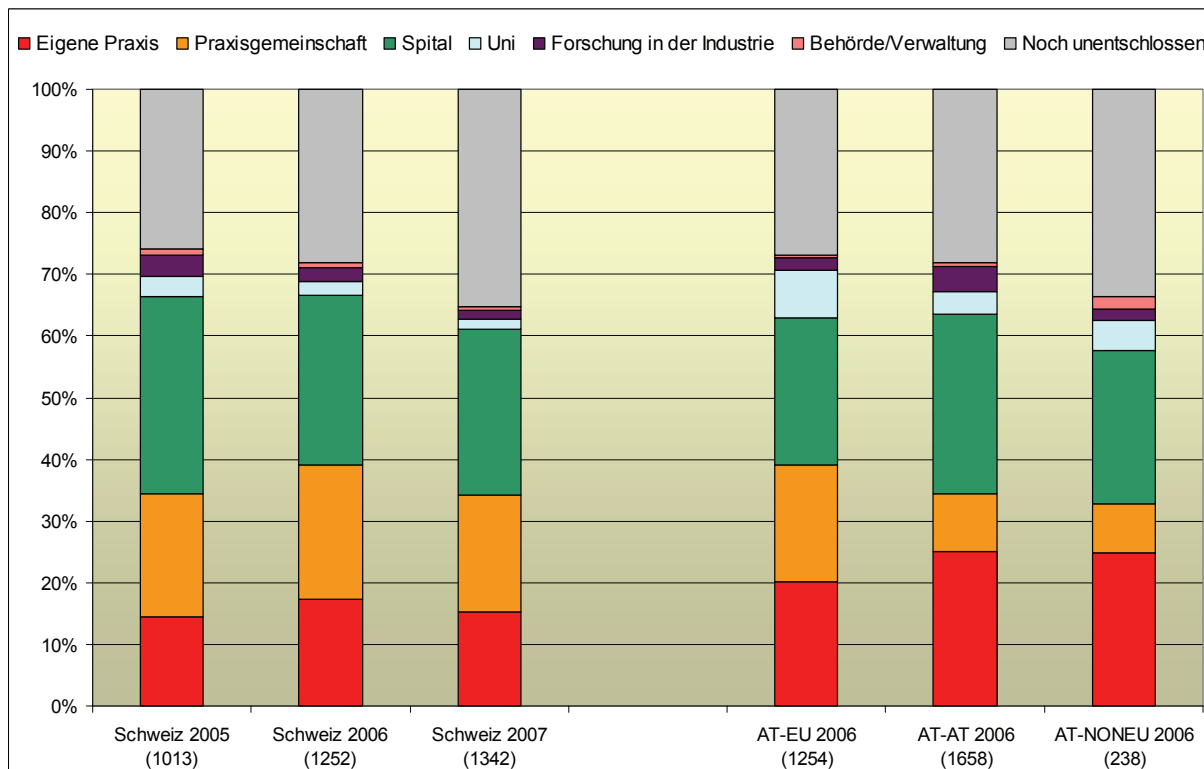


Abbildung 2: Ergebnisse der Befragungen zum zukünftig gewünschtem Arbeitsort (nur Bewerbungen Humanmedizin). Schweiz 2005 bis 2007 und Vergleichsdaten Österreich 2006 für die Quoten „EU-Bürger“ (AT-EU), „Bürger mit österreichischer Maturität“ (AT-AT) und „Nicht-EU-Bürger“ (AT-NONEU), in Klammern die Personenzahl.

Die Mehrzahl (rund 40%) möchte in jedem Jahr Spezialarzt für ein Fachgebiet werden. 2006 wurde die gleiche Befragung auch in Österreich durchgeführt, wo nur 7% Hausarzt und 50% Spezialarzt werden wollten, nur 14% sind unentschlossen. 25-30% der Bewerber sind in der Schweiz hinsichtlich der Berufsinteresse noch unentschlossen – hier geht es dann auch während des Studiums darum, das Interesse zu wecken.

Die Abbildung 2 verdeutlicht, dass auch der gewünschte Arbeitsort vor dem Studium in der Schweiz über die Jahre hinweg konstant bleibt. Kinder und Erwachsene als gewünschte Patientel halten sich in etwa die Waage, hier sind die meisten noch unentschlossen (Abbildung 3). Auch diese Ergebnisse stimmen grundsätzlich mit denen in Österreich überein.

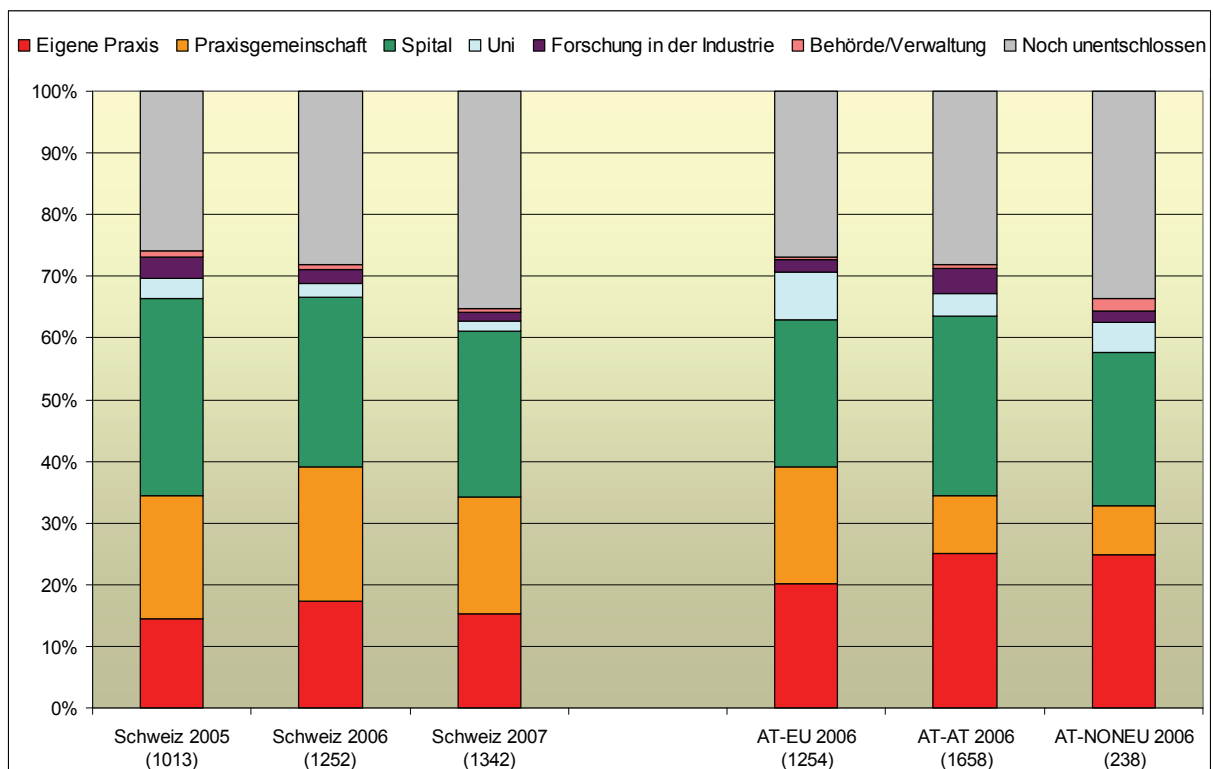


Abbildung 3: Ergebnisse der Befragungen zum zukünftig gewünschten Patientel (nur Bewerbungen Humanmedizin). Schweiz 2005 bis 2007 und Vergleichsdaten Österreich 2006 für die Quoten „EU-Bürger“ (AT-EU), „Bürger mit österreichischer Maturität“ (AT-AT) und „Nicht-EU-Bürger“ (AT-NONEU), in Klammern die Personenzahl.

Wenn man mit dem früheren „ganzheitlichen“ Fähigkeitskonzept der humanistischen Bildung argumentiert und dieses heute vermisst: Es ist nicht erwiesen, dass sich naturwissenschaftliche Begabung/Interessen und soziale Begabung/Interessen ausschliessen. Aus der Literatur ist eine hohe Korrelation von sozialer und allgemeiner Intelligenz bekannt.

Auch das fehlende **Interesse für Psychiatrie** scheint eher ein **Neigungsproblem** zu sein. Die „antipsychiatrische“ Diskussion lebt in der Öffentlichkeit momentan recht polemisch auf und der Berufsstand steht unter Kritik, ob es sich bei Psychiatrie um eine seriöse Wissenschaft handelt oder nicht. Wäre es nicht wichtiger, die Integration dieses Faches als vollwertiges Fach der Medizin zu betonen statt eine Trennung der Ausbildung vorzunehmen, die dann im Verdacht steht, auch eine Abgrenzung zu sein? Neigung und damit Nachfrage wird so nicht gefördert und es werden sich noch weniger Studierende für dieses Fach interessieren.

Immer wieder wird auch die Forderung gestellt, Elemente der **Berufseignung** im Sinne des „guten Arztes“ mit in den Zulassungstest aufzunehmen. Sehen wir einmal davon ab, dass die aktuellen gesetzlichen Grundlagen eine Beschränkung auf Studieneignung vorschreiben: Was bliebe offen? Die nicht-kognitiven Voraussetzungen oder „social skills“ (Empathie, soziale Kompetenz usw.) sind vermutlich für die einzelnen Fachgebiete ganz unterschiedlich

zu gewichten, sind wenig analysiert und auch durch Testverfahren nicht mit der gleichen Güte zu messen wie die kognitiven Voraussetzungen. Letztendlich muss der Medizinstudierende auch hier „nur“ bestimmte Fähigkeiten mitbringen. Die inhaltliche Ausformung und Differenzierung ist Aufgabe des Studiums, wobei eigene Neigungen und Interessen die Wahl des Fachgebietes dann mitbestimmen werden. Insofern wäre es vielleicht ausreichend, soziale Kompetenzen als soziale Fähigkeiten in geeigneter Weise mit zu erfassen und es dann der Ausbildung zu überlassen, dass diese Fähigkeiten auch umgesetzt werden.

Die Zeiten, als die Maturität ein exklusiver Abschluss war und allein dadurch die Universitäten keine Überlastungsprobleme hatten, sind vorbei. Eine Reduktion von Anforderungen, etwa der Verzicht auf naturwissenschaftliche Fähigkeiten in der Medizin, wäre das falsche Signal. Es scheint viel wahrscheinlicher, dass wir hierdurch auch Personen bevorzugt zulassen, die „ganzheitlich“ begabt sind. Man darf nicht vergessen, dass der EMS kein Wissenstest ist (wie viel man sich „allein in der Stube“ angelesen hat), sondern die **Fähigkeit zum Wissenserwerb** testet.

Es wäre vielen Menschen wichtig, dass der behandelnde Arzt auf der Basis bester biologisch-naturwissenschaftlicher Grundlagen und einschlägig fundierten Wissens handelt – und nicht nur empathisch-intuitiv oder begründet auf privater Erfahrung. **Die Fähigkeit zum Wissenserwerb braucht man lebenslang.**

4 Numerus clausus (NC) und Medizinstudium

Die Abbildungen 4 und 5 verdeutlichen die Anmeldeverläufe 1991 bis 2007. Weiterhin setzt sich die Zunahme der Bewerberzahlen in der Deutschschweiz nahezu „ungebremst“ fort. Man kann davon ausgehen, dass der Beruf des Mediziners nach wie vor als attraktiv und zukunftssicher erlebt wird. Auch in der Westschweiz hat es einen erneuten Anstieg der Anmeldezahlen gegeben. Es bleibt offen, wie lange die verschärfte inneruniversitäre Selektion nach dem ersten Studienjahr das Problem der Überlastung im ersten Studienabschnitt noch lösen kann.

Im Jahre 2007 hat es erstmals eine elektronische Anmeldung zum Medizinstudium bei der CRUS gegeben – bis zum Vorjahr erfolgte dies durch Einsendung eines Formulars. Die Papier-Anmeldeformulare mussten erst beschafft werden. Das überdeutliche Ansteigen der Anmeldezahlen 2007 und die bisher grössten Rückzüge bis zur Testanmeldung (siehe Kapitel 5.1 auf Seite 20) legen nahe, dass die „Schwelle“ für die elektronische Anmeldung wesentlich niedriger liegt. Trotz dieser Bedingungen ist die Zahl der Personen mit gültigem Testwert auch 2007 gegenüber den Vorjahren weiter gestiegen. Für 2008 soll das klassische Anmeldeverfahren erneut verwendet werden und daher wird ein leichter Rückgang bzw. ein Stagnieren der Anmeldezahlen im Februar auf dem Niveau von 2007 erwartet. Die Zahl der Bewerbungen mit gültigem Testwert wird eher weiter ansteigen. In der Abbildung 6 deutet sich für letztere Zahl an, dass der Anstieg aber bereits weniger deutlich als in den Vorjahren ist, weil ein Plateau wie bei den übrigen Studierendenzahlen erreicht wurde.

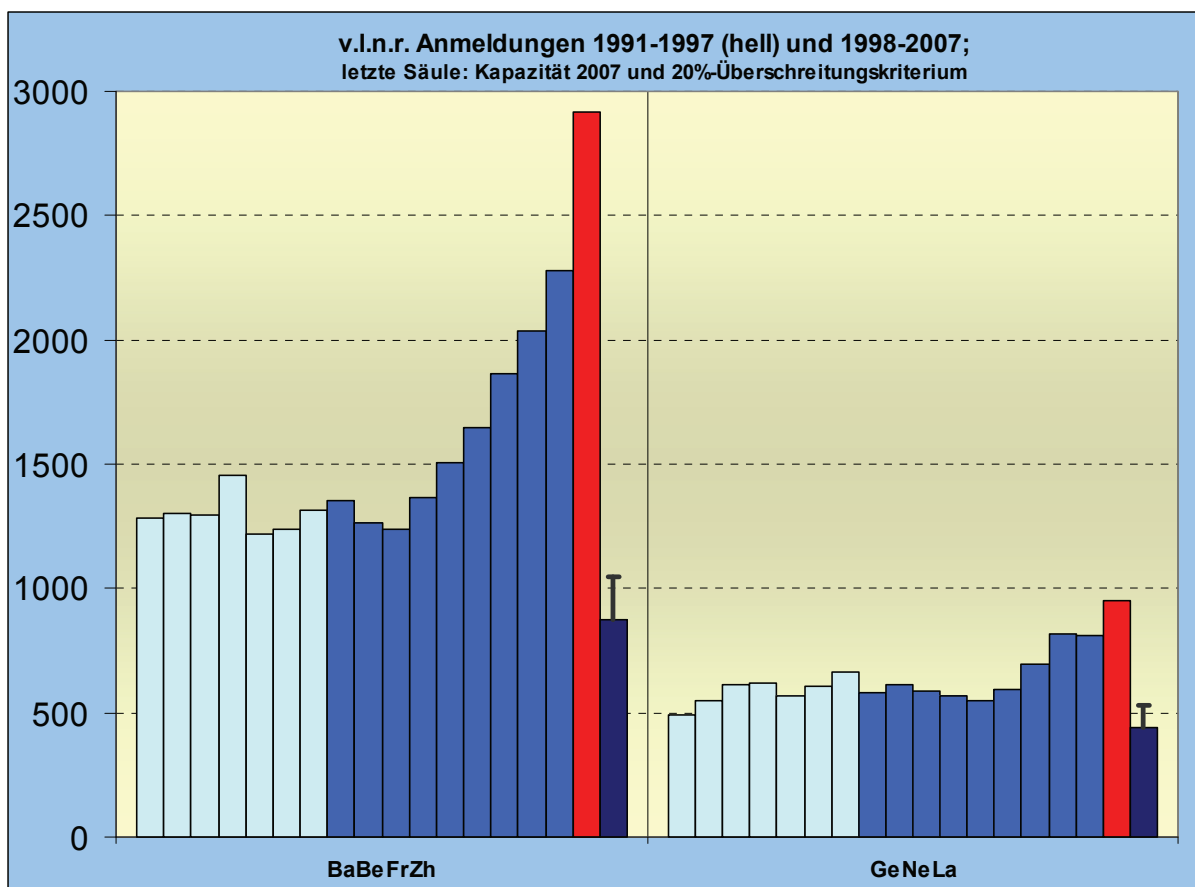


Abbildung 4: Anmeldungen zum Medizinstudium 1991 bis 2007 nach Gruppen (**B**asel, **B**ern, **F**reiburg, **Z**ürich vs. **G**enf, **N**euenburg, **L**ausanne). Hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule: Kapazität 2007 und Markierung 20%-Überschreitungskriterium als Grenze für NC.

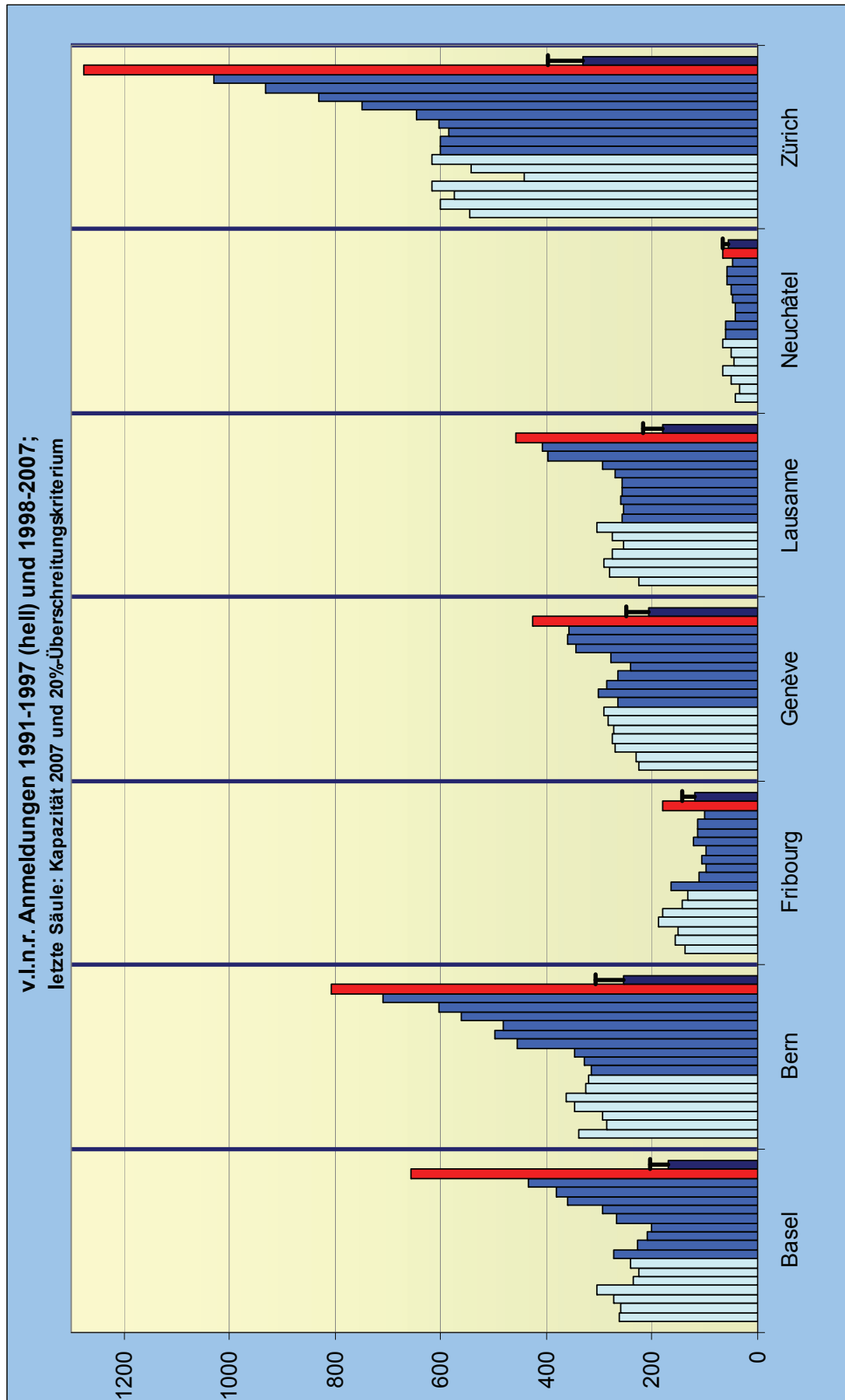


Abbildung 5: Anmeldezahlen pro Universität 1991 bis 2007 – hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule: Kapazität 2007 und Markierung 20%-Überschreitungskriterium als Grenze für NC.

Die Abbildung 6 zeigt auch, dass nicht mit einer Entspannung der Situation zu rechnen ist. Das Bundesamt für Statistik (BfS) veröffentlichte eine Prognose für die Studierendenzahlen der Universitäten und Hochschulen (alle Fächer der Schweiz). Wir verwenden den Stand September 2005 (Szenario „neutral HE-A-2004“). Setzt man die Zahl der Anmeldungen bei der CRUS (bzw. früher der SHK) im Jahre 1998 zu 100%, ergibt sich zunächst für 1999 und 2000 ein relativer Rückgang der Anmeldezahlen. Dass es sich hier um einen Dissuasionseffekt durch die Neueinführung des EMS handelt, wird durch das Gleichbleiben bzw. leichte Ansteigen der Studieneintritte 1999 und 2000 nahe gelegt. Auch wenn die Zahl der Studienplätze in einem realistischen Ausmass erhöht würde, würden auf lange Sicht deutlich mehr Bewerbungen erfolgen als Kapazitäten vorhanden wären. Es sollte damit gerechnet werden, dass ein Numerus clausus und damit der EMS noch für längere Zeit notwendig bleiben.

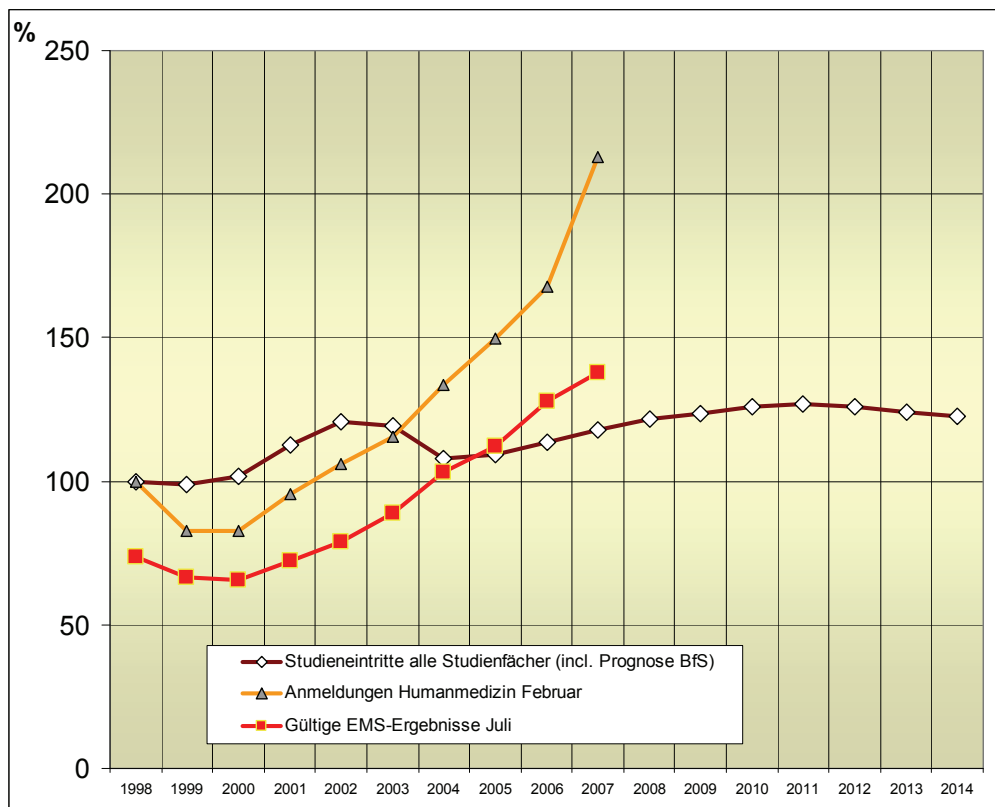


Abbildung 6: **Alle** Studieneintritte an Universitäten und Hochschulen (indexiert, 1998 = 100%, Stand September 2005, Szenario "neutral HE-A-2004") im Vergleich zu **Medizin**-Studienanmeldungen Februar und Bewerbungen mit gültigem EMS (indexiert, Anmeldungen 1998 = 100%) für Humanmedizin.

5 Anmeldung zum Medizinstudium und Test 2007

5.1 Anmeldungen, Testantritte und Zulassungen

	Humanmedizin									
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Kapazitäten (am NC beteiligte Univ.)	620	546	546	583	648	598	546	546	546	583
Anmeldungen bei CRUS im Februar	1020	845	844	974	1080	1177	1360	1525	1712	2171
... in % zu Kapazität	186	155	154	167	185	197	249	279	314	372
Anmeldungen zum EMS im Mai	797	717	678	752	827	917	1083	1182	1356	1452
+ Übernahme Testergebnis aus Vorjahr	0	5	33	27	40	47	38	41	45	41
= Total Bewerbungen Mai	797	722	711	779	867	964	1121	1223	1401	1493
... in % zu Kapazität	129	132	130	134	134	161	205	224	257	256
Absolvierung EMS 2007	750	673	637	709	770	871	1026	1107	1263	1374
Total Bewerbungen Juli - EMS und Ergebnis Vorjahr	750	678	670	736	805	907	1051	1143	1302	1405*
Rückzugsquote zwischen Februar und Juli (%)	26.5	19.8	20.6	24.4	25.5	22.9	22.7	25.0	23.9	34.8
Zugewiesene Studienplätze mit „Überbuchung“	669	631	652	672	758	705	624	640	653	685**
Abgewiesene Bewerbungen	81	42	14	58	47	202	426	503	649	720**
Prozent Bewerbungen, die Studienplatz erhalten	89	94	98	92	94	78	59	56	50	49**

Tabelle 1: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich für Humanmedizin (* 10 Personen wollen aus einem begonnenen Studium der Zahnmedizin in ein höheres Semester Humanmedizin wechseln. Sie nehmen am Test teil, zählen aber nicht als Bewerbungen um die Studienplätze 2007. Vergleichbare Fälle wurden hier auch in den Vorjahren entsprechend herausgerechnet; ** Stand Juli gemäss Modell.)

	Veterinärmedizin								
	1999	2000	2001	2002	2003	2004	2005	2006	2007
Kapazitäten (am NC beteiligte Univ.)	150	150	165	185	170	150	150	150	150
Anmeldungen bei CRUS im Februar	231	213	217	254	271	273	318	348	411
... in % zu Kapazität	153	142	132	137	159	182	212	232	274
Anmeldungen zum EMS im Mai	194	176	179	206	225	207	257	279	301
+ Übernahme Testergebnis aus Vorjahr	0	5	5	3	11	13	8	10	7
= Total Bewerbungen Mai	194	181	184	209	236	220	265	289	308
... in % zu Kapazität	129	121	112	113	139	147	177	193	205
Absolvierung EMS 2007	182	164	171	203	217	196	244	268	288
Total Bewerbungen Juli - EMS und Ergebnis Vorjahr	182	167	176	204	228	209	252	278	295
Rückzugsquote zwischen Februar und Juli (%)	21.2	20.7	18.9	19.3	15.9	23.4	20.8	20.1	28.2
Zugewiesene Studienplätze mit „Überbuchung“	160	167	176	202	194	175	180	170	173**
Abgewiesene Bewerbungen	22	0	0	2	34	34	72	108	122**
Prozent Bewerbungen, die Studienplatz erhalten	88	100	100	99	85	84	71	61	59**

Tabelle 2: Disziplinspezifische Statistiken für Veterinärmedizin und die hier am NC beteiligten Universitäten Bern und Zürich (In Freiburg und Basel wird diese Studienrichtung nicht angeboten; ** Stand Juli gemäss Modell.)

	Zahnmedizin			
	2004	2005	2006	2007
Kapazitäten (am NC beteiligte Univ.)	142	142	142	142
Anmeldungen bei CRUS im Februar	208	190	217	267
... in % zu Kapazität	146	134	153	188
Anmeldungen zum EMS im Mai	163	159	168	179
+ Übernahme Testergebnis aus Vorjahr	1	11	20	32
= Total Bewerbungen Mai	164	170	188	212
... in % zu Kapazität	115	120	132	149
Absolvierung EMS 2007	151	150	157	169
Total Bewerbungen Juli - EMS und Ergebnis Vorjahr	151	161	174	201
Rückzugsquote zwischen Februar und Juli (%)	26.9	15.2	19.3	24.7
Zugeteilte Studienplätze mit „Überbuchung“	151	161	172	165**
Abgewiesene Bewerbungen	0	0	2	36**
Prozent Bewerbungen, die Studienplatz erhalten	100	100	99	82**

Tabelle 3: **Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich für Zahnmedizin** (**Stand Juli gemäss Modell.)

„Modell“ heisst in den Tabellen 1 bis 3, dass mehr Personen einen Studienplatz erhalten als Kapazitäten vorhanden sind. Diese „Überbuchung“ beruht auf den Erfahrungen der Vorjahre und berücksichtigt alle Nichtantritte trotz Zulassung (z.B. weil bei Umleitungen der Studienort nicht zusagt). Ziel der Überbuchungen ist es, die Kapazitäten möglichst genau auszulasten und keine Nachrückverfahren notwendig werden zu lassen. Durch das Vorziehen des Semesterbeginns wird dies in Zukunft besonders wichtig. In den letzten Jahren konnten weitere Zuteilungsrunden zum Auslasten der Kapazität weitgehend vermieden werden. Diese würden die Zahl der abgewiesenen Personen verringern.

In Humanmedizin werden laut Modell erstmals mehr als 50% der Bewerber keinen Studienplatz zugewiesen bekommen. In Veterinärmedizin stabilisiert sich das Verhältnis aus dem Vorjahr. In Zahnmedizin treten erstmals nennenswerte Nichtzulassungen auf – hier konnte das Problem bis zum Vorjahr durch freiwillige Rückzüge bisher gelöst werden.

	Studienbeginn 2004			Studienbeginn 2005			Studienbeginn 2006			Studienbeginn 2007		
	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	299	234	21.7	350	267	23.7	386	275	28.8	565	340	39.8
Bern HM	357	277	22.4	396	299	24.5	458	376	17.9	515	357	30.7
Freiburg HM	107	89	16.8	105	82	21.9	94	61	35.1	161	107	33.5
Zürich HM	597	450	24.6	674	495	26.6	774	590	23.8	930	601	35.4
Total	1360	1050	22.8	1525	1143	25.0	1712	1302	23.9	2171	1405	35.3
Bern VM	137	111	19.0	152	121	20.4	187	143	23.5	195	137	29.7
Zürich VM	136	98	27.9	166	131	21.1	161	135	16.1	216	158	26.9
Total	273	209	23.4	318	252	20.8	348	278	20.1	411	295	28.2
Basel ZM	45	30	33.3	32	31	3.1	49	41	16.3	74	49	33.8
Bern ZM	65	52	20.0	56	47	16.1	66	59	10.6	73	62	15.1
Freiburg ZM	7	6	14.3	9	7	22.2	7	6	14.3	14	9	35.7
Zürich ZM	91	63	30.8	93	76	18.3	95	68	28.4	106	81	23.6
Total	208	151	27.4	190	161	15.3	217	174	19.8	267	201	24.7

Tabelle 4: Anmelungsverlauf für Universitäten, Human- (HM), Veterinär- (VM) und Zahnmedizin (ZM). Personen mit gültigem Testwert. Vermehrte Rückzüge 2007 vermutlich wegen neu eingeführtem elektronischem Anmeldeverfahren.

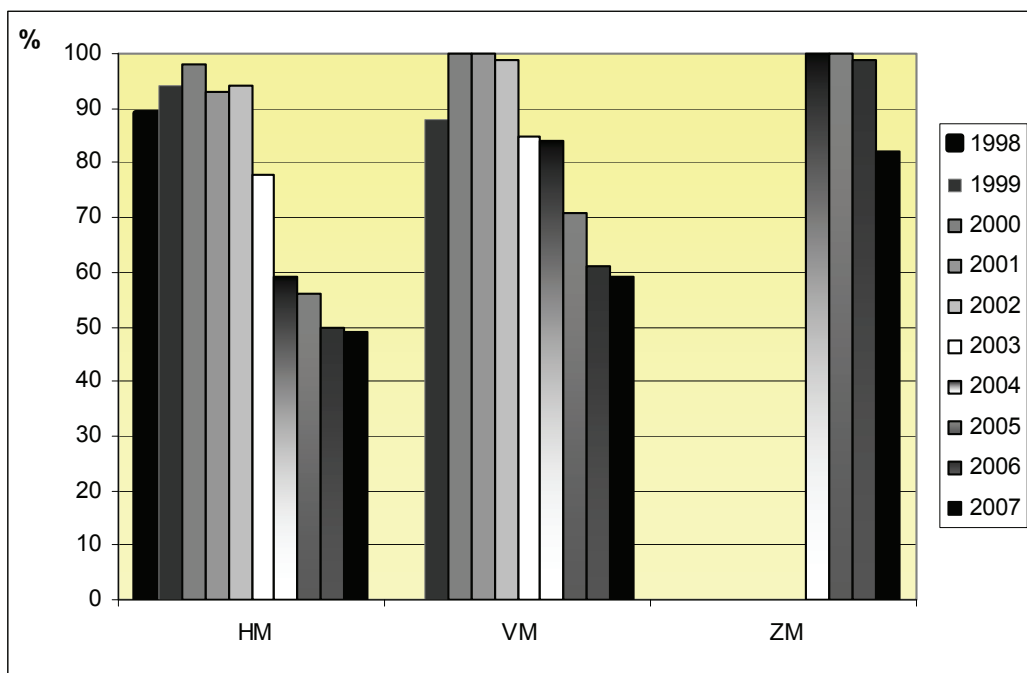


Abbildung 7: Prozent der Personen mit gültigem Testwert, denen ein Studienplatz zugewiesen werden kann (ohne Überbuchung – HM: Humanmedizin; VM: Veterinärmedizin; ZM: Zahnmedizin).

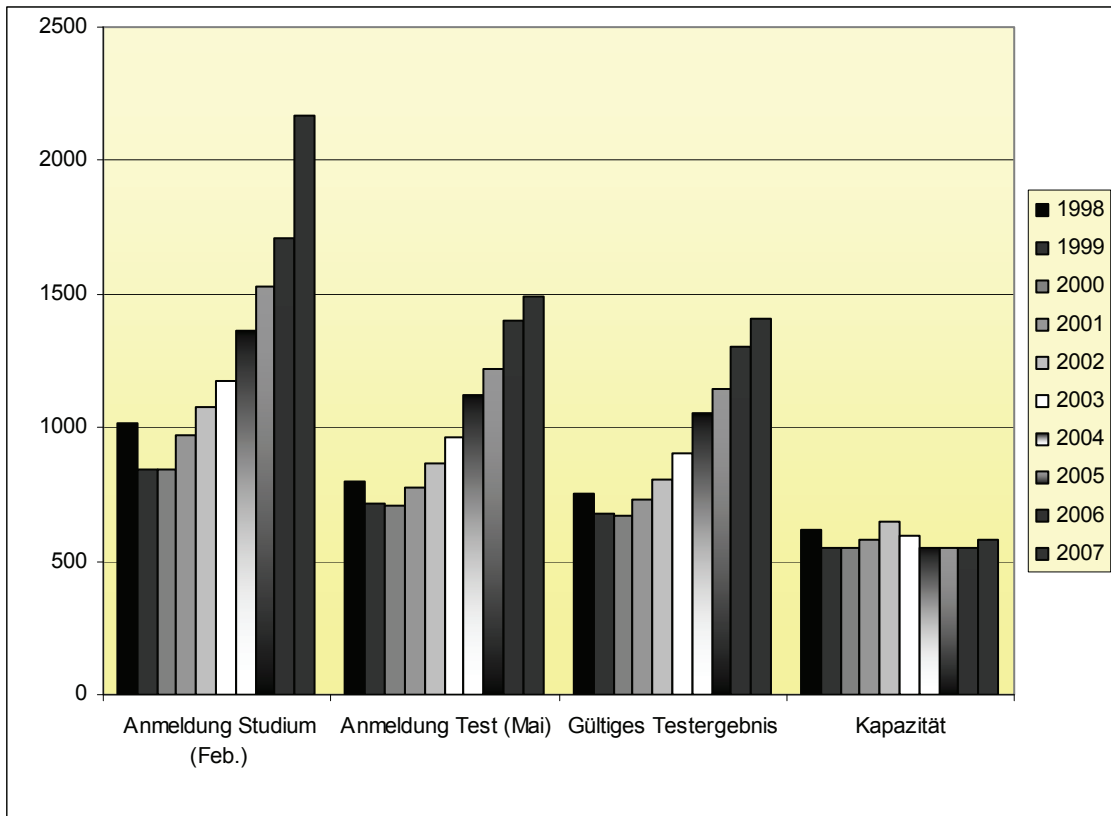


Abbildung 8: Anmeldeverlauf für Humanmedizin 1998 bis 2007 im Vergleich (mit NC).

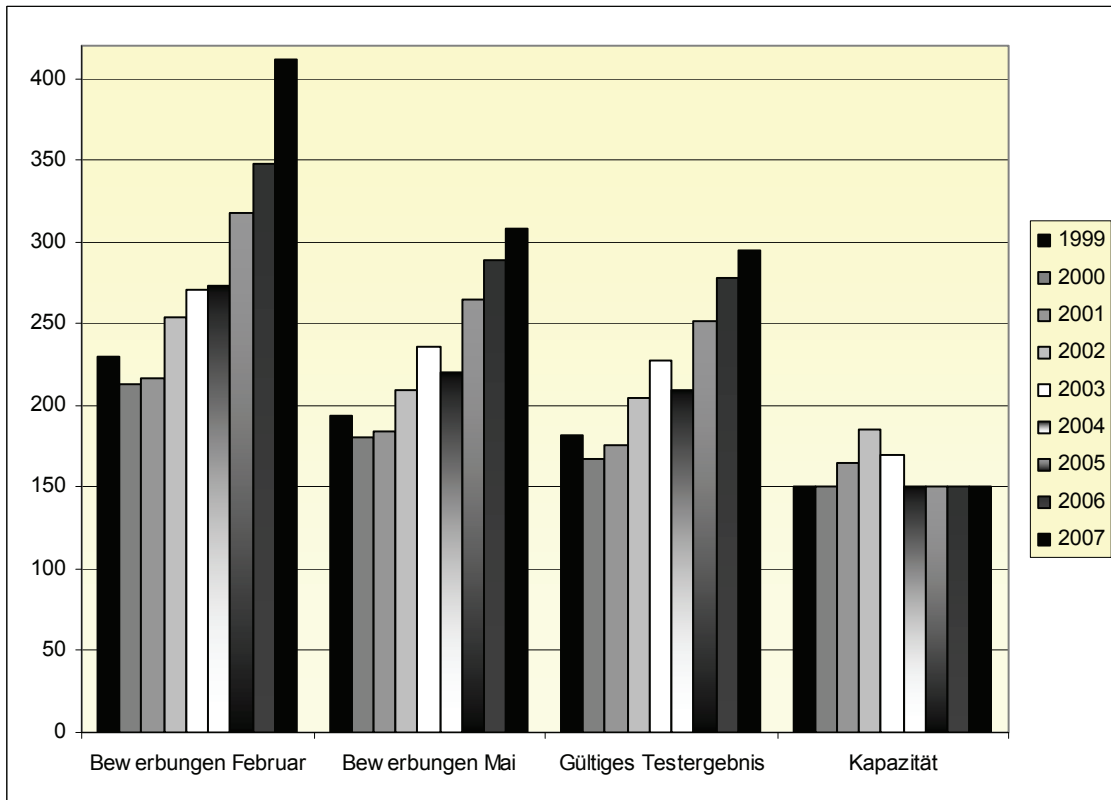


Abbildung 9: Anmeldeverlauf für Veterinärmedizin 1999 bis 2007 im Vergleich (mit NC).

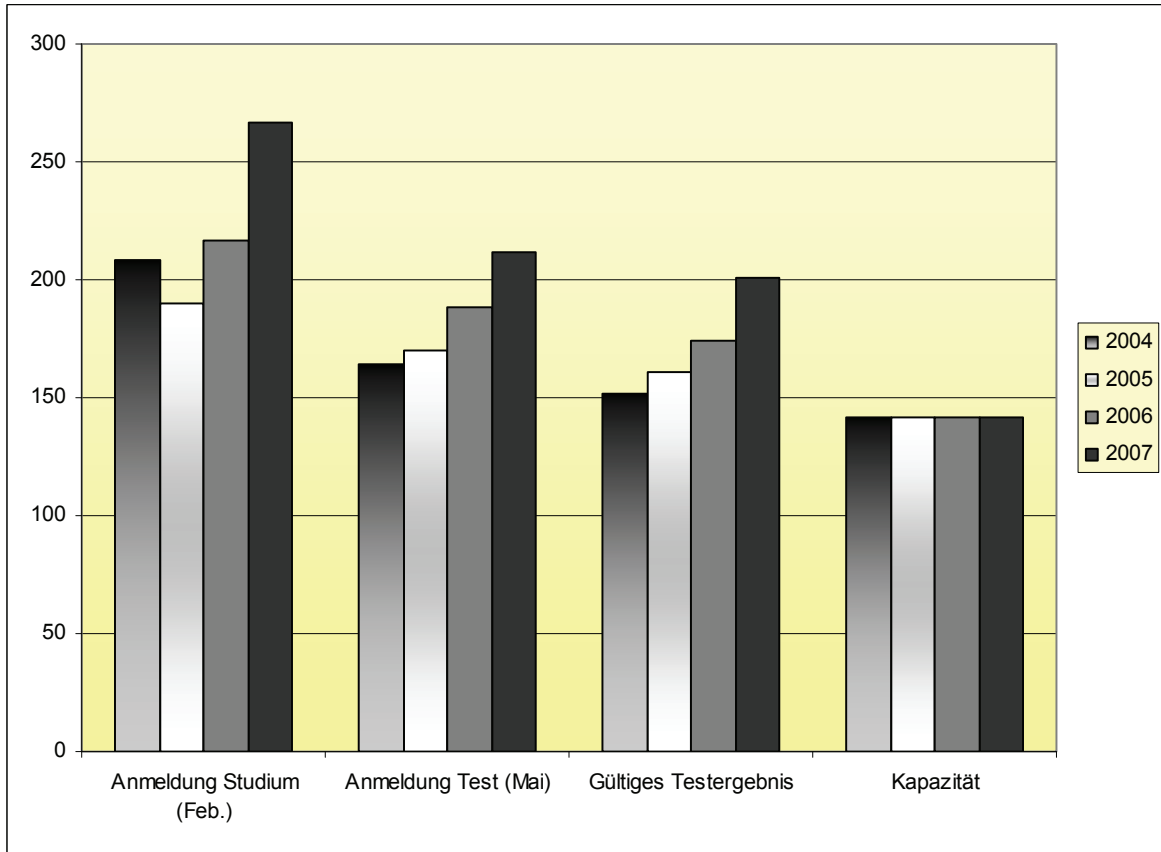


Abbildung 10: Anmeldeverlauf für Zahnmedizin 2004 bis 2007 im Vergleich (mit NC).

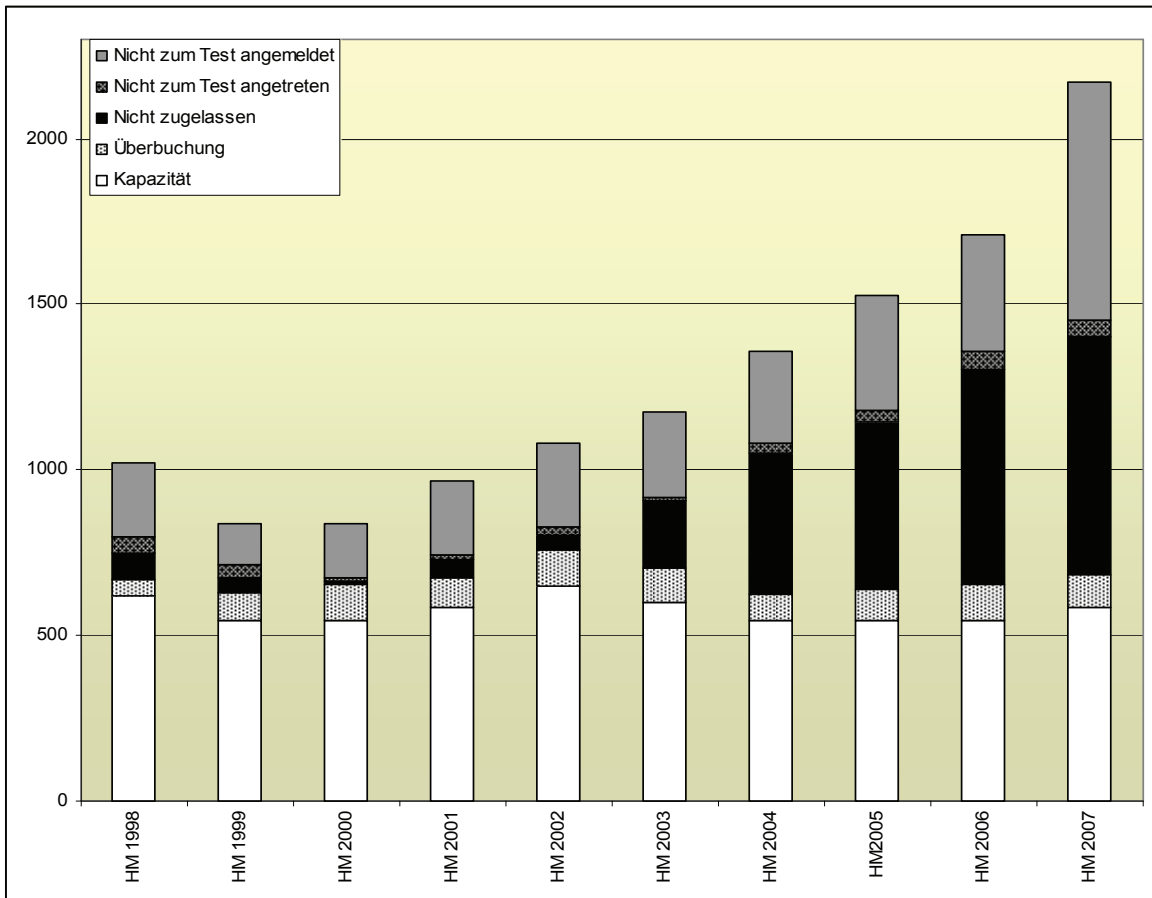


Abbildung 11: Aufteilung der Voranmeldungen (Februar 2007) für Humanmedizin (HM). Zugelassen wurden die Personen entsprechend Kapazität plus Überbuchung.

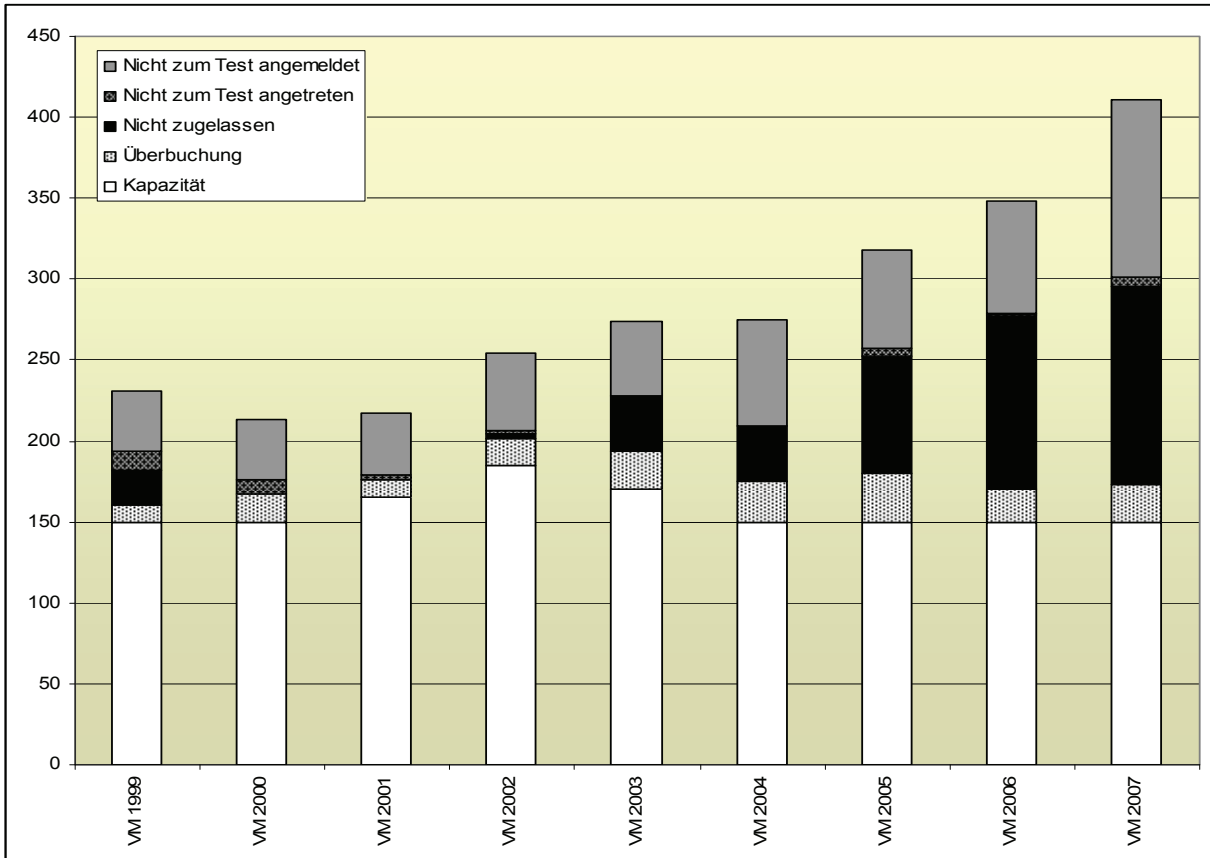


Abbildung 12: Aufteilung der Voranmeldungen (Februar 2007) für Veterinärmedizin (VM). Zugelassen wurden die Personen entsprechend Kapazität plus Überbuchung.

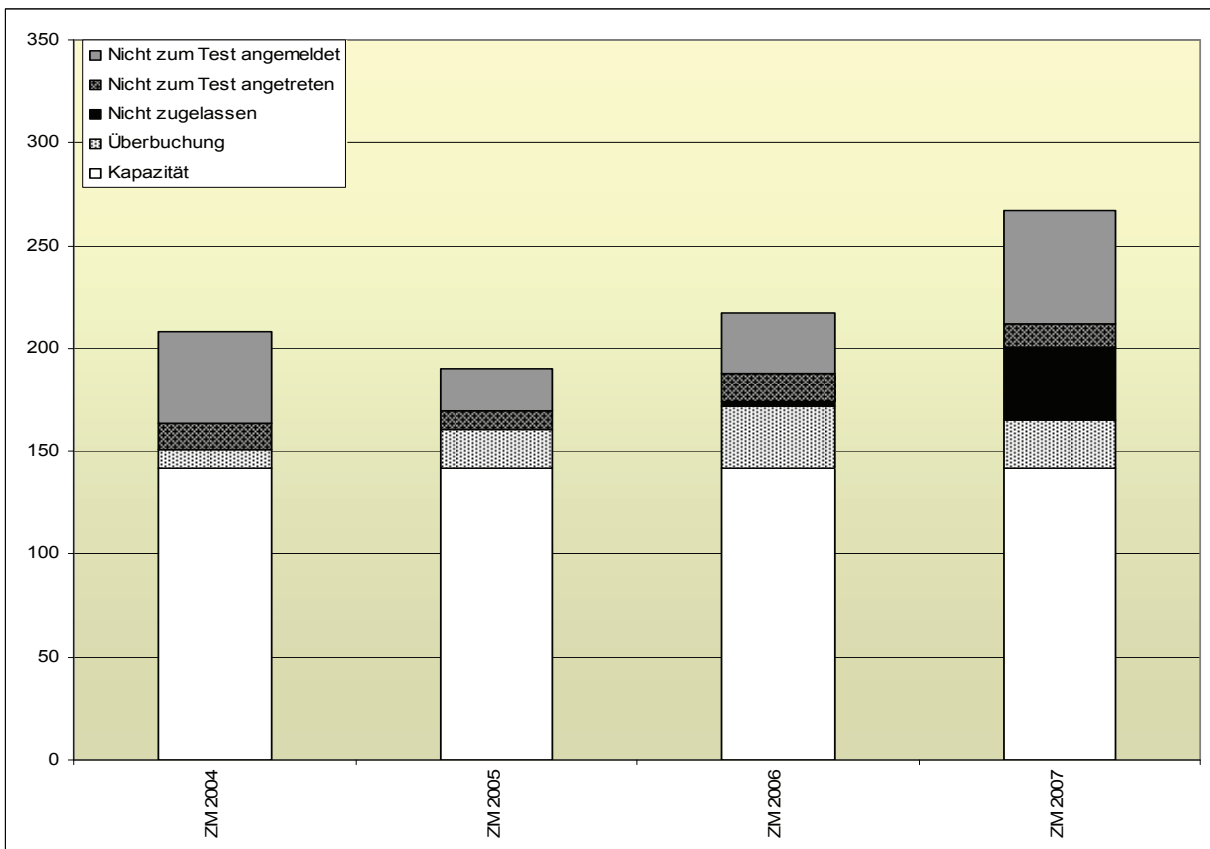


Abbildung 13: Aufteilung der Voranmeldungen (Februar 2007) für Zahnmedizin (ZM). Zugelassen wurden die Personen entsprechend Kapazität plus Überbuchung.

5.2 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme in 21 Testlokalen verteilt auf 8 Testorte.

Die Grösse der Testlokale variierte im bisher üblichen Rahmen. Die Zahl der Betreuer war proportional zur Teilnehmerzahl vorgegeben. Es gibt keine Hinweise seitens des Betreuungspersonals oder durch 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).

Interessant ist in diesem Zusammenhang ein Befund aus Innsbruck aus dem Jahre 2006, wo sich auch für sehr grosse Testlokale (350 Personen) kein Unterschied beim Testergebnis zu kleineren Testlokalen (70 Personen) ergab. Die in der Schweiz als gross geltenden Testlokale erreichen noch nicht die Grenze, ab der man nachteilige Effekte erwarten müsste.

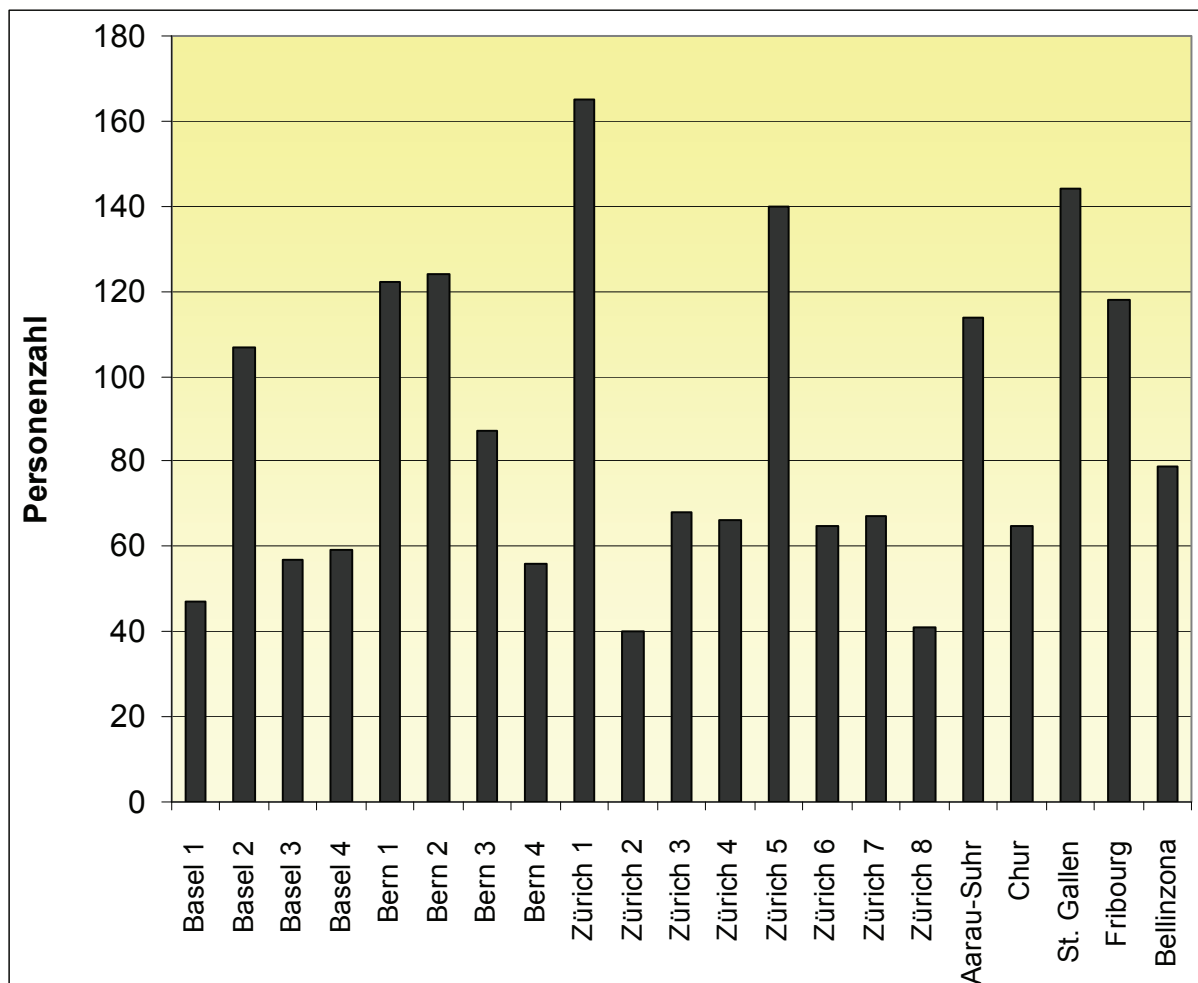


Abbildung 14: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2007.

5.3 Testorte, Wunschuniversitäten und Anreise

Die Wahl des Testortes hat bei der Zulassung keinen Einfluss auf den zukünftigen Studienort. Es wird empfohlen, denjenigen Testort zu wählen, der dem Wohnort am nächsten liegt. Längere Anreisen könnten sich negativ auf die Leistungsfähigkeit auswirken. Da der EMS gewollt auch eine Belastungssituation darstellt, spielen optimale äussere Bedingungen vermutlich auch eine grosse Rolle, um am Testtag die optimal möglichen Leistungen zu erreichen.

Um dies weiter abzuklären, wurde dieses Jahr nach der Testdurchführung erfragt, wann die Anreise zum EMS erfolgte. 40% der Testteilnehmer wohnen am Testort, 23 Personen sind mehr als einen Tag vorher und 74 Personen einen Tag vorher angereist – diese 6% werden nachfolgend zusammengefasst. 54% der Testteilnehmer sind am Testtag morgens angereist.

Testwert	Mittelwert	Standardabweichung	Anzahl
wohnhaft am Testort	100.6	10.0	682
einen/mehrere Tage vorher angereist	97.4	9.8	97
... davon nicht aus dem Ausland (Schweizerbürger)	99.0	9.2	80
am Testtag angereist	100.0	9.8	913

Tabelle 5: Anreise und Testwert 2007

Testort	BASEL	BELLIN-ZONA	BERN	CHUR	FRI-BOURG	ST. GALLEN	SUHR	ZÜRICH
wohnhaft am Testort	174	32	81	18	35	54	29	259
einen/mehrere Tage vorher angereist	10	7	21	6	19	2	0	32
am Testtag angereist	67	38	194	41	60	85	83	345

Tabelle 6: Anreise nach Testort 2007, Anzahl der Testteilnehmer je Gruppe

Tagesform	viel schlechter als sonst	schlechter als sonst	genau wie sonst	besser als sonst	viel besser als sonst	Gesamt
wohnhaft am Testort	48 7.1%	311 45.9%	287 42.4%	29 4.3%	2 .3%	677
einen/mehrere Tage vorher angereist	9 9.3%	52 53.6%	32 33.0%	3 3.1%	1 1.0%	97
... davon nicht aus dem Ausland	7 8.8%	45 56.3%	25 31.3%	2 2.5%	1 1.3%	80
am Testtag angereist	50 5.5%	449 49.4%	382 42.0%	26 2.9%	2 .2%	909

Tabelle 7: Anreise und Tagesform 2007, Anzahl und Prozentaufteilung zeilenweise

Der Mittelwert des Testwertes für die beiden **grossen** Gruppen „wohnhaft am Testort“ und „am Testtag angereist“ unterscheidet sich nicht signifikant. Auch die Tagesform, die in 5 Stufen erfragt wurde, unterscheidet sich für diese Gruppen nicht wesentlich. Dies ist ein

wichtiger Befund – die Testlokale liegen so verteilt, dass sie ohne Hinweis auf Nachteile von der Mehrheit erreichbar sind.

Die mit rund 6% **recht kleine** Gruppe mit Anreise ein oder mehrere Tage früher unterscheidet sich allerdings signifikant davon, es werden schlechtere Testleistungen erreicht. Nimmt man hier nur die Schweizerbürger, ist der Unterschied allerdings nur noch 1 Punkt und nicht mehr signifikant. Interessant ist, dass die Schweizerbürger, die einen oder mehr Tage vorher angereist sind, sich tendenziell von der Tagesform her am schlechtesten gefühlt haben (Tabelle 7). Es scheint also eine Rolle zu spielen, ob man auswärtig übernachtet oder übernachten muss – eventuell in ungewohnter Umgebung. Der subjektiv erlebte Einfluss für die Schweizerbürger scheint dabei deutlicher zu sein als die tatsächliche Auswirkung auf den Testwert.

Den grössten Anteil hat der Kanton Wallis mit 12 Personen (Total von 31 kantonalen Testteilnehmern). Hier könnte sich die Verbesserung der Verkehrsanbindung nach Bern positiv auswirken, um erst am Testtag anzureisen. Den zweitgrössten Anteil (11 Personen) hat der Kanton Luzern in dieser Gruppe (Total von 123 kantonalen Testteilnehmern). Hier wird 2008 ein neuer Testort eingerichtet. Inwieweit regionale Effekte diese Unterschiede überlagern, kann aufgrund der geringen Gruppengrösse nicht weiter analysiert werden. Wir veröffentlichen die Ergebnisse vor allem deshalb, um auf die Bedeutung dieses Faktors hinzuweisen, damit man sich im Zweifel die optimalen Bedingungen schafft.

Wie in den Vorjahren zeigen sich Übereinstimmungen zwischen Test- und Wunschstudienort, da die Wahl der Universität häufig nach der Wohnortnähe erfolgt. Bekanntlich kann auch eine Umleitung an eine andere Universität 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
BASEL	24 4.5%	227 59.3%	8 7.1%	11 1.4%	270 14.7%
BELLINZONA	7 1.3%	11 2.9%	15 13.3%	46 5.7%	79 4.3%
BERN	339 64.2%	18 4.7%	22 19.5%	10 1.2%	389 21.2%
CHUR	13 2.5%	12 3.1%	3 2.7%	37 4.6%	65 3.5%
FRIBOURG	58 11.0%	4 1.0%	51 45.1%	5 .6%	118 6.4%
ST. GALLEN	17 3.2%	26 6.8%	4 3.5%	97 12.0%	144 7.9%
SUHR	28 5.3%	49 12.8%	1 .9%	36 4.5%	114 6.2%
ZÜRICH	42 8.0%	36 9.4%	9 8.0%	565 70.0%	652 35.6%
Total	528	383	113	807	1831

Tabelle 8: Testorte und Wunschuniversität (erste Wahl) 2007

5.4 Wunschuniversität und Testort nach Wohnkanton

In der folgenden Tabelle sind die Testorte nach Wohnkanton für die einzelnen Disziplinen aufgeschlüsselt. Angegeben sind auch die Vergleichswerte seit 1998 bzw. für die Jahre mit NC.

Wohnkanton/ Wohnort	Humanmedizin 2007: gewünschter Studienort					Total									
	BE	BS	FR	ZH	Total	2006	2005	2004	2003	2002	2001	2000	1999	1998	
AG	21	57	2	44	124	99	81	93	54	52	42	62	62	66	
AI	0	1	0	2	3	4	1	4	0	2	1	3	0	2	
AR	4	2	1	5	12	7	4	8	6	6	8	4	6	4	
BE	189	6	4	5	204	206	174	179	135	139	148	88	87	101	
BL	2	77	0	2	81	56	64	54	44	37	41	38	52	50	
BS	2	73	1	3	79	77	61	48	41	35	31	26	37	40	
FR	7	0	39	0	46	45	54	42	44	28	24	33	39	34	
GE	0	1	1	0	2	1	0	0	1	1	1	0	0	1	
GL	0	0	0	1	1	7	5	1	6	3	3	3	4	4	
GR	7	9	3	25	44	40	34	41	46	18	30	16	22	19	
JU	0	2	2	1	5	1	0	3	1	0	0	0	1	0	
LU	43	29	5	21	98	70	65	53	52	62	42	29	34	43	
NE	0	0	1	0	1	0	3	1	1	1	0	0	1	2	
NW	3	5	0	4	12	2	4	5	9	4	3	2	3	6	
OW	1	1	0	3	5	6	5	5	2	7	3	9	2	3	
SG	13	24	3	71	111	86	81	55	48	50	41	51	51	64	
SH	4	1	1	9	15	9	11	6	10	10	15	5	4	7	
SO	11	23	1	2	37	55	42	42	38	35	21	13	22	26	
SZ	5	2	3	16	26	17	19	18	11	11	15	16	16	21	
TG	3	4	2	31	40	42	38	24	22	28	30	20	16	22	
TI	4	10	14	22	50	45	26	34	30	18	31	21	23	30	
UR	6	1	0	2	9	8	3	1	3	4	4	2	1	3	
VD	1	0	8	1	10	7	3	1	2	1	2	2	0	3	
VS	9	2	7	4	22	20	10	7	5	13	12	11	12	22	
ZG	2	2	0	23	27	30	20	17	12	16	19	6	4	6	
ZH	6	3	3	271	283	295	270	263	234	178	130	169	159	165	
FL	1	1	0	3	5	7	9	5	3	3	5	4	2	3	
Übriges Ausland	1	7	3	11	22	21	20	16	11	8	7	4	13	3	
Total	345	343	104	582	1374	1263	1107	1026	871	770	709	637	673	750	

Tabelle 9: Gewünschte Studienorte Humanmedizin nach Wohnkantonen 2007 und Vergleich mit 1998 bis 2006. Das Fürstentum Liechtenstein (FL) wird als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wohnkanton/ Wohnort	Veterinärmedizin 2007: gewünschter Studienort			Total							
	Bern	Zürich	Total	2006	2005	2004	2003	2002	2001	2000	1999
AG	3	13	16	27	25	11	13	8	7	9	22
AI	0	1	1	0	0	1	0	3	0	1	0
AR	0	1	1	2	2	2	2	0	2	0	0
BE	34	1	35	38	24	29	32	30	39	26	26
BL	11	4	15	7	10	12	8	7	4	3	7
BS	4	3	7	1	9	4	2	4	7	2	2
FR	9	1	10	8	9	6	8	9	3	11	9
GE	12	0	12	13	6	2	4	6	9	11	13
GL	0	2	2	2	4	0	1	1	0	2	0
GR	2	7	9	10	6	10	13	4	6	3	7
JU	8	0	8	7	0	4	4	2	1	1	1
LU	10	6	16	14	6	11	7	13	5	12	10
NE	6	0	6	6	10	4	7	6	3	6	4
NW	1	2	3	3	2	1	2	2	0	0	0
OW	1	0	1	1	0	0	0	0	0	0	0
SG	1	19	20	12	19	14	13	9	7	4	10
SH	0	6	6	1	1	1	2	0	3	3	0
SO	5	1	6	10	7	7	7	8	4	7	6
SZ	0	5	5	6	2	4	2	4	0	2	2
TG	0	12	12	12	8	5	11	6	8	3	4
TI	2	9	11	8	10	6	9	9	7	6	6
UR	0	1	1	1	1	2	1	1	1	0	0
VD	16	3	19	25	31	15	15	13	21	15	11
VS	4	1	5	7	5	7	4	7	7	7	3
ZG	0	5	5	7	3	4	6	3	4	3	6
ZH	1	49	50	36	39	29	42	46	19	25	32
FL	0	0	0	1	0	0	0	1	2	0	0
Übriges Ausland	2	4	6	3	5	1	2	1	2	2	1
Total	132	156	288	268	244	196	217	203	171	164	182

Tabelle 10: Gewünschte Studienorte Veterinärmedizin nach Wohnkantonen 2007 und Vergleich mit 1999 bis 2006. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Wohnkanton / Wohnort	Zahnmedizin 2007: gewünschter Studienort					Total		
	BE	BS	FR	ZH	Total	2006	2005	2004
AG	2	6	0	3	11	14	10	14
AI	0	1	0	0	1	0	1	0
AR	0	0	0	1	1	2	2	4
BE	30	0	0	1	31	32	27	31
BL	0	11	0	0	11	5	11	10
BS	1	4	0	0	5	7	6	4
FR	2	0	4	0	6	3	4	1
GL	0	1	0	1	2	1	3	0
GR	2	2	0	2	6	5	3	8
JU	0	0	2	0	2	0	0	0
LU	1	2	0	6	9	10	5	2
NE	1	0	0	0	1	1	0	0
NW	0	0	0	0	0	0	1	1
OW	0	0	0	0	0	0	0	2
SG	3	2	0	12	17	13	12	4
SH	0	0	0	0	0	0	5	3
SO	2	8	1	2	13	10	2	9
SZ	1	1	0	1	3	4	1	2
TG	0	0	0	2	2	4	9	6
TI	0	0	0	10	10	4	3	8
UR	1	1	0	1	3	2	2	0
VS	4	0	0	0	4	5	4	4
ZG	0	0	0	2	2	2	1	1
ZH	0	0	2	23	25	28	33	33
FL	1	0	0	0	1	1	2	2
Übriges Ausland	0	1	0	2	3	4	3	2
Total	51	40	9	69	169	157	150	151

Tabelle 11: Gewünschte Studienorte Zahnmedizin nach Wohnkantonen 2007 und Vergleich mit 2004 bis 2006. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Die nachfolgende Tabelle zeigt, welche Testorte aus den einzelnen Kantonen (für alle Disziplinen zusammengefasst) gewählt werden und dient auch der Kapazitätsplanung für die Testorte. Ein zukünftiger Testort Luzern ist in Planung. Er wird nach dieser Aufteilung vor allem Zürich und Bern entlasten, was beabsichtigt ist.

Wohnkanton	Testort								
	BASEL	BELLIN-ZONA	BERN	CHUR	FRIBOURG	ST. GALLEN	SUHR	ZÜRICH	Total
AG	23	0	1	0	0	0	80	47	151
AI	1	0	0	0	0	4	0	0	5
AR	0	0	0	0	0	14	0	0	14
BE	3	0	257	0	5	0	0	5	270
BL	103	0	1	2	1	0	0	0	107
BS	86	0	1	0	0	0	1	3	91
FR	1	0	21	0	40	0	0	0	62
GE	1	0	2	0	11	0	0	0	14
GL	1	0	0	0	0	0	0	4	5
GR	1	4	1	47	0	0	0	6	59
JU	0	0	0	0	15	0	0	0	15
LU	13	0	35	1	1	1	22	50	123
NE	0	0	0	0	8	0	0	0	8
NW	2	0	4	0	0	0	1	8	15
OW	0	0	3	0	0	0	1	2	6
SG	3	0	1	13	0	104	0	27	148
SH	0	0	1	0	0	0	0	20	21
SO	18	0	26	0	0	0	8	4	56
SZ	2	1	3	0	0	0	0	28	34
TG	3	0	1	0	0	17	0	33	54
TI	0	70	0	0	1	0	0	0	71
UR	0	0	3	0	0	0	0	10	13
VD	0	0	1	0	26	0	0	2	29
VS	0	0	22	0	7	0	0	2	31
ZG	0	0	0	0	0	0	0	34	34
ZH	2	2	1	0	0	1	1	351	358
FL	0	0	1	2	0	2	0	1	6
Übriges Ausland	7	2	3	0	3	1	0	15	31
Total	270	79	389	65	118	144	114	652	1831

Tabelle 12: Testorte nach Wohnkantonen 2007. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

5.5 Testabsolvierung nach Alter und Geschlecht

Disziplin	Geburtsjahre	Geschlecht				Gesamt
		männlich		weiblich		
		Anzahl	Anzahl der Zeilen (%)	Anzahl	Anzahl der Zeilen (%)	
Humanmedizin	1954-1983	54	51.9%	50	48.1%	104
	1984-1986	153	48.7%	161	51.3%	314
	1987-1990	345	36.1%	611	63.9%	956
	Gesamt	552	40.2%	822	59.8%	1374
Veterinärmedizin	1954-1983	6	18.8%	26	81.3%	32
	1984-1986	11	16.7%	55	83.3%	66
	1987-1990	34	17.9%	156	82.1%	190
	Gesamt	51	17.7%	237	82.3%	288
Zahnmedizin	1954-1983	15	60.0%	10	40.0%	25
	1984-1986	22	44.9%	27	55.1%	49
	1987-1990	29	30.5%	66	69.5%	95
	Gesamt	66	39.1%	103	60.9%	169
Gesamt	1954-1983	75	46.6%	86	53.4%	161
	1984-1986	186	43.4%	243	56.6%	429
	1987-1990	408	32.9%	833	67.1%	1241
	Gesamt	669	36.5%	1162	63.5%	1831

Tabelle 13: Testabsolventen nach Geburtsjahr und Geschlecht 2007.

Der Männeranteil in Humanmedizin stabilisiert sich bei 40%, der Zuwachs war in diesem Jahr für beide Geschlechter etwa gleich. In Veterinärmedizin bleibt der Geschlechteranteil bei über 80% Frauen. In Zahnmedizin ist das Geschlechterverhältnis wieder auf dem Niveau von 2005 und 2006 – bei ebenfalls einem Frauenanteil von 60%; ebenso in Humanmedizin.

Betrachtet man gleichzeitig das Alter und das Geschlecht, kommt der Zuwachs bei den Bewerbungen vor allem aus dem Kreis der Jüngeren. Es gibt bisher keine Hinweise darauf, dass sich ein Kreis von „Dauerbewerbern“ bildet.

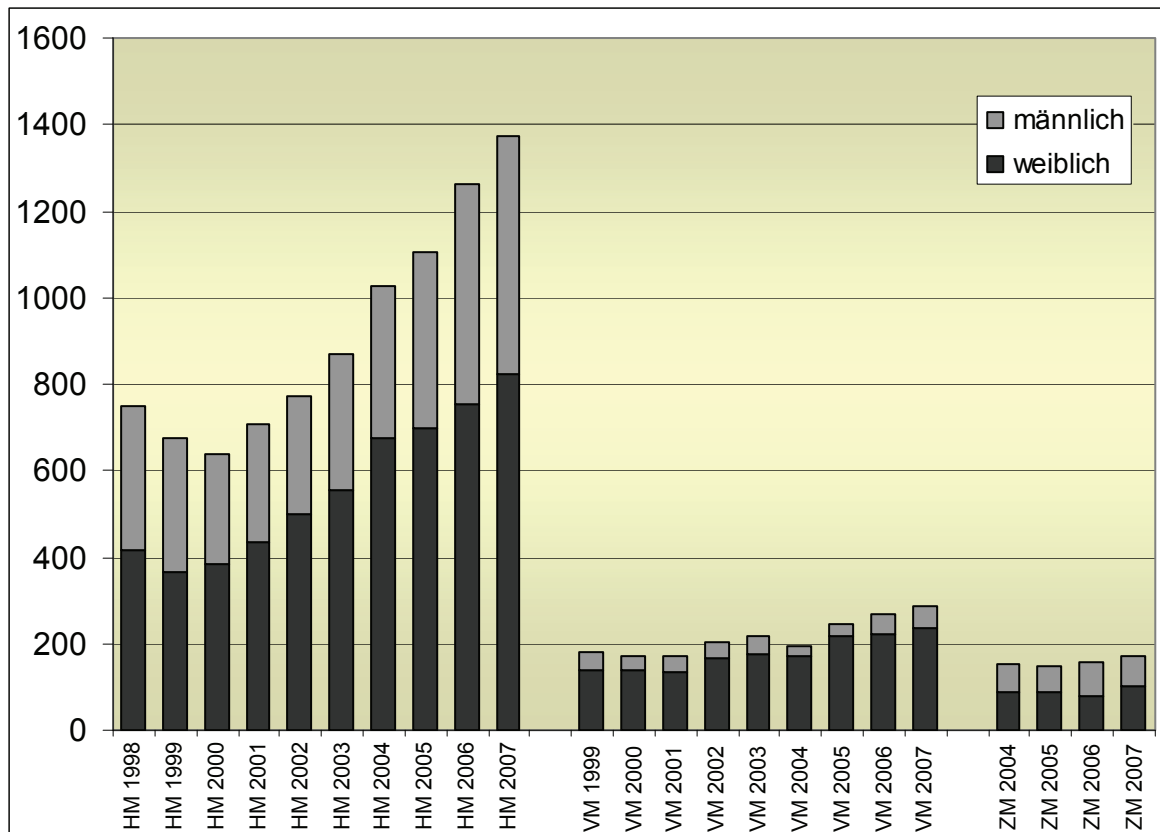


Abbildung 15: Anzahl männlicher und weiblicher Bewerber für Humanmedizin (HM), Veterinärmedizin (VM) und Zahnmedizin (ZM), Jahre mit NC.

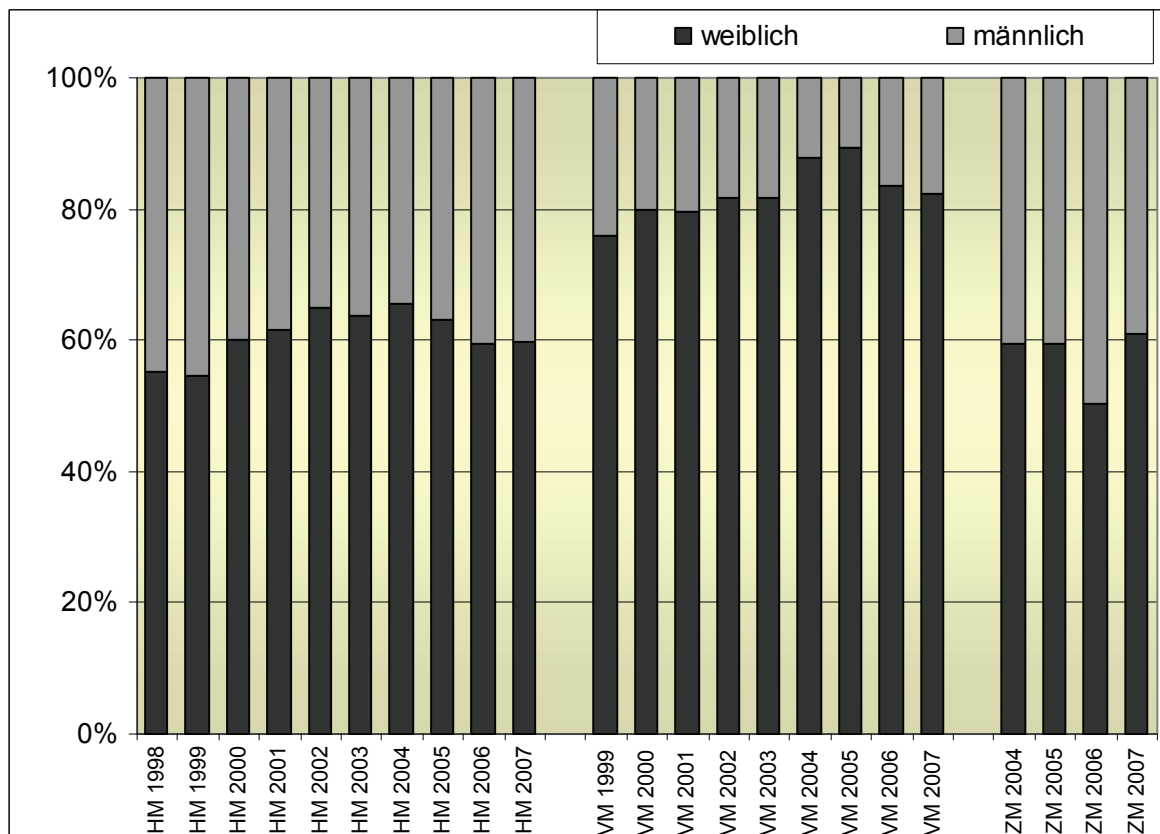


Abbildung 16: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (HM), Veterinärmedizin (VM) und Zahnmedizin (ZM), Jahre mit NC, bezogen auf 100% pro Jahr und Disziplin.

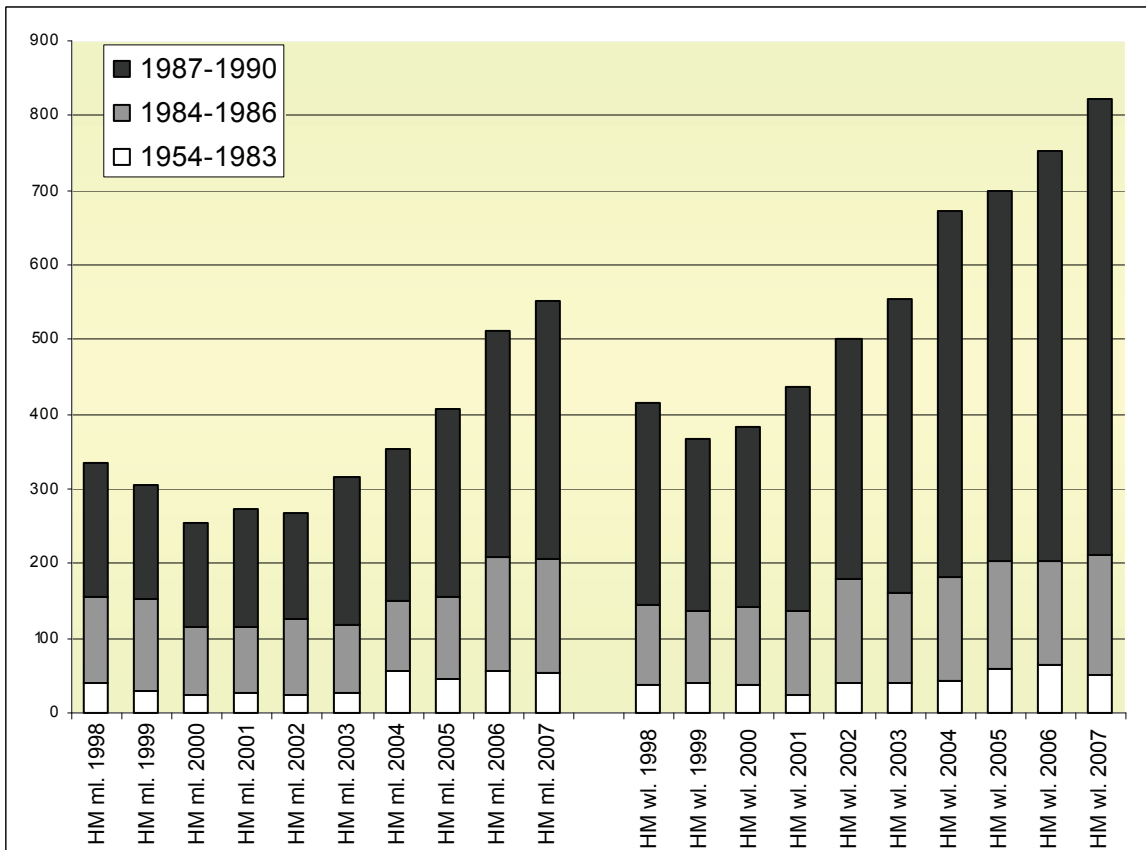


Abbildung 17: Zusammensetzung der Bewerbergruppen nach Geschlecht (mi: männlich, wl: weiblich) und Alter (Geburtsjahre) für Humanmedizin, Absolutzahlen.

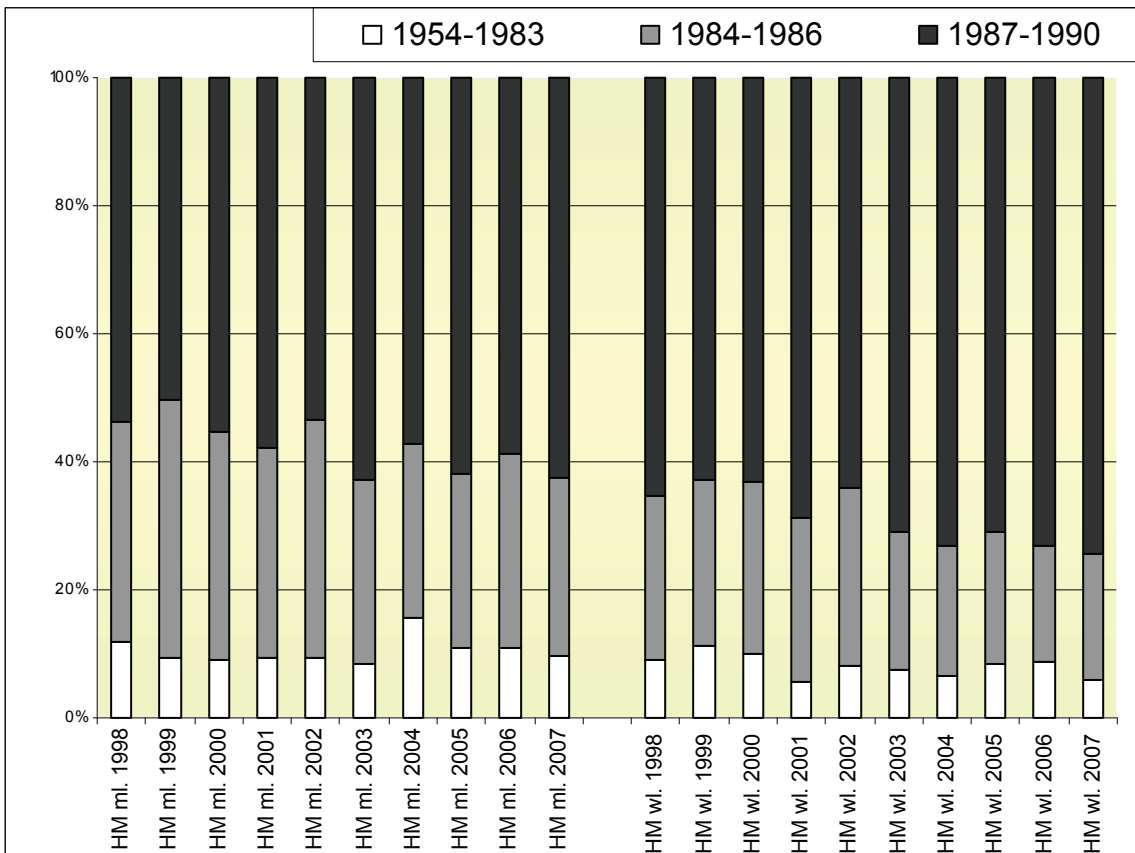


Abbildung 18: Zusammensetzung der Bewerbergruppen nach Geschlecht (mi: männlich, wl: weiblich) und Alter (Geburtsjahre) für Humanmedizin bezogen auf 100% pro Kohorte.

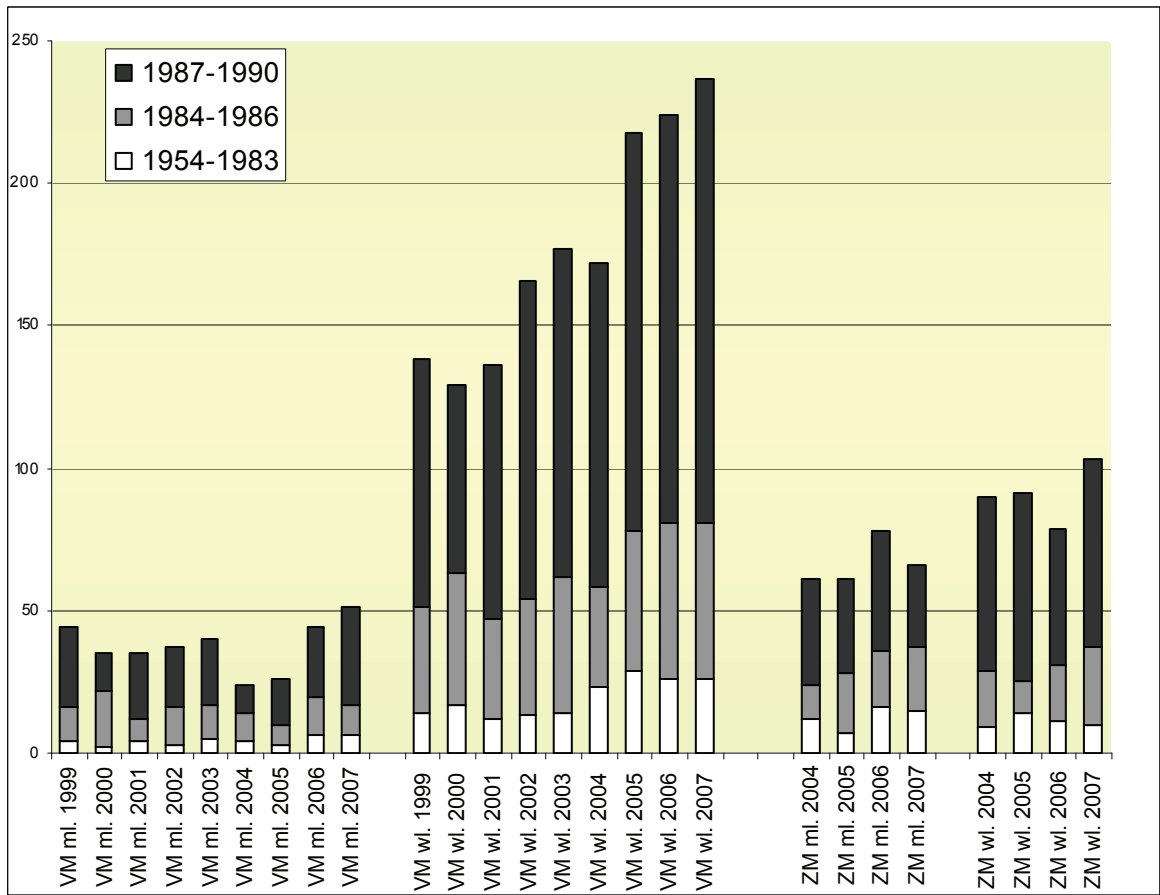


Abbildung 19: Zusammensetzung der Bewerbergruppen nach Geschlecht (ml: männlich, wl: weiblich) und Alter (Geburtsjahre) für Veterinär- und Zahnmedizin, Absolutzahlen.

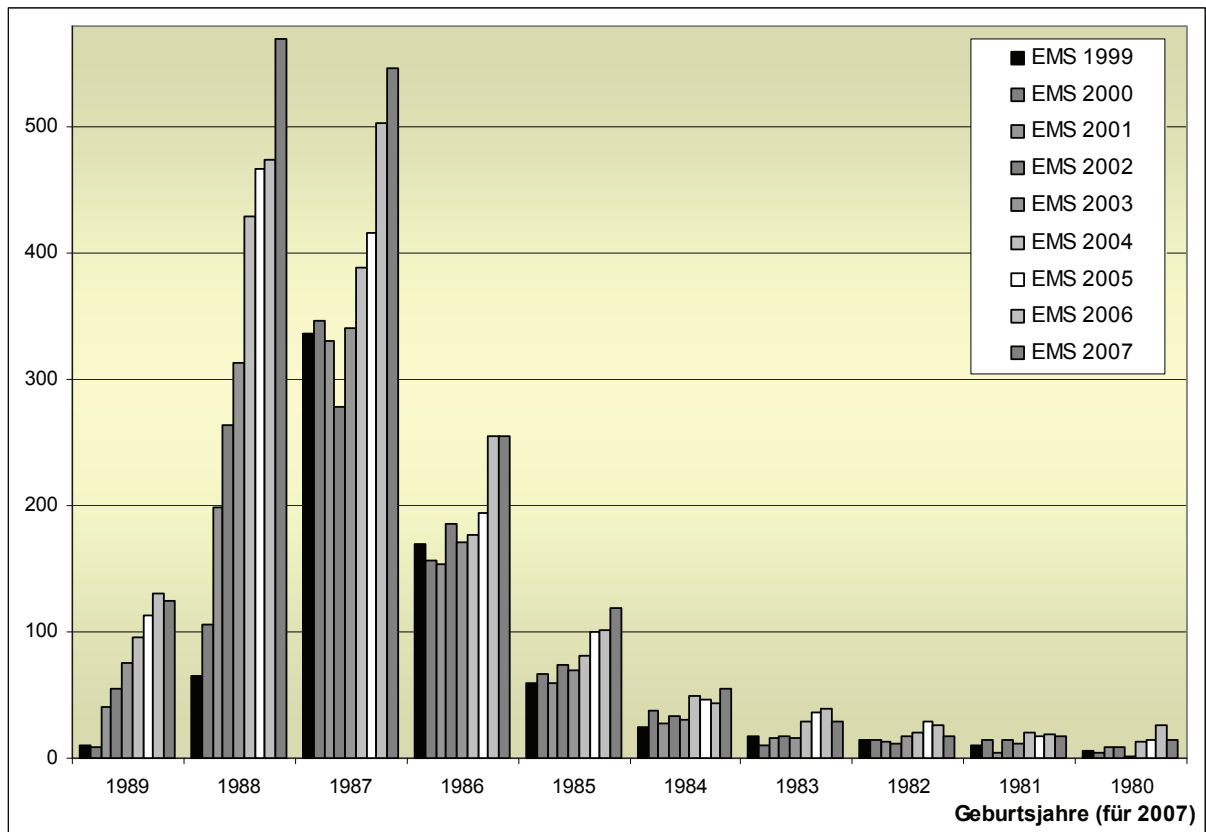


Abbildung 20: Bewerbungen nach Geburtsjahren für 2007 und Entsprechungen für die Jahre 1999 bis 2006 (jeweils um entsprechende Jahre versetzt).

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 Testergebnis zählt.

Die Zahl der Übernahmen der Testergebnisse aus dem Vorjahr bleibt in etwa konstant (Tabelle 14). 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 erweist. 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 6 Personen hat 2007 der übertragene Testwert aus 2006 nicht für eine Zulassung ausgereicht.

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)
	2006	45	94	128	109.9	7.23	41 (4 nicht)
	2007	33	94	126	111.4	7.7	29 (4 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)
	2006	10	90	109	102.1	5.56	8 (2 nicht)
	2007	7	94	112	102.7	6.2	6 (1 nicht)
Zahn- medizin	2004	1			93.0		1
	2005	11	78	99	91.7	7.10	11
	2006	20	89	117	96.7	6.47	20
	2007	32	89	106	98.2	4.11	31 (1 nicht)

Tabelle 14: Testwertstatistiken für Personen mit Übernahme des Testergebnisses aus dem Vorjahr. Zugelassen 1. Runde sind Personen, die sofort einen Platz zugeteilt erhielten (ohne Nachrücker).

5.7 Erneute Testteilnahme

Erste Testteilnahme	Zweite Testteilnahme								
	1999	2000	2001	2002	2003	2004	2005	2006	2007
1998	12	2	1	1	1	1	0	0	9
1999		19	3	0	0	2	1	1	2
2000			12	0	3	1	1	1	1
2001				11	3	1	2	1	0
2002					32	13	1	3	1
2003						78	21	10	3
2004							151	26	15
2005								184	34
2006									257
Gesamt	12	21	16	12	39	96	177	226	313

Tabelle 15: Wiederholte Testteilnahme: Anzahl der Personen.

Es ist möglich, erneut zum Test anzutreten und dann jeweils 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 deshalb auch Personen erneut an, um ihr Ergebnis zu verbessern. Wegen der Verschärfung des NC nimmt die Zahl der erneuten Antritte auch dieses Jahr zu (Tabelle 15) – das Risiko, dass ein knapp über der Zulassungsgrenze liegender Wert in einem Jahr dann im nächsten Jahr nicht mehr reicht, wird grösser. Bisher ist die Wahrscheinlichkeit eines besseren 2. Testergebnisses deutlich höher. Dieser Unterschied zwischen erster und zweiter Messung scheint sich beim Testwert im Mittel bei 8 zu stabilisieren (Tabelle 16).

Zeitraum	N	1. Testung		2. Testung		Differenz	
		m	s	m	s	m	s
1998-2003	100	86.9	7.8	96.6	9.3	9.7	6.5
2004	96	89.0	7.0	97.0	9.9	8.0	8.2
2005	177	93.7	6.6	101.2	8.4	7.6	9.3
2006	226	92.4	6.9	100.7	8.8	8.3	7.9
2007	313	93.2	7.0	101.9	9.3	8.7	7.9
Gesamt	912	92.0	7.3	100.4	9.2	8.4	8.1

Tabelle 16: Testwertvergleich zwischen erster und zweiter Testteilnahme 1998 bis 2007.

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.

		Jahr					Gesamt
		1998-2003	2004	2005	2006	2007	
Testwert-differenz	verschlechtert	4	15	30	25	36	110
	gleich geblieben	0	3	8	6	8	25
	verbessert	96	78	139	195	269	777
Gesamt		100	96	177	226	313	912

Tabelle 17: Zahl der Verbesserungen/Verschlechterungen des Testwerts bei Testwiederholungen.

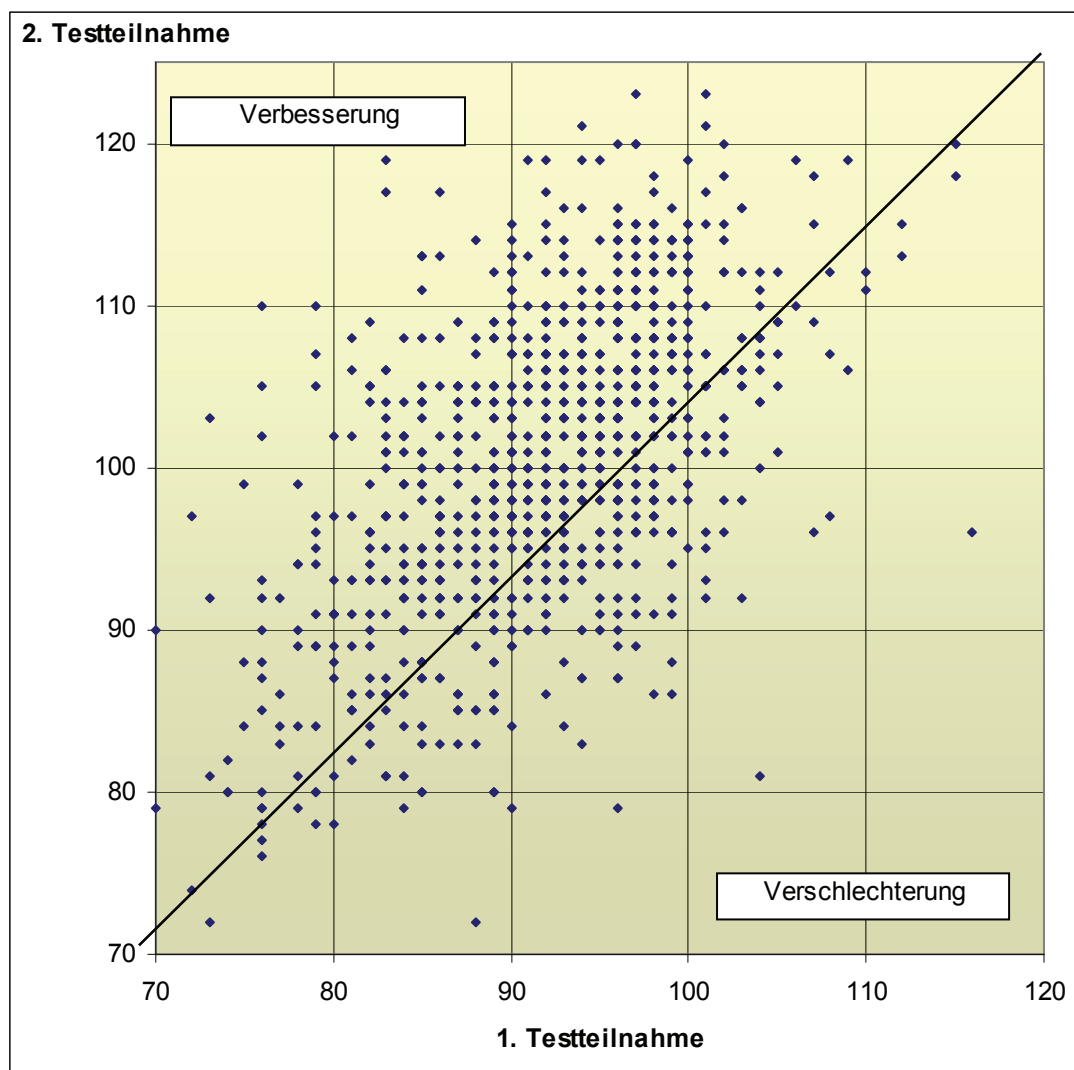


Abbildung 21: Testwerte für Personen, die zweimal am Test teilnahmen (1998 bis 2007).

5.8 Sprachgruppen

Nach wie vor treten vor allem deutschsprachige Kandidaten zum EMS an; für die französisch- und italienischsprachigen Kandidaten 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.

Der Chancengleichheit der Sprachgruppen wird wie bisher grosse Aufmerksamkeit gewidmet:

- 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 eine Bearbeitung der veröffentlichten Originalversionen in der jeweiligen Sprache. Das ZTD unterstützt dies durch eine Kurz-Anleitung für die Testleiter dieser Probeläufe, die den Schulen zur Verfügung gestellt wird.

			Geburtsjahre			Total
			1961-1982	1983-1985	1986-1989	
Humanmedizin	deutsch	männlich	50	135	309	494
		weiblich	49	155	564	768
		Total	99	290	873	1262
	französisch	männlich	3	8	18	29
		weiblich	0	5	23	28
		Total	3	13	41	57
	italienisch	männlich	1	10	18	29
		weiblich	1	1	24	26
		Total	2	11	42	55
Veterinärmedizin	deutsch	männlich	6	9	26	41
		weiblich	25	48	108	181
		Total	31	57	134	222
	französisch	männlich	0	2	6	8
		weiblich	1	7	38	46
		Total	1	9	44	54
	italienisch	männlich	0	0	2	2
		weiblich	0	0	10	10
		Total	0	0	12	12
Zahnmedizin	deutsch	männlich	15	22	24	61
		weiblich	9	25	55	89
		Total	24	47	79	150
	französisch	männlich	0	0	1	1
		weiblich	1	1	4	6
		Total	1	1	5	7
	italienisch	männlich	0	0	4	4
		weiblich	0	1	7	8
		Total	0	1	11	12

Tabelle 18: Sprachgruppen der Teilnehmenden 2007 nach Disziplin, Alter und Geschlecht aufgeschlüsselt.

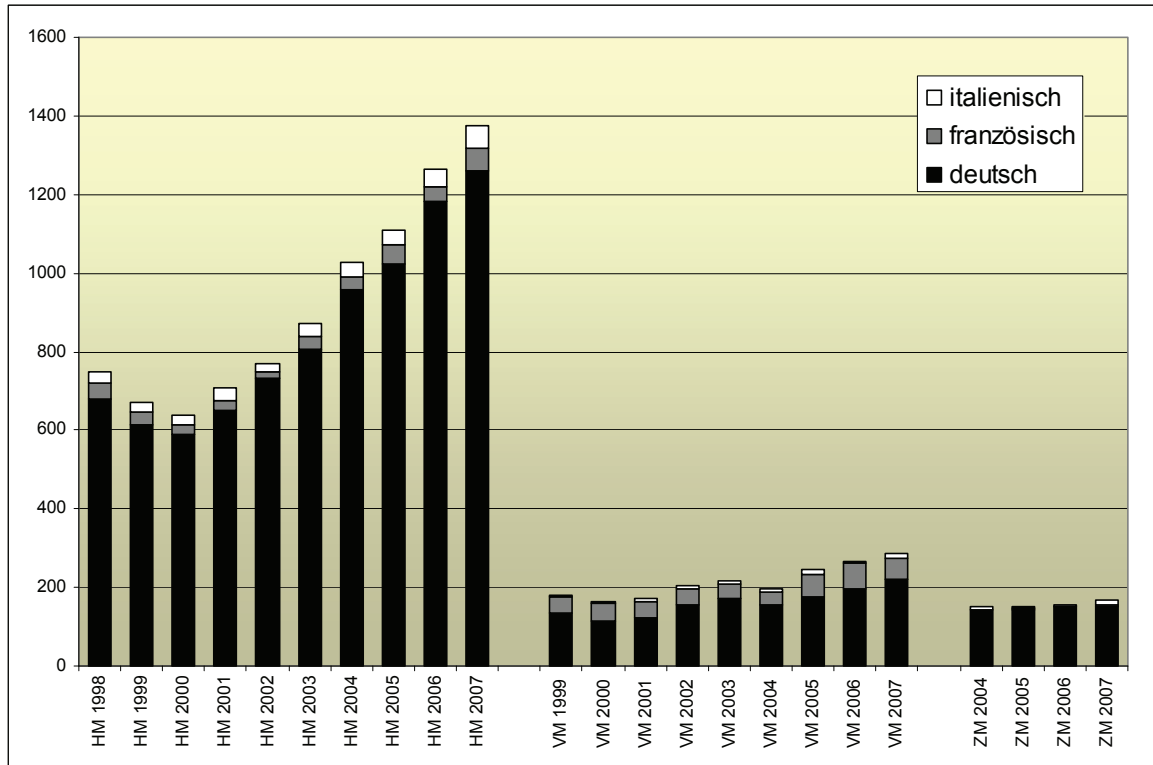


Abbildung 22: Vergleich der Sprachgruppen 1998 bis 2007, Absolutzahlen.

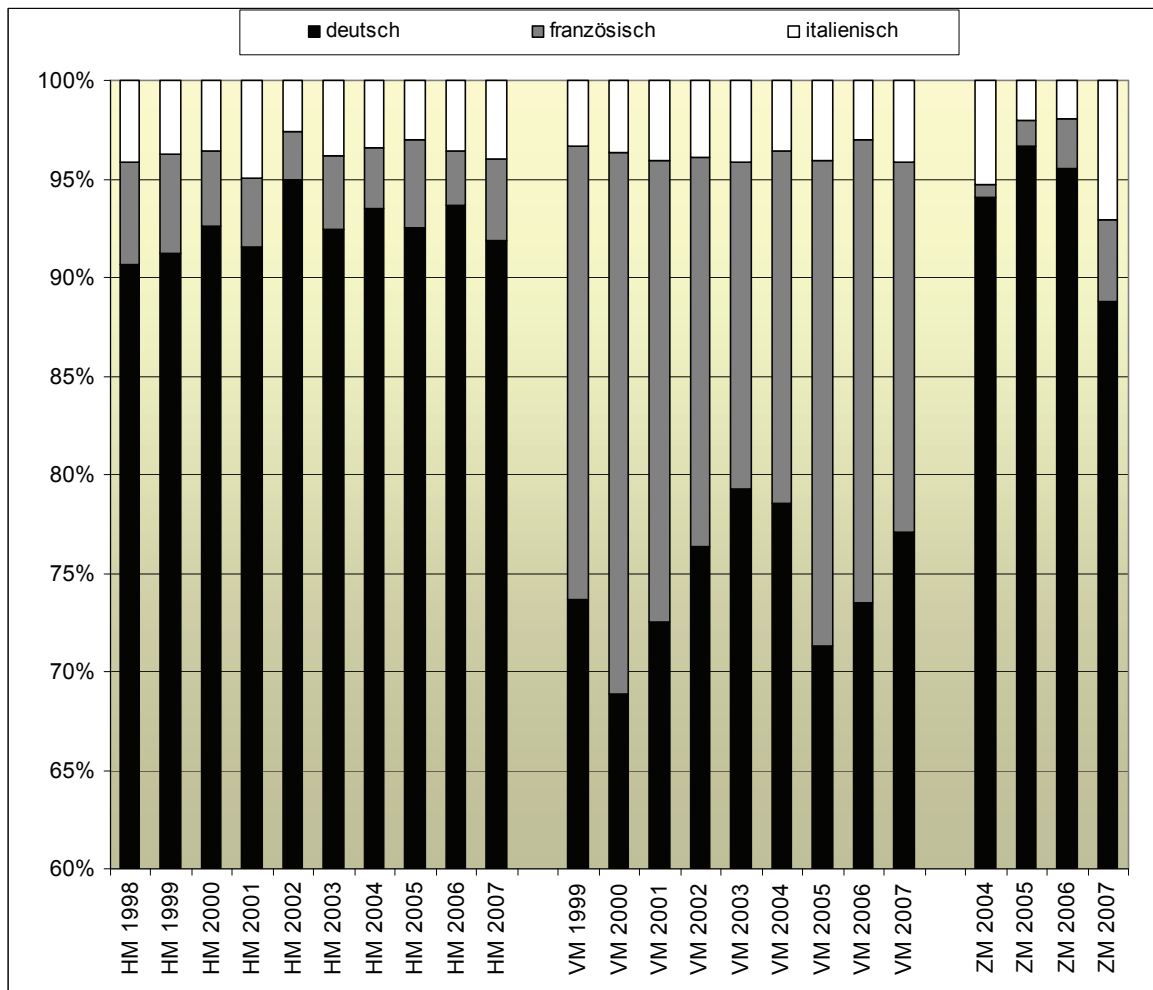


Abbildung 23: Vergleich der Sprachgruppen 1998 bis 2007 Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%).

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.

Geb.- Jahr	Maturitätsjahr																						
	83	86	87	88	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05		06	07
54																1							1
63																						1	1
64	1																						1
65		1																					1
66		1	1											1									3
67		1	1																			1	3
68				2																			2
69			1												1							2	4
70					1																	1	2
71					1																1		2
72						1	1									1							3
73															2			1				1	4
74								2	1				1									1	5
75												1											1
76										2	1						1			1	2	1	8
77											2	4	1				1			1	2	1	12
78												1	1	1			1				2	5	11
79													1	5				2	2		5	3	18
80														1	4	1			1	1	1	6	15
81															1		1	1	1	2	4	7	17
82																4	4	1	1			8	18
83																1	4	7	2	1	6	8	29
84																	3	13	13	10	8	8	55
85																		7	16	36	35	25	119
86																			12	61	126	56	255
87											1					1		5	2	19	317	201	546
88																	2	8	1	3	89	467	570
89									1							1		1	2	2	1	110	118
90																				0	0	7	7
	1	3	3	2	2	1	1	2	2	3	3	5	5	8	8	10	14	49	53	137	599	920	1831

Tabelle 19: Geburtsjahr und Maturitätsjahr. Dicke Linien trennen die Altersgruppen (bei der ältesten Gruppe zusätzlich eine Teilung nach dem Maturitätszeitpunkt).

6 Beschreibung des verwendeten Eignungstests

6.1 Aufbau des Tests

Der EMS besteht seit 2005 aus zehn Untertests, die zu einem Gesamtwert verrechnet werden, „Planen und Organisieren“ ist als neuer Untertest dazugekommen. Im Jahr 2006 wurde die Reihenfolge optimiert, um den Vormittags- und den Nachmittagsteil in der Dauer anzugleichen.

Die Tabelle 20 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Untertests. Aufgabenbeispiele finden Sie im Anhang.

Bezeichnung der Untertests	Aufgaben	Max. Punktzahl	Bearbeitungszeit
Quantitative und formale Probleme	20	20	50 min
Schlauchfiguren	20	20	12 min
Textverständnis	18	18	45 min
Planen und Organisieren	20	20	60 min
Konzentriertes und sorgf. Arbeiten	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
Medizinisch-naturw. Grundverst.	20	20	50 min
<i>Gedächtnistests:</i>			
Figuren reproduzieren	20	20	5 min
Fakten reproduzieren	20	20	7 min
Muster zuordnen	20	20	18 min
Diagramme und Tabellen	20	20	50 min
Gesamttest	198	198	ca. 5 Std.
Gesamtdauer (inkl. Pause) 9.00 bis ca. 17.00 Uhr			

Tabelle 20: Struktur und Ablauf des EMS 2007.

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 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 seitdem 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.

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. Dieser erlaubt die gesetzlich geforderte Studienzulassung nach Eignung;
- 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 Schwierigkeitsspektrum 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 erfolgte in Deutschland im Rahmen so genannter „**Einstreuaufgaben**“. Nur bei ausreichender Bewährung wurden solche Aufgaben in nachfolgenden Testversionen für die Werteberechnung verwendet. Im Unterschied zu vielen „Übungsaufgaben“, die im so genannten 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 werden von der Menge der insgesamt angestrichenen Zeichen abgezogen. Die verbleibende Menge sind die „Richtigen“, die 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 24). 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.

In diesem Jahr können aufgrund guter Kennwerte alle Aufgaben gewertet werden. Die Berechnung und Verwendung von Gütekriterien für die Itemselektion ist ein wichtiges Merkmal eines Tests. Insofern bevorzugen wir intern die Strategie, im Zweifel eine Aufgabe lieber auszuschließen und die Testgüte damit ausreichend hoch zu halten. Das Standardisierungsverfahren zur Bildung des Testwertes gewährleistet auch bei unterschiedlichen Aufgabenzahlen vergleichbare Testwerte zwischen den Jahren.

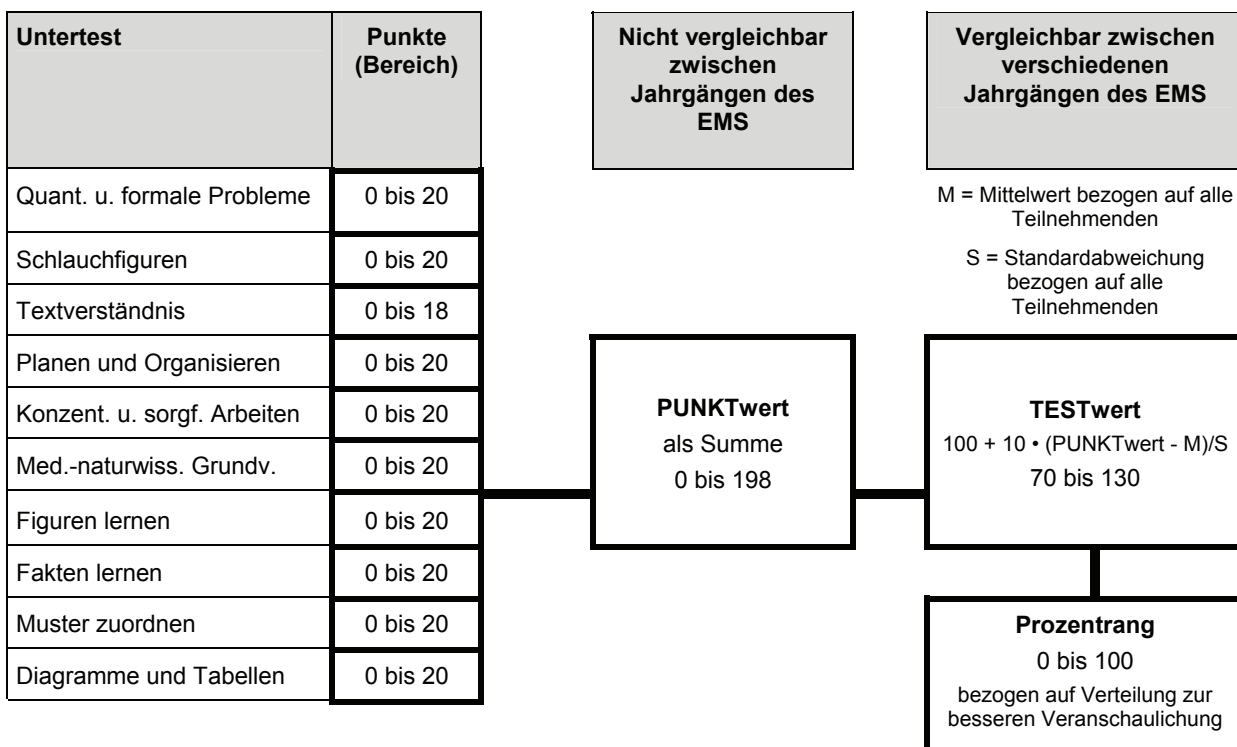


Abbildung 24: Punktwerte der einzelnen Untertests 2007 und ihre Zusammenführung über den Punktwert zum Testwert und Prozentrang.

6.3 Mittlerer Rangplatz der Untertests

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.

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 am Ende der Zuteilung noch zur Verfügung standen.

Aus diesem Grunde hat der Beirat EMS beschlossen, ab 2005 weiterhin 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 wird ebenfalls 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 belohnen gute Leistungen in einem schwierigen Untertest stärker, indem die dort vergebenen niedrigen Ränge höheres Gewicht erhalten:

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 50.5. Haben im anderen Untertest nur 10 Personen diesen Wert erreicht, ist der Rangplatz 5.5. Letzterer wird höher bewertet, weil die 20 Punkte im zweiten Untertest schwerer zu erzielen waren und deshalb wertvoller sind.

Zwecks Vergleichbarkeit der Rangplätze zwischen den Jahren wird ein mittlerer Rangplatz auf der Basis von 1000 Personen verwendet.

- Der Rangplatz bewegt sich 2007 zwischen 41 und 962. Niedrige Werte stehen für bessere Leistungen.
- Er korreliert mit dem Testwert mit $-.98$ was praktisch für eine Äquivalenz spricht – siehe auch Abbildung 25.
- Er differenziert allerdings besser zwischen den Personen und lässt eine kapazitätsgenaue Auswahl zu.

Das seit 2005 angewandte Prinzip, bei gleichem Testwert den Rangplatz zu verwenden, kann daher auch für die Folgejahre beibehalten werden.

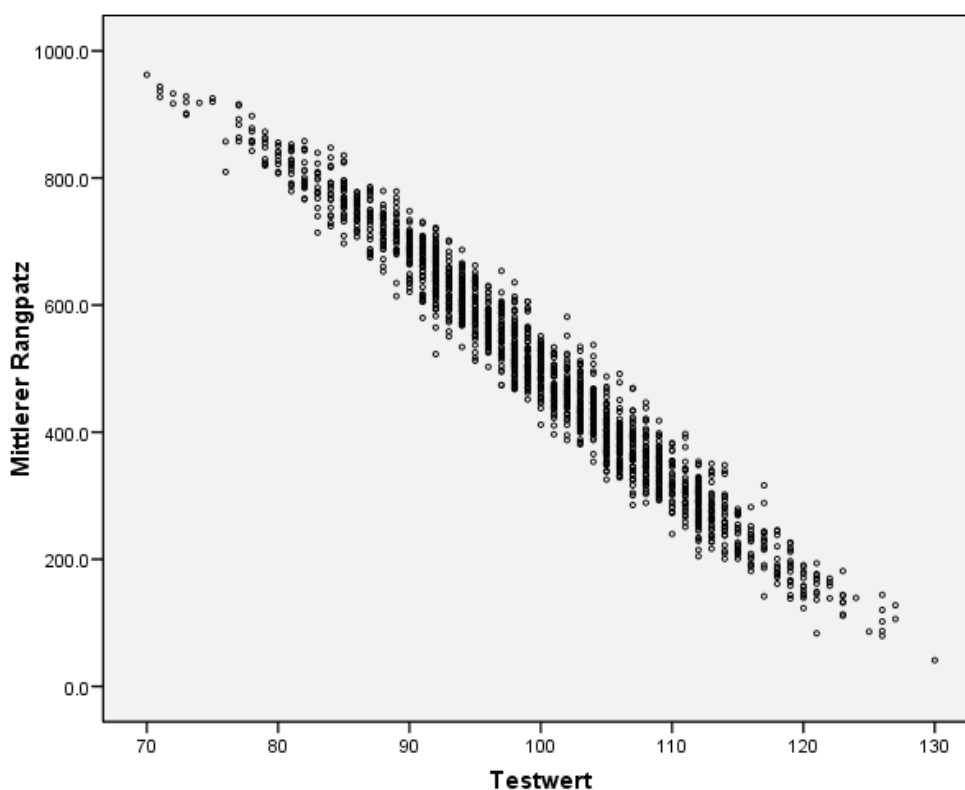


Abbildung 25: Beziehung zwischen Testwert und mittlerem Rangplatz für 2007.

7 Testanwendung in der Schweiz 2007

7.1 Verteilungsprüfung

Die Verteilung des Punktwerts entspricht 2007 wiederum einer Normalverteilung. Metrische Verfahren und die Standardisierung zum Testwert durch lineare z-Transformation sind daher zulässig. Die Punktwerte der Einzeltests sind nicht normalverteilt (Tabelle 21).

	Punktwert	Quantitative und formale Probleme	Schlauchfiguren	Textverständnis	Planen und Organisieren	Med.-naturwiss. Grundv.	Figuren lernen	Fakten lernen	Muster zuordnen	Diagramme und Tabellen	Konzentr. und sorgf. Arbeiten
Mittelwert	108.23	10.61	13.83	9.33	10.21	10.42	11.28	11.15	11.10	10.15	10.03
Stand.-abweich.	21.47	3.61	3.59	3.36	3.52	3.35	3.36	3.85	2.96	3.51	4.72
Median	109	10	14	9	10	10	11	11	11	10	10
Modalwert	114	9	14	10	10	11	12	11	11	10	10
Spannweite	135	19	18	18	19	19	20	19	20	19	20
25. Perzentil	93	8	11	7	8	8	9	8	9	8	7
75. Perzentil	123	13	17	12	13	13	14	14	13	13	13
K-S: extremste Differenz	0.03	0.08	0.08	0.08	0.06	0.07	0.07	0.07	0.09	0.07	0.07
K-S: Z-Wert	1.197	3.602	3.428	3.429	2.769	2.902	2.993	3.054	3.655	2.976	3.195
K-S: Asymp. Sig.	0.114	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Tabelle 21: Kennwerte der Punktwertskalen und Verteilungsprüfung auf Normalverteilung (K-S: Kolmogorov-Smirnov-Test).

Die in den folgenden Diagrammen (Abbildung 26 und Abbildung 27) dargestellten Verteilungen der Punktwerte für die Untertests zeigen, dass alle Skalen gut bis sehr gut differenzieren. Der Untertest „Schlauchfiguren“ war der leichteste Untertest (Modalwert 14), die übrigen Untertests bewegen sich um den Modalwert 10.

Im Untertest „Konzentriertes und sorgfältiges Arbeiten“ sind die relativen Häufungen an beiden Enden der Verteilung auf die Systematik der Umrechnung der Rohwerte in den Normwert zurückzuführen (vgl. die Darstellung dieses Tests ab Seite 91).

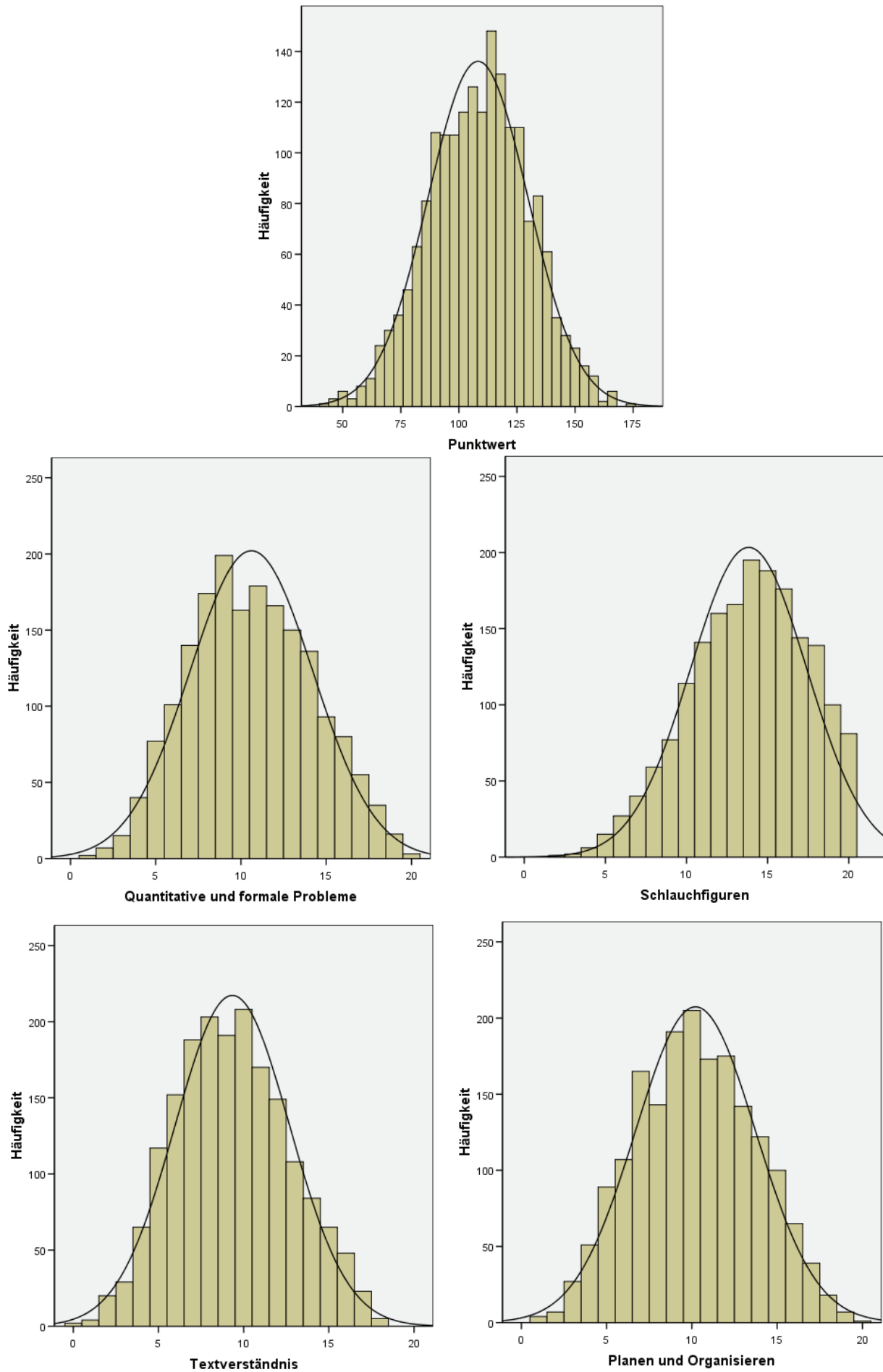


Abbildung 26: Häufigkeitsverteilungen für Punktwert und Punkte der Untertests (1).

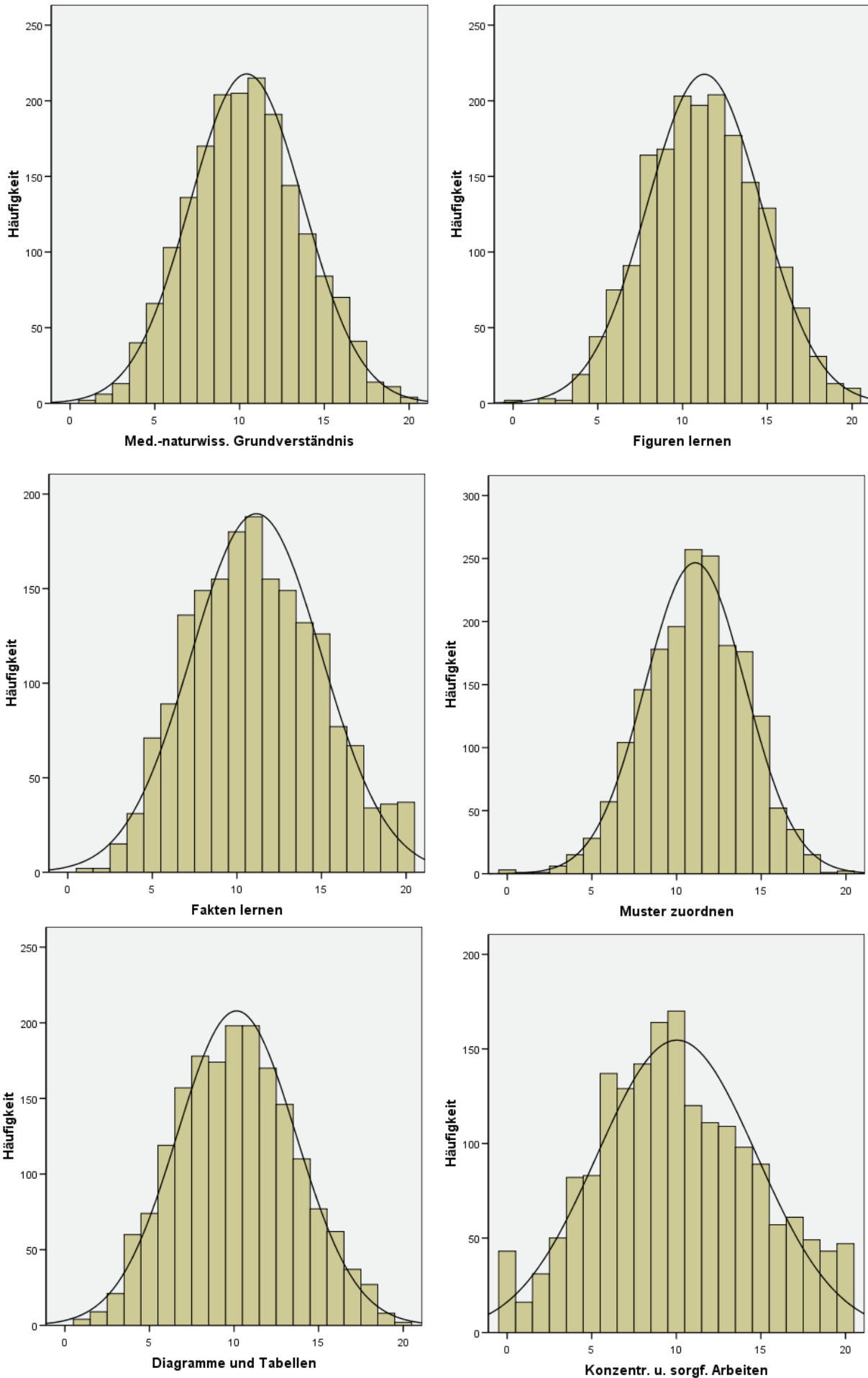


Abbildung 27: Häufigkeitsverteilungen für Punkte der Untertests (2).

7.2 Vergleich der Testfassungen 1998 bis 2007

Durch die Transformation der Punktwerte in Testwerte können diese **Testwerte** aus verschiedenen Jahren direkt 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 22 zeigt den Vergleich für alle bisherigen Testdurchführungen seit 1998 in der Schweiz mit den geschätzten Punktwerten aus Deutschland (Schätzung aufgrund der Item-Schwierigkeiten von Items, die ggf. in einer anderen Reihenfolge und Zusammenstellung angewendet worden sind). Für die Berechnung der Punktwerte in der Schweiz wurden zum Zweck besserer Vergleichbarkeit nur die Ergebnisse der deutschsprachigen Gruppe berücksichtigt:

		2007	2006	2005	2004	2003	2002	2001	2000	1999	1998
Muster zuordnen	CH	11.2	11.3	12.7	10.5	10.8	11.3	10.1	10.3	11.6	13.2
	D	9.3	9.7	10.4	10.3	10.2	10.9	9.7	9.3	11.0	12.2
Med.-naturwiss. Grundverst.	CH	10.7	10.7	11.3	10.4	10.5	9.9	10.1	10.7	11.3	12.5
	D	10.8	10.5	11.0	10	10.2	9.5	10.7	10.8	11.2	11.6
Schlauchfiguren	CH	14.1	13.8	12.3	12.1	12.7	12.1	13.3	13.2	12.9	13.3
	D	12.3	12.2	10.9	11.8	12	11.8	12.2	12.3	11.9	11.8
Quant. und form. Probleme	CH	10.9	10.8	10.6	9.9	10.5	11.1	10.9	11.8	11.6	11.9
	D	10.3	9.9	9.8	9.4	9.6	9.9	9.8	10.2	9.7	9.9
Textverständnis	CH	9.6	9.1	9.2	8.9	10.2	9.2	8.5	9.0	8.9	10.3
	D	9.6	8.7	8.9	9.0	8.9	8.6	8.6	9.3	8.9	10.2
Figuren lernen	CH	11.4	11.7	13.4	13.0	13.6	15.7	10.6	10.2	12.7	12.6
	D	9.3	9.3	11.5	11.5	10.7	11.8	9.2	9.3	11.8	11.6
Fakten lernen	CH	11.4	12.0	12.6	12.4	11.4	10.7	10.8	10.1	11.8	11.6
	D	9.0	9.3	11.1	11.3	10.3	9.7	9.3	8.9	11.0	11.2
Diagramme und Tabellen	CH	10.5	9.4	10.5	10.7	11.7	10.8	10.3	10.4	10.5	11.4
	D	9.9	9.6	10.7	10.1	10.7	10.5	10.0	9.7	10.3	11
Total	CH	89.8	88.8	92.7	87.9	91.4	90.9	84.6	85.7	91.3	96.8
	D	80.5	79.1	83.7	83.4	82.6	82.7	79.5	79.8	85.8	89.5

Tabelle 22: Mittelwerte (Punkte) der Untertests 1998 bis 2006; deutsche Sprachgruppe Schweiz (Humanmedizin) und Schätzung aufgrund der Item-Schwierigkeitswerte aus Deutschland – für 8 Untertests liegen Werte vor.

Die folgenden Abbildungen veranschaulichen den Vergleich zwischen deutschen und deutschschweizerischen Ergebnissen für die Untertests. Die Differenzen bewegen sich – von wenigen „Ausreissern“ abgesehen – auf weitgehend konstantem Niveau. Die Kurven verlaufen gleichförmig, die relativen Schwierigkeiten stimmen also nach wie vor gut überein.

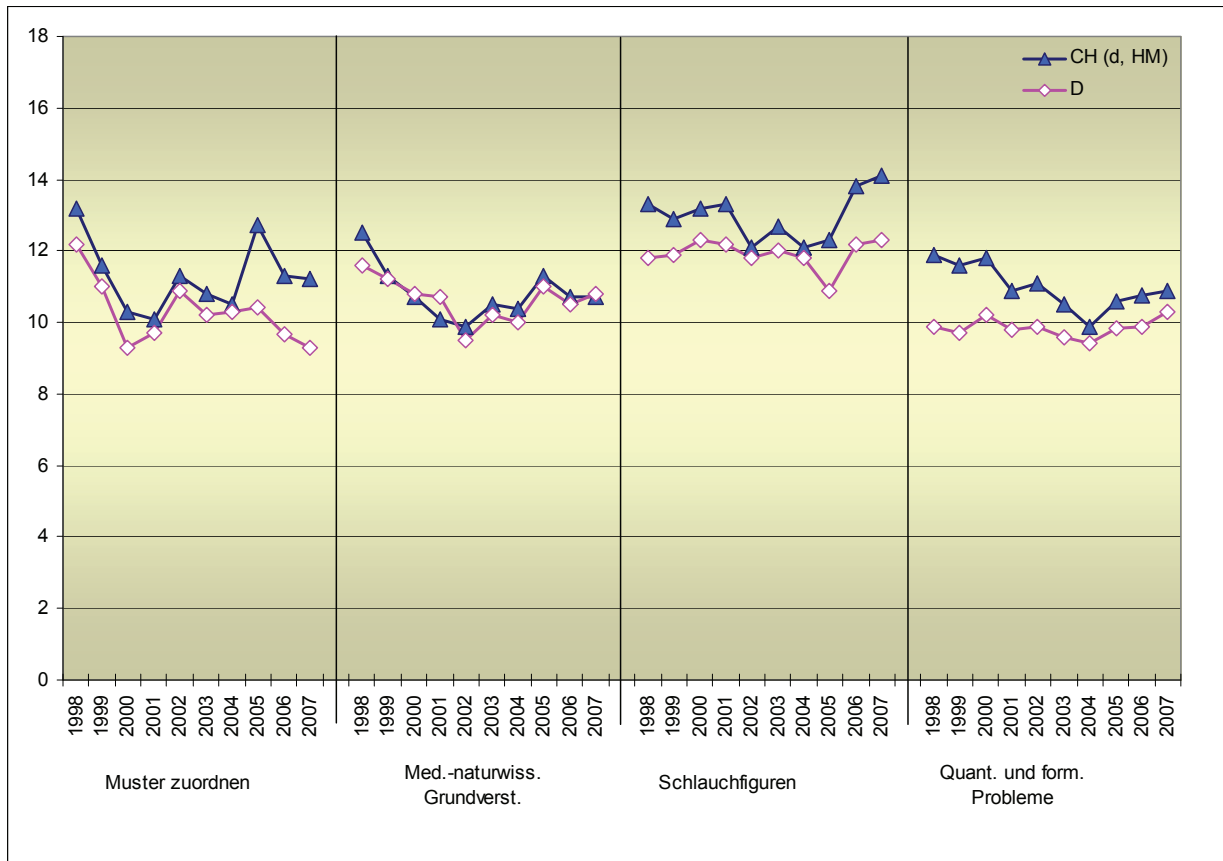


Abbildung 28: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig, Humanmedizin) und Vergleichswerte Deutschland, 1998 bis 2007.

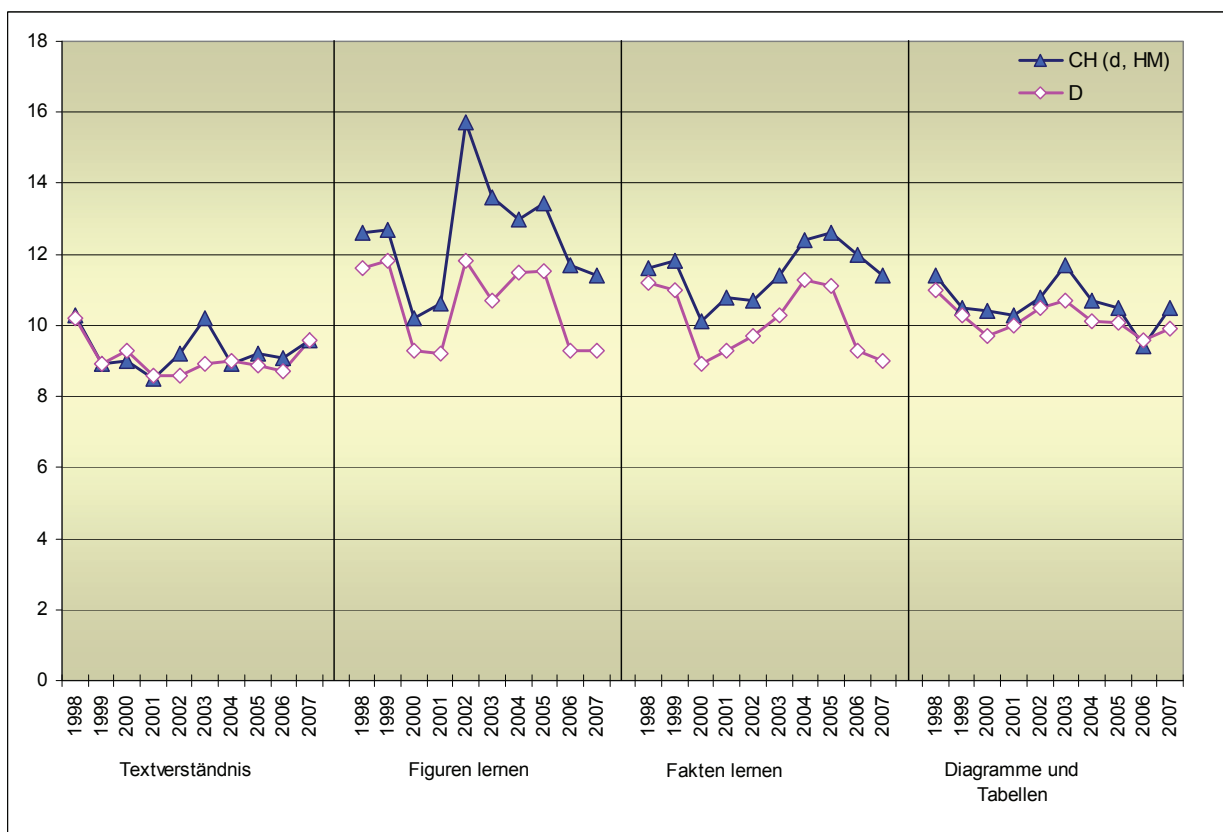


Abbildung 29: Mittelwertvergleiche der Untertests Schweiz (deutschsprachig, Humanmedizin) und Vergleichswerte Deutschland 1998 bis 2007.

7.3 Äquivalenz der Sprachversionen

Der Test wurde nach den gleichen Kriterien wie in den Vorjahren von der deutschen Vorlage adaptiert. Eine ausführliche Beschreibung des Verfahrens kann dem Bericht 5 (1999) entnommen werden. Zielsetzung bleibt, möglichst übereinstimmende Aufgabenschwierigkeiten zwischen den Sprachversionen zu schaffen.

Chancengleichheit als Gleichbehandlung der Sprachgruppen muss gewährleistet sein, die Lösungswahrscheinlichkeit einer Aufgabe muss *bei gleicher Fähigkeit* tatsächlich auch vergleichbar sein. Hervorzuheben ist, dass seit 2005 in französischer und italienischer Sprache die Schlussredaktion des Tests von zweisprachigen Lehrpersonen durchgeführt wird, die an Maturitätsschulen unterrichten. Sie wurden auch dahingehend geschult, dass es vor allem auf eine Übertragung der Schwierigkeit in die jeweilige Sprache ankommt. Da sie zudem mit dem Sprachniveau der Maturitätsstufe durch ihre Tätigkeit gut vertraut sind, gewinnt der Test zusätzlich an Güte.

Die Gleichheit ist nicht allein durch Übersetzung zu beeinflussen, sondern hängt von mehreren Faktoren ab. Bei allen Adaptationen muss mit Unterschieden hinsichtlich der Aufgabenschwierigkeiten zwischen den Sprachversionen gerechnet werden. Abbildung 30 fasst zusammen, welche Ursachen für diese Differenzen verantwortlich sein können.

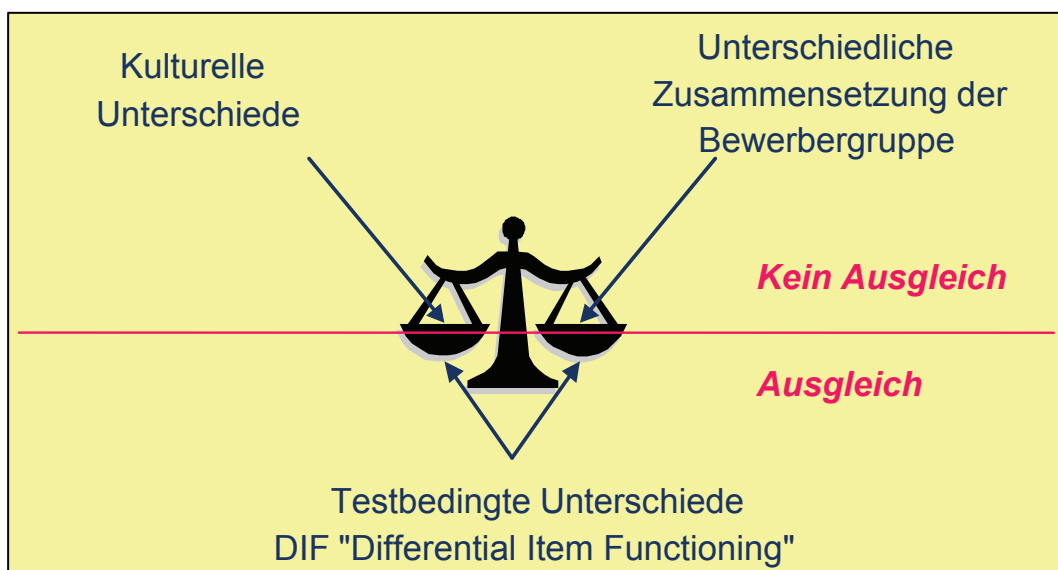


Abbildung 30: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen.

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

Testbedingte Unterschiede werden durch die Adaptation verursacht und sollen möglichst vollständig ausgeglichen werden. Beispiele wären eine Vereinfachung oder Erschwerung der Aufgabe bei der Übersetzung (Wortwahl, Satzgliederung o.ä.). Erneut wird zu diesem Zweck das bewährte DIF-Verfahren (Differential Item Functioning) angewendet.

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 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. Französisch- und italienischsprachige Personen können den NC durch ein Studium in ihrer Muttersprache an einer anderen Universität vermeiden. 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 sechs von zehn Untertests. Der Sprachausgleich beschränkt sich demnach auf die stärker sprachabhängigen Untertests „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“, „Diagramme und Tabellen“ sowie „Planen und Organisieren“.
2. Den Ausgleich nur vorzunehmen, wenn signifikante Mittelwertunterschiede im entsprechenden Untertest vorhanden sind. Geprüft werden dazu die Abweichungen der jeweiligen Zielsprache von der deutschen Sprachgruppe.

Adaptationsbedingte Effekte sind nur in den sechs **sprachabhängigen Untertests** zu erwarten, während die vier **sprachunabhängigen Untertests** davon kaum betroffen sein können.

Spricht man von DIF, so muss dieses 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 Verschiebung der Schwierigkeiten aller Items eines Untertests. Es ist unwahrscheinlich, dass diese systematischen Unterschiede testbedingt sind. Testbedingte Unterschiede als sprachliche Besonderheiten zeigen sich in spezifischen Abweichungen einzelner Items. Systematische Unterschiede sind dagegen mit grösserer Wahrscheinlichkeit den Faktoren „Kulturunterschiede“ und „unterschiedliche Repräsentativität“ zuzurechnen.

Es ergibt sich für das Ausgleichsverfahren die nachfolgende Zielstellung:

Das Risiko, dass testbedingte Unterschiede die Sprachunterschiede systematisch beeinflussen, soll verringert werden. Andererseits soll kein Ausgleich von

Unterschieden aufgrund unterschiedlicher Zusammensetzung der Stichproben und der kulturellen Unterschiede erfolgen.

7.3.1 Sprachvergleich für die Untertests

Wie jedes Jahr treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und auch in den wenig sprachabhängigen Untertests auf. Erneut liegen die Punktzahlen der deutschsprachigen Teilnehmer mehrheitlich über jenen der übrigen Kandidaten. Im Untertest „Textverständnis“ erzielen allerdings die italienischsprachigen Teilnehmer erneut die besten Ergebnisse. Dies war bereits 2006 zu beobachten.

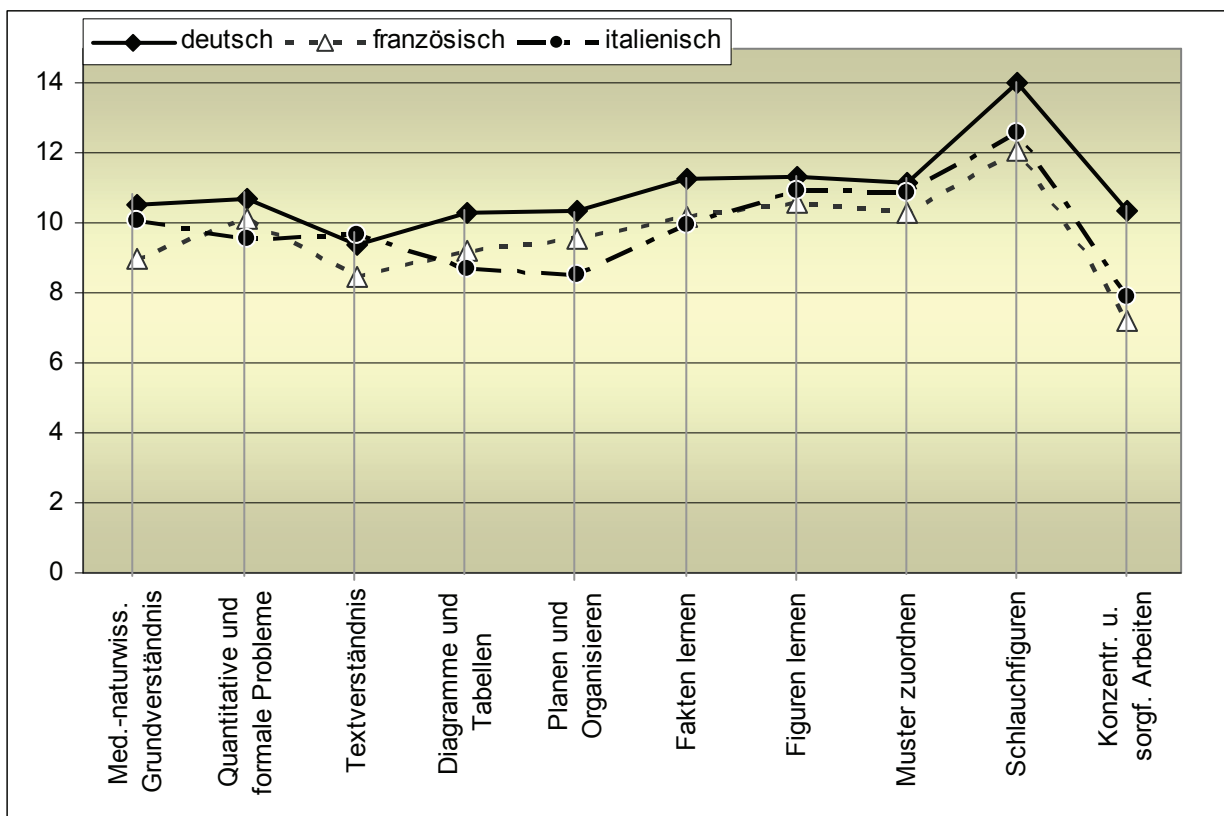


Abbildung 31: Mittelwerte der Untertests für die Sprachgruppen 2007 (unkorrigiert).

Die Abbildung 32 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern für alle Testanwendungen seit 1998. In der Abbildung 33 ist dieser Vergleich zwischen deutsch- und italienischsprachigen Personen dargestellt. Die senkrechten Linien bezeichnen die Spannweite (Minimum bis Maximum) der Differenzen über alle Jahrgänge seit 1998, die Verbindungslinie steht für die Differenzen aus dem aktuellen Jahrgang. Höhere Differenzwerte sprechen für bessere Ergebnisse der deutschen Sprachgruppe.

Grundsätzlich treten in diesem Jahr kaum extreme Differenzen auf, sie ordnen sich in die bisher festgestellten Spannweiten der Differenzen ein. Einzige Ausnahme ist „Konzentriertes und sorgfältiges Arbeiten“ für die französische Sprachgruppe.

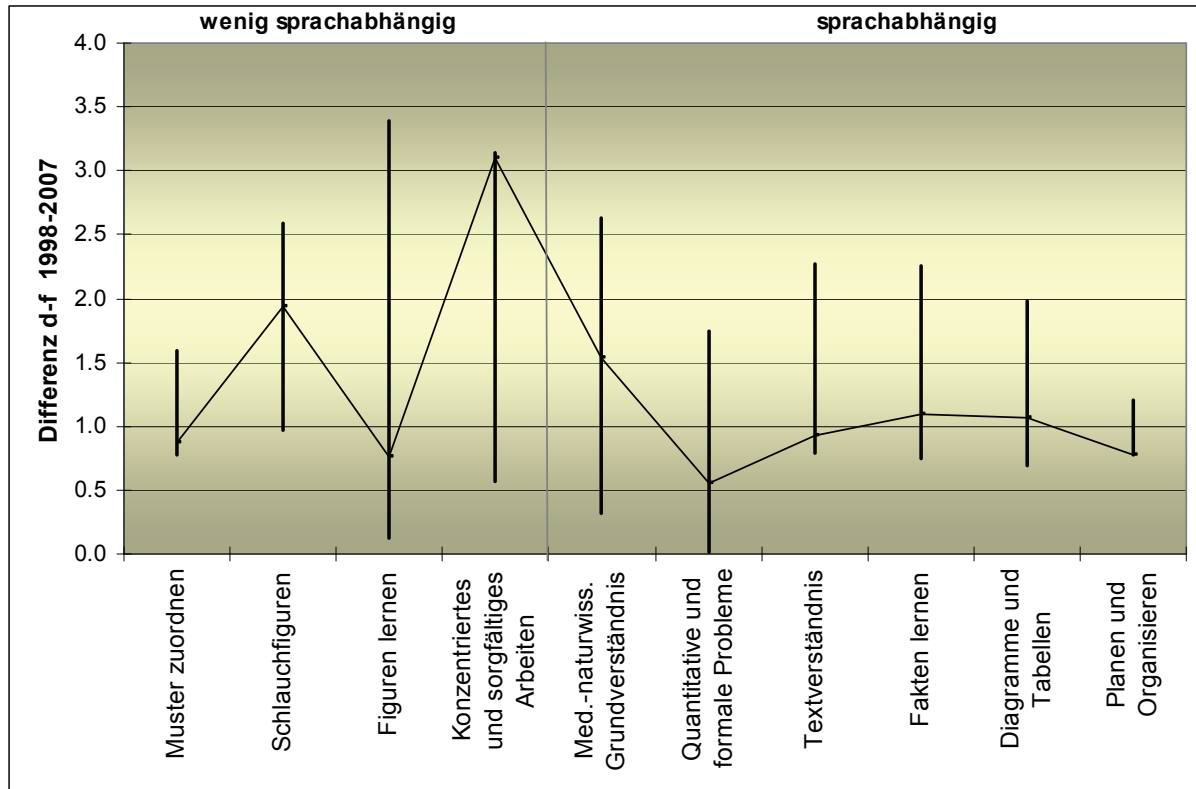


Abbildung 32: Differenz Testwert **deutsch-französisch** 2007 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2007 (vertikale Linien).

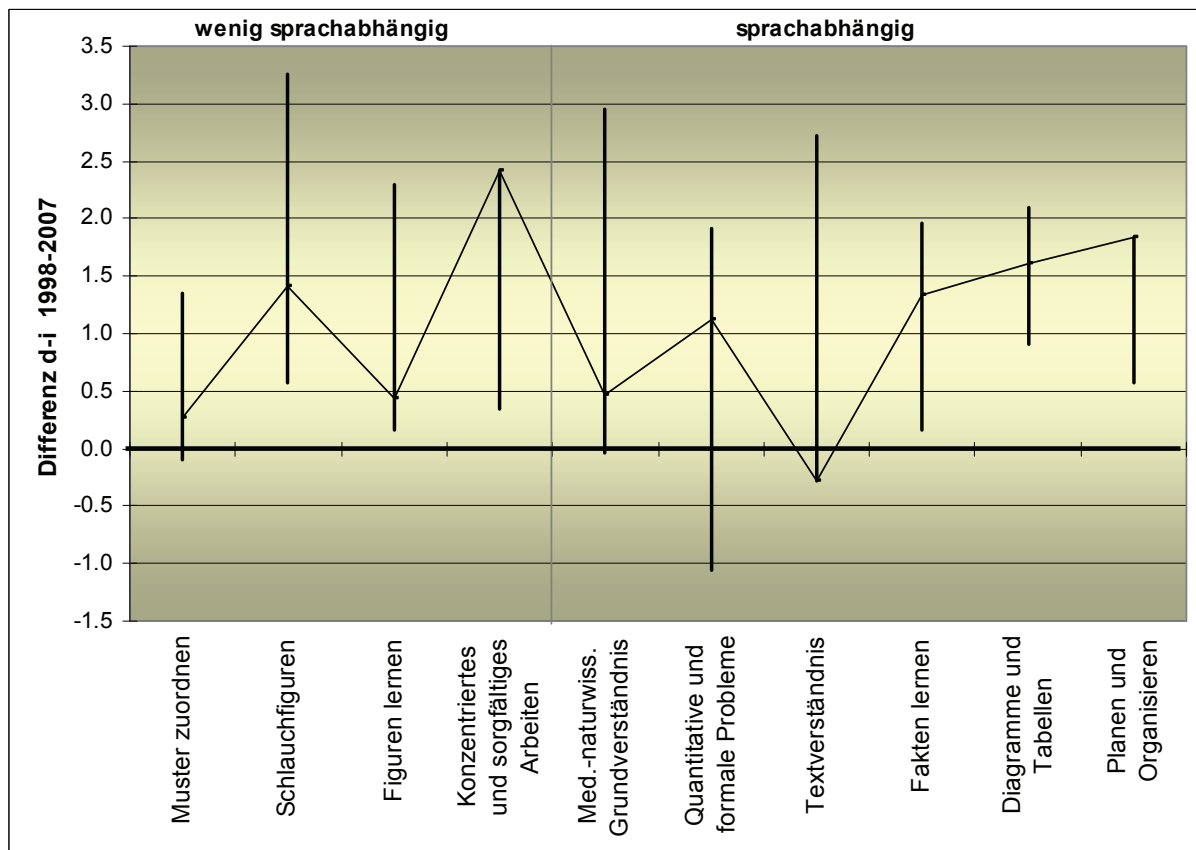


Abbildung 33: Differenz Testwert **deutsch-italienisch** 2007 (unkorrigiert) und Spannweite dieser Differenz (min-max) der Jahre 1998 bis 2007 (vertikale Linien).

7.3.2 Darstellung des Korrekturverfahrens

7.3.2.1 Identifikation und 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 zwischen den Sprachgruppen signifikant unterscheiden** (Vergleich deutsch - französisch, beziehungsweise deutsch - italienisch mittels t-Test). Wenn keine Mittelwert- und Varianzunterschiede vorhanden sind, trägt der Untertest auch nicht zu Mittelwertunterschieden beim Testwert bei. Eine Korrektur allfälliger DIF-Items (die vorhanden sein können) würde dann eventuell 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 korrekt 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 falscher Antwort 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. Nur Personen, welche das betreffende Item nicht gelöst haben, erhalten den Bonus.

7.3.2.2 Welche Untertests ausgleichen?

Die folgenden Tabellen zeigen die Sprachvergleiche für die sprachabhängigen Untertests. Referenz ist die deutschsprachige Testversion. Die französischsprachigen Teilnehmer (Tabelle 23) erzielten in fünf von sechs sprachabhängigen Untertests signifikant geringere Punktwerte als die deutschsprachigen Teilnehmer. Für die italienischsprachigen Kandidaten (Tabelle 24) sind vier Untertests signifikant unterschiedlich.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	1634	10.54	3.35	.000	Ja
	f	118	9.00	3.19		
Quantitative und formale Probleme	d	1634	10.69	3.62	.112	Nein
	f	118	10.14	3.50		
Textverständnis	d	1634	9.38	3.39	.004	Ja
	f	118	8.45	2.98		
Fakten lernen	d	1634	11.27	3.87	.003	Ja
	f	118	10.18	3.82		
Diagramme und Tabellen	d	1634	10.29	3.53	.001	Ja
	f	118	9.22	3.18		
Planen und Organisieren	d	1634	10.34	3.53	.009	Ja
	f	118	9.56	3.06		

Tabelle 23: Signifikanzprüfung deutsche vs. französische Sprachform.

Untertest	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	1634	10.54	3.35	.233	Nein
	i	79	10.08	3.25		
Quantitative und formale Probleme	d	1634	10.69	3.62	.007	Ja
	i	79	9.57	3.49		
Textverständnis	d	1634	9.38	3.39	.476	Nein
	i	79	9.66	3.06		
Fakten lernen	d	1634	11.27	3.87	.000	Ja
	i	79	9.94	3.11		
Diagramme und Tabellen	d	1634	10.29	3.53	.000	Ja
	i	79	8.68	3.03		
Planen und Organisieren	d	1634	10.34	3.53	.000	Ja
	i	79	8.51	3.43		

Tabelle 24: Signifikanzprüfung deutsche vs. italienische Sprachform.

Mit der Bestimmung der auszugleichenden Untertests ist noch nichts darüber ausgesagt, ob es sich bei den Differenzen um in der Stichprobe „real“ vorhandene Differenzen oder um mögliche (!) adaptationsbedingte Unterschiede handelt. Zu diesem Zweck ist eine Analyse der Aufgaben der betreffenden Untertests notwendig.

7.3.2.3 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. Dies würde bedeuten, dass sowohl die Itemschwierigkeiten wie auch deren Reihenfolge in beiden Sprachversionen vergleichbar sind.

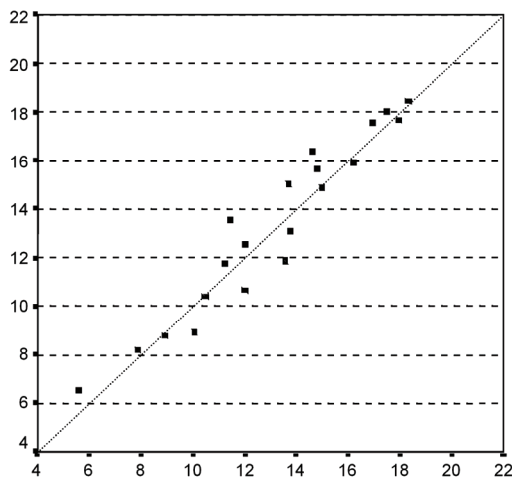


Abbildung 34: Beispiel für einen Delta-Plot.

Die Diagonale wird in den folgenden Abbildungen durch eine gestrichelte 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 demzufolge 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 Aufgaben 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.

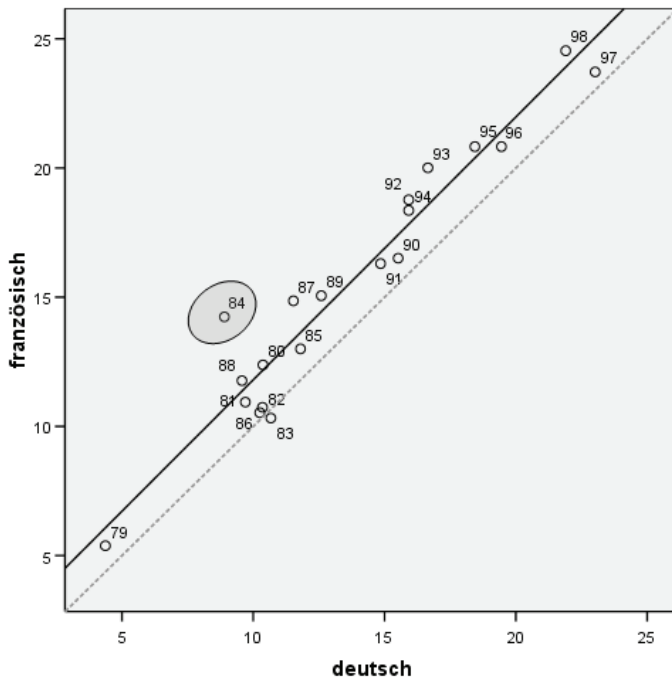
Regressionsgeraden werden in der Folge als durchgezogene Linien dargestellt. Massgeblich für die Entscheidung, ob bei einem Item DIF vorliegt, ist die Distanz des entsprechenden Punktes von dieser Geraden. Die Distanz D wird nach der Formel:

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

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.

7.3.2.4 Medizinisch-naturwissenschaftliches Grundverständnis



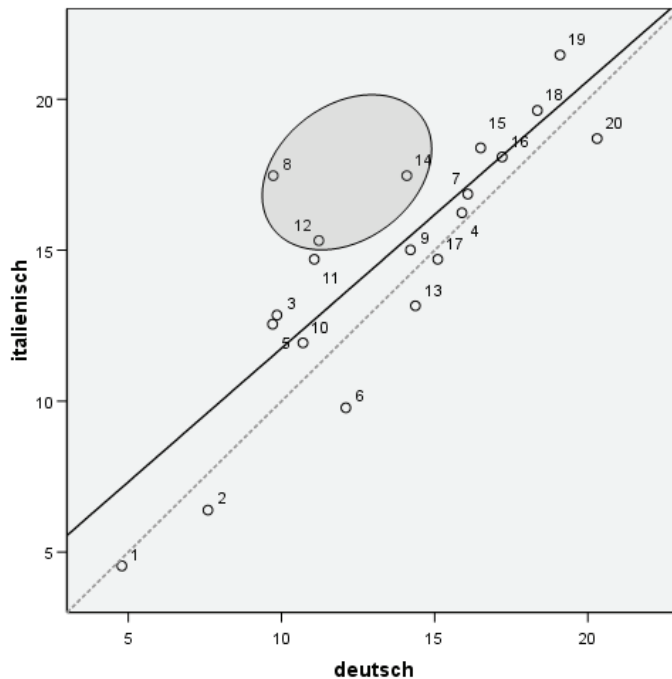
Im Untertest „medizinisch-naturwissenschaftliches Grundverständnis“ fällt ein Item auf, welches die kritische Distanz zur Regressionsgeraden überschreitet. Für die Frage 84 erhalten die französischsprachigen Teilnehmer entsprechende Bonuspunkte, wenn sie diese Aufgaben nicht gelöst haben. Insgesamt ist die Adaptation dieses Untertests sehr nahe am Original ausgefallen – die Distanzen zur Regressionsgeraden sind hier gering.

Abbildung 35: 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
79	0.91	0.86	0.92	4.37	5.38	3.93	0.49			
80	0.66	0.58	0.61	10.37	12.38	11.62	-0.14			
81	0.69	0.64	0.63	9.70	10.94	11.01	0.39			
82	0.66	0.65	0.46	10.25	10.53	15.32	1.07			
83	0.65	0.66	0.80	10.67	10.32	7.01	1.51			
84	0.72	0.50	0.77	8.90	14.24	7.62	-2.49		0.22	
85	0.60	0.55	0.43	11.80	13.00	15.93	0.44			
86	0.66	0.64	0.57	10.35	10.74	12.55	1.00			
87	0.61	0.47	0.63	11.53	14.86	11.01	-1.05			
88	0.69	0.60	0.67	9.57	11.77	10.08	-0.28			
89	0.57	0.47	0.48	12.59	15.06	14.70	-0.44			
90	0.45	0.41	0.49	15.52	16.51	14.39	0.64			
91	0.47	0.42	0.56	14.85	16.30	12.85	0.31			
92	0.43	0.31	0.44	15.92	18.77	15.62	-0.66			
93	0.40	0.26	0.39	16.65	20.01	16.86	-1.01			
94	0.43	0.33	0.41	15.92	18.36	16.55	-0.37			
95	0.33	0.23	0.13	18.44	20.83	23.32	-0.31			
96	0.29	0.23	0.29	19.45	20.83	19.32	0.41			
97	0.14	0.11	0.16	23.02	23.72	22.40	0.93			
98	0.18	0.08	0.23	21.90	24.54	20.86	-0.44			

Tabelle 25: DIF-Analyse „Medizinisch-naturwissenschaftliches Grundverständnis“ (d: deutsch; f: französisch; i: italienisch).

7.3.2.5 Quantitative und formale Probleme



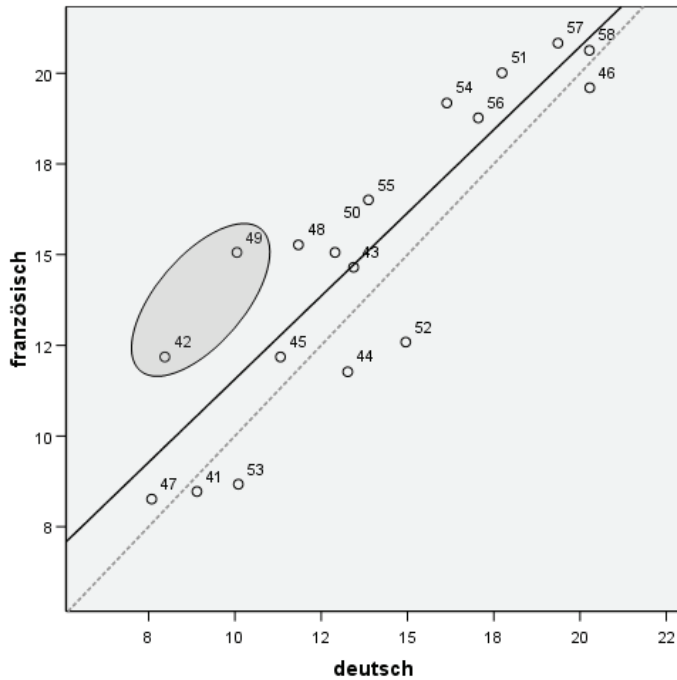
Die erzielten Punktwerte der deutsch- und italienischsprachigen Teilnehmer im Untertest „quantitative und formale Probleme“ unterscheiden sich signifikant. Auf Ebene der einzelnen Aufgaben können drei Items identifiziert werden, die auffällig von einer systematischen Verschiebung – und damit von der Schwierigkeitshierarchie der Referenzgruppe – abweichen.

Abbildung 36: Delta-Plot für „Quantitative und formale Probleme“ (deutsch-italienisch).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
1	0.89	0.88	0.90	4.79	4.97	4.54	Kein Ausgleich	1.94	Kein Ausgleich	
2	0.77	0.81	0.82	7.60	6.61	6.39		2.42		
3	0.68	0.72	0.56	9.85	8.88	12.85		-0.92		
4	0.43	0.44	0.42	15.89	15.68	16.24		0.55		
5	0.69	0.62	0.57	9.71	11.35	12.55		-0.78		
6	0.59	0.70	0.68	12.10	9.29	9.78		2.87		
7	0.42	0.47	0.39	16.08	15.06	16.86		0.21		
8	0.69	0.45	0.37	9.73	15.48	17.47		-4.46		0.32
9	0.50	0.51	0.47	14.21	14.03	15.01		0.35		
10	0.65	0.56	0.59	10.70	12.80	11.93		0.33		
11	0.63	0.49	0.48	11.07	14.45	14.70		-1.49		
12	0.62	0.55	0.46	11.22	13.00	15.32		-1.86		0.17
13	0.49	0.54	0.54	14.37	13.21	13.16		1.85		
14	0.51	0.42	0.37	14.09	16.09	17.47		-1.57		0.14
15	0.41	0.36	0.33	16.50	17.74	18.39		-0.66		
16	0.38	0.42	0.34	17.20	16.30	18.09		0.03		
17	0.46	0.45	0.48	15.10	15.48	14.70		1.18		
18	0.33	0.28	0.28	18.35	19.60	19.63		-0.36		
19	0.30	0.25	0.20	19.09	20.42	21.47		-1.25		
20	0.25	0.23	0.32	20.30	20.83	18.70		1.62		

Tabelle 26: DIF-Analyse „Quantitative und formale Probleme“ (d: deutsch; f: französisch; i: italienisch).

7.3.2.6 Textverständnis



Im Untertest „Textverständnis“ weisen im unteren Schwierigkeitsbereich zwei Items eine kritische Differenz auf (42 und 49). Die Items 41, 44, 52 und 53 wurden andererseits durch die Adaptation möglicherweise vereinfacht – diese werden jedoch nicht ausgeglichen.

Abbildung 37: Delta-Plot für „Textverständnis“ (deutsch-französisch).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
41	0.72	0.74	0.85	8.90	8.47	5.77	1.54			
42	0.76	0.58	0.70	7.97	12.18	9.47	-1.82		0.17	
43	0.53	0.48	0.62	13.44	14.65	11.31	0.05			
44	0.54	0.60	0.48	13.27	11.77	14.70	2.07			
45	0.62	0.58	0.63	11.32	12.18	11.01	0.44			
46	0.25	0.28	0.28	20.28	19.60	19.63	1.03			
47	0.77	0.75	0.85	7.59	8.26	5.77	0.81			
48	0.60	0.46	0.49	11.84	15.27	14.39	-1.48			
49	0.67	0.47	0.85	10.06	15.06	5.77	-2.54		0.21	
50	0.56	0.47	0.37	12.90	15.06	17.47	-0.62			
51	0.36	0.26	0.35	17.74	20.01	17.78	-0.99			
52	0.47	0.57	0.59	14.95	12.59	11.93	2.59			
53	0.67	0.73	0.75	10.10	8.67	8.24	2.20			
54	0.42	0.30	0.41	16.14	19.18	16.55	-1.46			
55	0.52	0.41	0.52	13.87	16.51	13.78	-1.03			
56	0.38	0.31	0.37	17.05	18.77	17.47	-0.55			
57	0.29	0.23	0.29	19.36	20.83	19.32	-0.51			
58	0.25	0.24	0.27	20.27	20.63	19.93	0.26			

Tabelle 27: DIF-Analyse „Textverständnis“ (d: deutsch; f: französisch; i: italienisch).

7.3.2.7 Fakten lernen

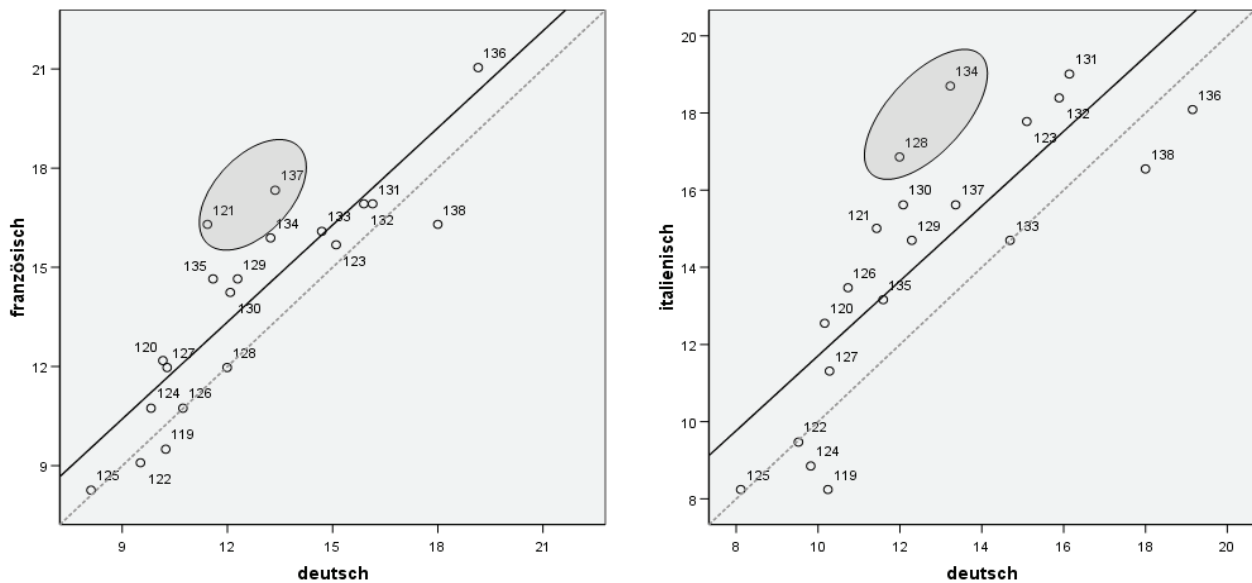


Abbildung 38: Delta-Plot für „Fakten lernen“ (deutsch-französisch und deutsch-italienisch).

Für die französisch- und italienischsprachigen Teilnehmer fallen hier jeweils zwei Aufgaben in den „auszugleichenden Bereich“. Für die italienischsprachige Gruppe fallen darüber hinaus besonders im sehr leichten Bereich eine Reihe von Items auf, die möglicherweise zusätzlich vereinfacht wurden.

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
119	0.66	0.69	0.75	10.24	9.50	8.24	1.52	2.66		
120	0.67	0.58	0.57	10.16	12.18	12.55	-0.45	-0.49		
121	0.62	0.42	0.47	11.43	16.30	15.01	-2.51	-1.38	0.2	
122	0.69	0.71	0.70	9.52	9.09	9.47	1.31	1.28		
123	0.46	0.44	0.35	15.10	15.68	17.78	0.50	-0.81		
124	0.68	0.64	0.72	9.82	10.74	8.85	0.34	1.93		
125	0.75	0.75	0.75	8.11	8.26	8.24	0.91	1.18		
126	0.64	0.64	0.53	10.73	10.74	13.47	0.98	-0.76		
127	0.66	0.59	0.62	10.28	11.97	11.31	-0.22	0.48		
128	0.59	0.59	0.39	11.99	11.97	16.86	0.98	-2.31		0.20
129	0.58	0.48	0.48	12.29	14.65	14.70	-0.73	-0.55		
130	0.59	0.50	0.44	12.08	14.24	15.62	-0.58	-1.36		
131	0.42	0.39	0.30	16.14	16.92	19.01	0.34	-0.97		
132	0.43	0.39	0.33	15.89	16.92	18.39	0.17	-0.70		
133	0.48	0.42	0.48	14.69	16.09	14.70	-0.09	1.11		
134	0.54	0.43	0.32	13.23	15.89	18.70	-0.96	-2.78		0.23
135	0.61	0.48	0.54	11.59	14.65	13.16	-1.22	0.06		
136	0.30	0.22	0.34	19.15	21.04	18.09	-0.50	1.79		
137	0.54	0.37	0.44	13.36	17.33	15.62	-1.90	-0.47	0.16	
138	0.35	0.42	0.41	18.00	16.30	16.55	2.09	2.10		

Tabelle 28: DIF-Analyse „Fakten lernen“ (d: deutsch; f: französisch; i: italienisch).

7.3.2.8 Diagramme und Tabellen

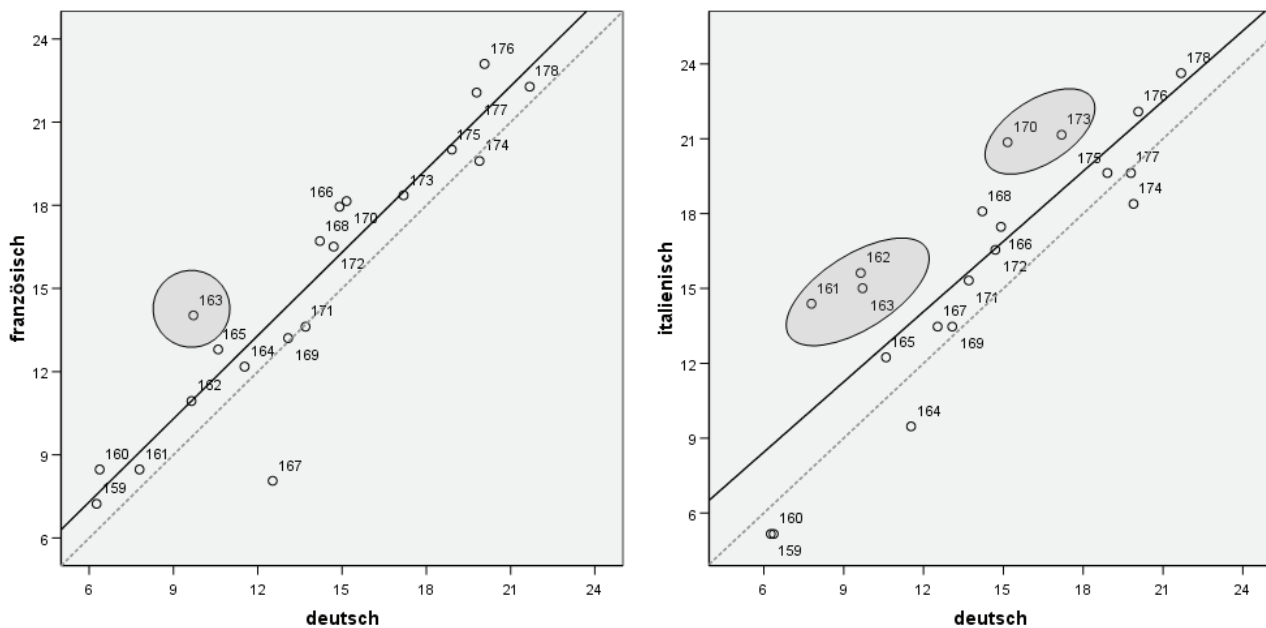


Abbildung 39: Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch und deutsch-italienisch).

In der französischsprachigen Gruppe fällt ein Item in den kritischen Bereich. Item 168 fällt dagegen in dieser Sprachgruppe massiv leichter aus. Eine Überprüfung dieses Items konnte keine Begründung für den beobachteten Effekt liefern.

In der italienischsprachigen Gruppe fallen 5 Items durch DIF auf.

Item	Schwierigkeit			Δ -Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
159	0.83	0.79	0.87	6.26	7.23	5.16	0.23	2.57		
160	0.82	0.74	0.87	6.37	8.47	5.16	-0.57	2.64		
161	0.76	0.74	0.49	7.79	8.47	14.39	0.44	-3.11		0.27
162	0.69	0.64	0.44	9.64	10.94	15.62	0.00	-2.75		0.25
163	0.69	0.51	0.47	9.71	14.03	15.01	-2.13	-2.25	0.18	0.22
164	0.61	0.58	0.70	11.53	12.18	9.47	0.46	3.03		
165	0.65	0.56	0.58	10.59	12.80	12.24	-0.64	0.37		
166	0.47	0.35	0.37	14.91	17.95	17.47	-1.23	-0.49		
167	0.57	0.75	0.53	12.53	8.06	13.47	4.08	0.80		
168	0.50	0.40	0.34	14.21	16.71	18.09	-0.85	-1.42		
169	0.55	0.54	0.53	13.08	13.21	13.47	0.83	1.18		
170	0.46	0.34	0.23	15.16	18.15	20.86	-1.20	-2.79		0.23
171	0.52	0.53	0.46	13.70	13.62	15.32	0.98	0.26		
172	0.48	0.41	0.41	14.70	16.51	16.55	-0.36	0.04		
173	0.38	0.33	0.22	17.19	18.36	21.16	0.09	-1.63		0.16
174	0.27	0.28	0.33	19.89	19.60	18.39	1.13	2.25		
175	0.31	0.26	0.28	18.91	20.01	19.63	0.14	0.68		
176	0.26	0.14	0.18	20.07	23.10	22.09	-1.22	-0.32		
177	0.27	0.18	0.28	19.79	22.07	19.63	-0.69	1.28		
178	0.19	0.17	0.11	21.68	22.28	23.63	0.50	-0.35		

Tabelle 29: DIF-Analyse „Diagramme und Tabellen“ (d: deutsch; f: französisch; i: italienisch).

7.3.2.9 Planen und Organisieren

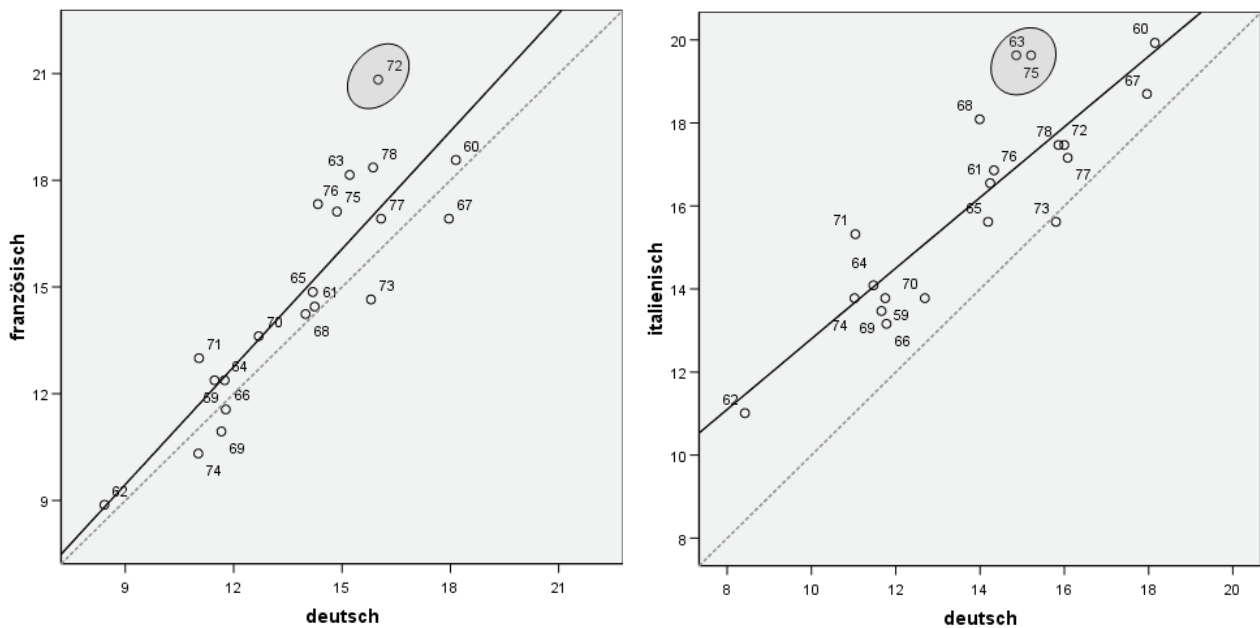


Abbildung 40: Delta-Plot für „Planen und Organisieren“ (deutsch-französisch).

In der französischsprachigen Gruppe weist ein Item (72) DIF auf, in der italienischsprachigen sind es zwei Items (63 und 75).

Item	Schwierigkeit			Δ-Wert			Differenz		Bonus	
	d	f	i	d	f	i	f	i	Bonus f	Bonus i
59	0.60	0.58	0.52	11.75	12.38	13.78	0.07	0.39		
60	0.34	0.32	0.27	18.15	18.57	19.93	0.66	-0.15		
61	0.50	0.49	0.41	14.24	14.45	16.55	0.53	-0.11		
62	0.74	0.72	0.63	8.42	8.88	11.01	-0.04	0.34		
63	0.46	0.34	0.28	15.21	18.15	19.63	-1.25	-1.82		0.18
64	0.61	0.58	0.51	11.47	12.38	14.09	-0.14	-0.03		
65	0.50	0.47	0.44	14.19	14.86	15.62	0.22	0.57		
66	0.60	0.61	0.54	11.78	11.56	13.16	0.65	0.88		
67	0.35	0.39	0.32	17.96	16.92	18.70	1.62	0.66		
68	0.51	0.50	0.34	13.99	14.24	18.09	0.48	-1.44		
69	0.61	0.64	0.53	11.66	10.94	13.47	0.97	0.57		
70	0.56	0.53	0.52	12.69	13.62	13.78	-0.06	1.00		
71	0.63	0.55	0.46	11.04	13.00	15.32	-0.87	-1.24		
72	0.43	0.23	0.37	16.00	20.83	17.47	-2.46	0.33	0.20	
73	0.44	0.48	0.44	15.80	14.65	15.62	1.55	1.61		
74	0.63	0.66	0.52	11.02	10.32	13.78	0.92	-0.08		
75	0.47	0.38	0.28	14.86	17.12	19.63	-0.81	-2.05		0.20
76	0.50	0.37	0.39	14.33	17.33	16.86	-1.35	-0.28		
77	0.42	0.39	0.38	16.08	16.92	17.16	0.23	0.62		
78	0.43	0.33	0.37	15.86	18.36	17.47	-0.90	0.24		

Tabelle 30: DIF-Analyse „Planen und Organisieren“ (d: deutsch; f: französisch; i: italienisch).

7.3.3 Effekte der Korrektur

Die durch den Ausgleich resultierenden Bonuspunkte können Tabelle 31 entnommen werden. Durchschnittlich werden den französischsprachigen Teilnehmern 0.7 Punkte, den italienischsprachigen Teilnehmern 1.52 Punkte gutgeschrieben. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der Bonus beträgt also 0 bis 1 Punkte für französischsprachige Teilnehmer bzw. 0 bis maximal 3 Punkte für italienischsprachige Teilnehmer. Nur eine italienischsprachige Person erhielt diesen Bonus von 3 Punkten. Die Auswirkungen der Korrektur sind damit wie in den Vorjahren insgesamt gering.

		n	Minimum	Maximum	m	s
Franz.	Textverständnis	118	.00	.38	.20	.13
	Planen und Organisieren	118	.00	.20	.14	.09
	Med. naturw. Grundverständnis	118	.00	.22	.10	.14
	Fakten	118	.00	.36	.17	.20
	Diagramme und Tabellen	118	.00	.18	.09	.13
	GESAMT	118	.00	1.34	.70	.31
Ital.	Quantitative und formale Probleme	79	.00	.63	.39	.21
	Planen und Organisieren	79	.00	.38	.25	.12
	Fakten	79	.00	.43	.23	.15
	Diagramme und Tabellen	79	.00	1.13	.66	.29
	GESAMT	79	.40	2.57	1.52	.47

Tabelle 31: Mittelwerte und Standardabweichungen der Korrekturwerte für die Personen der französisch- und italienischsprachigen Gruppen.

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Testwert ** $D > (F, I)$	deutsch	1634	109.38	21.40	41	176
	französisch	118	97.46	19.82	46.34	152.39
	italienisch	79	100.30	19.40	49.39	140.74
Muster zuordnen ** $D > F$	deutsch	1634	11.16	2.95	0	20
	französisch	118	10.29	2.86	3	18
	italienisch	79	10.90	3.12	0	18
Med.-naturwiss. Grundverständnis ** $(D, I) > F$	deutsch	1634	10.54	3.35	1	20
	französisch	118	9.10	3.14	1.22	17
	italienisch	79	10.08	3.25	4	17
Schlauchfiguren ** $D > (F, I)$	deutsch	1634	14.01	3.53	3	20
	französisch	118	12.08	3.82	2	20
	italienisch	79	12.61	3.51	4	20
Quantitative und formale Probleme <i>homogen</i>	deutsch	1634	10.69	3.62	1	20
	französisch	118	10.14	3.50	4	20
	italienisch	79	9.96	3.35	2.63	19.14
Textverständnis <i>homogen</i>	deutsch	1634	9.38	3.39	0	18
	französisch	118	8.65	2.91	0.38	15
	italienisch	79	9.66	3.06	2	16
Figuren lernen* $D > F$	deutsch	1634	11.35	3.38	0	20
	französisch	118	10.58	2.95	2	16
	italienisch	79	10.91	3.39	2	18
Fakten lernen ** $D > (F, I)$	deutsch	1634	11.27	3.87	1	20
	französisch	118	10.35	3.77	3.16	20
	italienisch	79	10.16	3.05	2.43	17
Diagramme und Tabellen ** $D > (F, I)$	deutsch	1634	10.29	3.53	1	20
	französisch	118	9.31	3.14	2.18	18
	italienisch	79	9.34	2.82	3.9	17.16
Konzentriertes und sorgf. Arbeiten ** $D > (F, I)$	deutsch	1634	10.34	4.71	0	20
	französisch	118	7.24	4.09	0	20
	italienisch	79	7.92	3.86	0	20
Planen und Organisieren** $D > I$	deutsch	1634	10.34	3.53	1	20
	französisch	118	9.70	3.06	3.2	17.2
	italienisch	79	8.76	3.39	2.2	17

Tabelle 32: Testwert und Punktwerte der Untertests für die Sprachgruppen. **/*: signifikanter Einfluss des Faktors „Sprache“ bei varianzanalytischer Prüfung (1%-, bzw. 5%-Niveau); darunter sind signifikante Unterschiede bei multiplem Mittelwertvergleich angegeben. In französischer und italienischer Sprachgruppe korrigierte Punktwerte.

Die Mittelwertsdifferenzen für jeden Untertest **nach** dem Sprachausgleich sind in Tabelle 32 dargestellt. Die Ergebnisse multipler Mittelwertvergleiche können 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.

In Abbildung 41 und 42 sind die Ausmasse der (ungerundeten) Bonuspunkte für die bisherigen Testdurchführungen dargestellt (1998 wurde kein DIF-Verfahren angewendet). Die senkrechten Linien bezeichnen die Spannweite in jedem Jahr (Minimum bis Maximum), der jeweilige Mittelwert wird durch die Markierung gekennzeichnet.

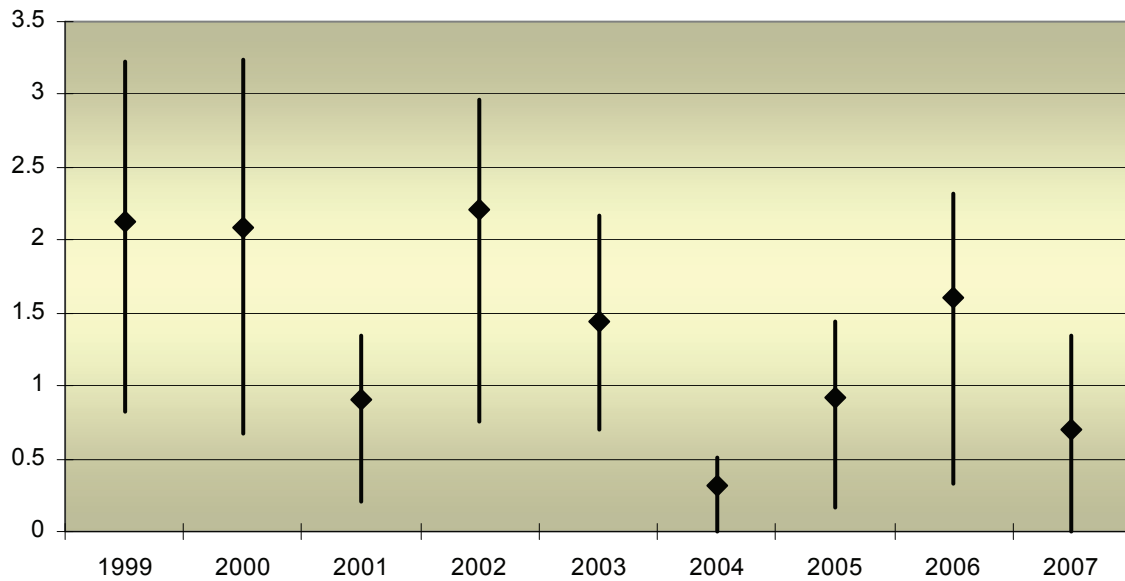


Abbildung 41: Bonuspunkte (ungerundet) über die Jahre für die französischsprachigen Kandidaten (Minimum – Mittelwert – Maximum).

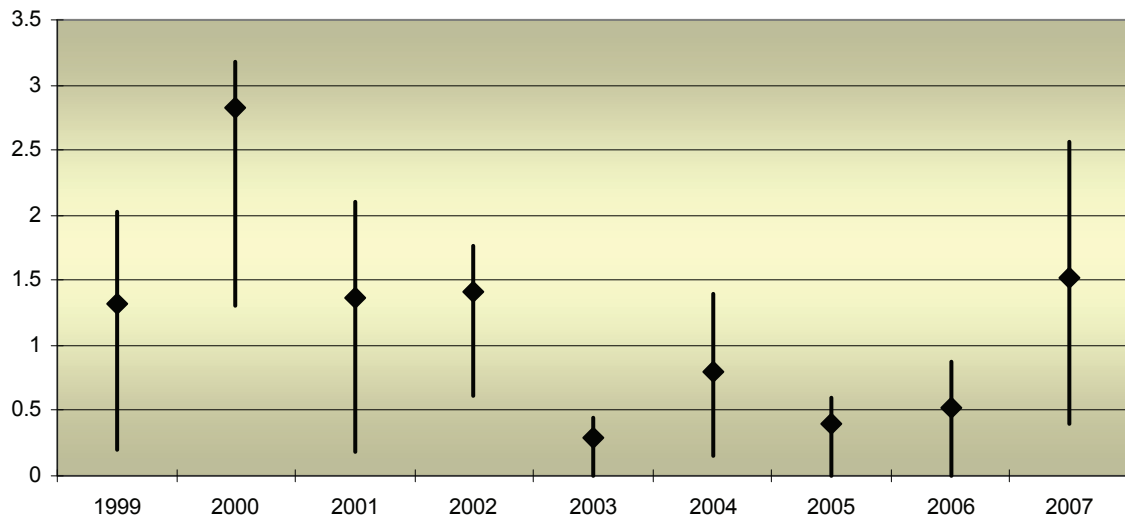


Abbildung 42: Bonuspunkte (ungerundet) über die Jahre für die italienischsprachigen Kandidaten (Minimum – Mittelwert – Maximum).

Der Vergleich für die französischsprachige Gruppe zeigt, dass es sich 2007 um einen Jahrgang mit vergleichsweise kleiner „Bonuspunktezuteilung“ handelt. Bei der italienischsprachigen Version fällt der Ausgleich etwas höher aus als in den vergangenen Durchführungen – liegt aber in der Grössenordnung der Jahre 1999 bis 2002.

Es muss sichergestellt sein, dass Unterschiede zwischen den Sprachgruppen nicht im Test selber begründet liegen. Dies kann durch eine Analyse des Verhältnisses zwischen sprachabhängigen und wenig sprachabhängigen Untertests belegt werden. Die Differenzen zwischen der deutschsprachigen Referenzgruppe und den anderssprachigen Teilnehmern werden nach folgender Formel berechnet:

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

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 geringere Punktzahlen erzielt als die deutschsprachige Referenzgruppe.

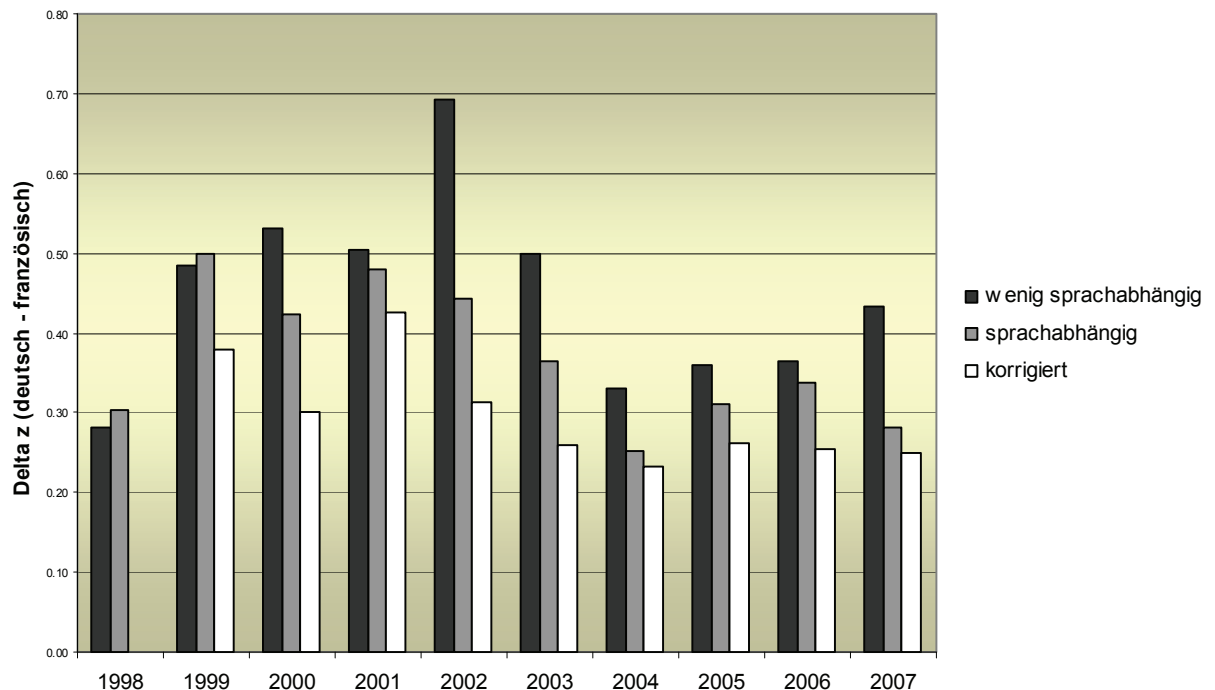


Abbildung 43: Abweichung (Delta z) für den Testwert zwischen deutscher und französischer Sprachgruppe. Für sprachabhängige Untertests sind korrigierte und unkorrigierte Werte dargestellt.

Erneut sind in der französischen Gruppe die Unterschiede in den wenig sprachabhängigen Untertests bereits vor dem Ausgleich grösser als jene der sprachabhängigen Untertests. Die Korrektur verringert diese zusätzlich, so dass eine Benachteiligung ausgeschlossen werden kann.

Die Leistungsunterschiede zwischen den Sprachgruppen werden also deutlicher von den weniger sprachabhängigen Untertests bestimmt. Da die Faktorenanalysen einen gemeinsamen Faktor für alle Untertests mit 40% Varianzaufklärung identifiziert haben, ist die Erwartung begründet, dass sich die Unterschiede in den wenig sprachabhängigen Untertests ebenfalls so in den sprachabhängigen Untertests zeigen.

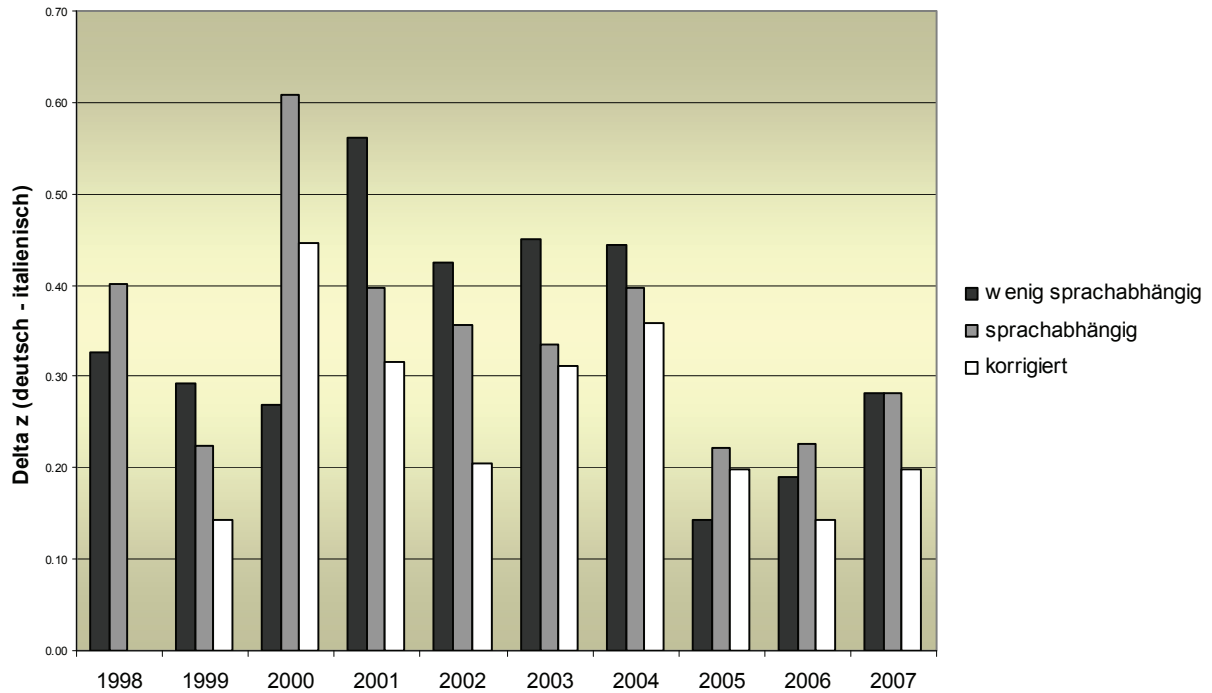


Abbildung 44: Abweichung (Delta z) für den Testwert zwischen deutscher und italienischer Sprachgruppe. Für sprachabhängige Untertests sind korrigierte und unkorrigierte Werte dargestellt.

Bei den italienischsprachigen Teilnehmern fallen in diesem Jahrgang die Unterschiede in den sprachabhängigen und wenig sprachabhängigen Untertests vor dem Ausgleich praktisch identisch aus. Durch den Ausgleich liegen die korrigierten Werte unter jenen der wenig sprachabhängigen Untertests – es darf also auch hier davon ausgegangen werden, dass diese Sprachgruppe nicht benachteiligt wurde.

7.4 Vergleichbarkeit der Testlokale

Die Durchführungsbedingungen an den einzelnen Testorten sind standardisiert. Aus den Protokollen der Testabnahme ergeben sich keine Hinweise, die als Einschränkungen dieser Chancengleichheit zu bewerten wären. Insbesondere die protokollierten Zeiten sind der Vorgabe entsprechend. Testwert und die Werte von drei Untertests unterscheiden sich signifikant zwischen den Testlokalen. Diese Unterschiede sind allerdings mit regionalen Unterschieden konfundiert.

In Tabelle 34 zeigt sich, dass nur ein einziges Testergebnis „Quantitative und formale Probleme“, Basel 1 deutlicher abweicht. Da auch eine genauere Prüfung keine Hinweise auf unterschiedliche Testbedingungen liefert, ergibt sich kein Hinweis auf eine Ungleichbehandlung. Mit 47 Personen ist das Testlokal vergleichsweise klein, sodass sich einzelne abweichende Werte hier stärker auswirken.

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testwert	<i>Zwischen Gruppen</i>	3592.559	18	199.587	2.037	0.006
	<i>Innerhalb Gruppen</i>	158264.122	1615	97.996		
	<i>Total</i>	161856.681	1633			
Muster zuordnen	<i>Zwischen Gruppen</i>	188.300	18	10.461	1.202	0.250
	<i>Innerhalb Gruppen</i>	14051.754	1615	8.701		
	<i>Total</i>	14240.054	1633			
Med.-naturwiss. Grundverständnis	<i>Zwischen Gruppen</i>	409.844	18	22.769	2.054	0.006
	<i>Innerhalb Gruppen</i>	17902.525	1615	11.085		
	<i>Total</i>	18312.370	1633			
Schlauchfiguren	<i>Zwischen Gruppen</i>	360.706	18	20.039	1.615	0.049
	<i>Innerhalb Gruppen</i>	20044.998	1615	12.412		
	<i>Total</i>	20405.704	1633			
Quant. und formale Probleme	<i>Zwischen Gruppen</i>	440.168	18	24.454	1.883	0.014
	<i>Innerhalb Gruppen</i>	20967.992	1615	12.983		
	<i>Total</i>	21408.160	1633			
Textverständnis	<i>Zwischen Gruppen</i>	306.112	18	17.006	1.484	0.086
	<i>Innerhalb Gruppen</i>	18507.355	1615	11.460		
	<i>Total</i>	18813.467	1633			
Figuren lernen	<i>Zwischen Gruppen</i>	260.305	18	14.461	1.271	0.197
	<i>Innerhalb Gruppen</i>	18376.354	1615	11.379		
	<i>Total</i>	18636.659	1633			
Fakten lernen	<i>Zwischen Gruppen</i>	309.243	18	17.180	1.150	0.297
	<i>Innerhalb Gruppen</i>	24132.171	1615	14.943		
	<i>Total</i>	24441.413	1633			
Diagramme und Tabellen	<i>Zwischen Gruppen</i>	287.217	18	15.957	1.282	0.190
	<i>Innerhalb Gruppen</i>	20103.282	1615	12.448		
	<i>Total</i>	20390.499	1633			
Konzentr. u. sorgf. Arbeiten	<i>Zwischen Gruppen</i>	575.186	18	31.955	1.445	0.101
	<i>Innerhalb Gruppen</i>	35720.337	1615	22.118		
	<i>Total</i>	36295.523	1633			
Planen und Organisieren	<i>Zwischen Gruppen</i>	328.407	18	18.245	1.469	0.092
	<i>Innerhalb Gruppen</i>	20058.084	1615	12.420		
	<i>Total</i>	20386.490	1633			

Tabelle 33: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den (deutschsprachigen) Testlokalen.

Die Mittelwerte der einzelnen Lokale sind in Tabelle 34 aufgeführt, wobei jene Werte jeweils dunkel schattiert sind, die im multiplen Mittelwertsvergleich nach Tukey signifikant tiefere als mindestens ein anderes Testlokal erzielten.

	Testwert	Muster zuordnen	Med.-naturwiss. Grundverständnis	Schlauchfiguren	Quant. und formale Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagramme und Tabellen	Konzentr. u. sorgf. Arbeiten	Planen und Organisieren	Lokalgrösse	
Testlokal	1	98.60	11.47	9.72	13.98	9.36	8.51	11.72	11.36	9.40	10.02	9.49	47
	2	100.01	11.27	9.80	14.11	10.08	8.98	12.01	11.77	9.88	10.78	9.46	107
	3	97.54	10.56	10.37	13.30	9.72	8.53	10.58	11.35	9.23	9.37	9.95	57
	4	101.36	11.31	11.17	14.51	10.25	9.93	11.83	11.19	10.34	9.88	10.61	59
	5	99.57	10.86	10.22	13.83	10.49	9.45	11.16	11.11	10.30	9.99	9.88	122
	6	102.13	11.18	10.90	14.51	10.98	9.87	11.73	11.81	10.88	10.19	10.77	124
	7	99.61	11.45	10.03	14.30	10.08	8.71	11.38	11.44	10.11	10.02	9.90	87
	8	99.32	10.88	10.73	13.63	10.25	8.70	11.36	10.88	10.46	9.75	10.11	56
	9	100.53	11.33	10.52	13.99	10.78	9.48	11.24	10.89	10.18	10.61	10.42	165
	10	103.75	11.93	11.33	14.48	11.23	10.33	11.40	11.68	10.95	11.73	11.20	40
	11	102.12	11.19	11.25	14.32	11.35	9.93	11.43	11.41	10.91	10.41	10.49	68
	12	100.61	11.21	10.82	13.26	10.77	9.56	11.35	10.91	10.06	11.18	10.30	66
	13	101.59	11.37	11.14	14.19	11.03	9.35	11.66	11.49	10.57	10.45	10.38	140
	14	103.74	11.31	11.05	14.83	11.65	9.91	11.94	12.08	11.00	11.43	10.89	65
	15	101.03	11.69	10.49	14.22	10.67	9.21	10.91	11.40	10.30	11.63	9.90	67
	16	100.00	11.37	10.27	13.02	10.83	8.95	11.68	11.29	10.59	10.20	10.34	41
	17	101.50	11.10	10.92	14.45	11.46	9.66	11.10	11.62	10.36	9.74	10.98	114
	18	99.82	11.18	10.26	13.91	10.38	9.48	10.60	10.72	10.14	10.49	10.74	65
	19	98.37	10.41	9.76	13.18	10.74	9.28	10.84	10.41	9.90	9.65	10.53	144

Tabelle 34: Darstellung der detaillierten Untertestergebnisse nach Testlokalen (deutschsprachig). Markiert: Mittelwerte für Testwert und die Punktwerte der Untertests im multiplen Mittelwertsvergleich nach Tukey, die signifikant tiefer als in mindestens einem anderen Testlokal liegen.

7.5 Vergleich für die 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. Beide Teilgruppen sind unterschiedlich repräsentativ in der Bewerberkohorte vertreten.

Der Männeranteil in der Humanmedizin sank zwischen 1999 bis 2002 immer weiter ab. 2003 war erstmals wieder ein geringer Anstieg des Anteils an Kandidaten zu verzeichnen, der aber 2004 weiter abfiel. 2007 stabilisiert sich der Männeranteil bei ca. 40%.

Die mittleren Testwerte (in Klammer die Standardabweichung) nach Geschlecht, Sprache und Disziplin können Tabelle 35 entnommen werden.

		Sprache			Disziplin			Gesamt
		deutsch	französisch	italienisch	HM	VM	ZM	
Geschlecht	männlich	101.8 (9.8) N=596	98.9 (9.4) N=38	97.2 (10.0) N=35	102.0 (9.7) N=552	100.5 (10.8) N=51	96.7 (8.8) N=66	101.4 (9.8) N=669
	weiblich	99.8 (10.0) N=1038	93.1 (8.5) N=80	95.6 (8.1) N=44	100.2 (9.9) N=822	97.5 (9.9) N=237	95.1 (9.8) N=103	99.2 (10.0) N=1162

Tabelle 35: Statistiken für den Testwert (1. Zeile Mittelwert, 2. Zeile Standardabweichung, 3. Zeile Personenzahl) nach Geschlechtern für Sprache und Disziplin.

Von 1998 bis 2006 betrug der Unterschied beim Testwert für die zahlenmässig grosse deutsche Sprachgruppe in der Schweiz zwischen 1.4 und 2.2 Punkten zugunsten der Männer (1.6, 1.5, 1.4, 2.2, 1.3, 1.4, 2.0, 1.6, 1.9). Der Unterschied von 2.0 Testwertpunkten im Jahr 2007 liegt also im Rahmen der Vorjahre.

		Humanmedizin			Veterinärmedizin			Zahnmedizin		
		Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch	Deutsch	Franz.	Italienisch
männlich	m	102.4	100.2	97.9	101.2	95.4	106.00	97.4	92.00	87.8
	s	9.7	9.9	9.4	11.4	7.0		8.4		12.2
	n	494	29	29	41	8	2	61	1	4
weiblich	m	100.5	95.8	96.2	99.1	91.8	95.1	95.5	90.8	88.00
	s	9.9	8.3	8.6	9.8	8.3	8.8	10.1	9.6	6.0
	n	768	28	26	181	46	10	89	6	8

Tabelle 36: Kombinierte Statistiken für den Testwert nach Geschlechtern, Sprache und Disziplin. *Kursiv: Wegen zu geringer Personenzahl nicht für Vergleichszwecke heranzuziehen.*

Sprache	Disziplin	Levene-Test Gleichheit der Varianzen			t-Test Gleichheit des Mittelwerts		
			F	Sig.	t	df	Sig.
Deutsch	Humanmedizin	Varianzen homogen	1.24	0.27	3.36	1260	.00
	Veterinärmedizin	Varianzen homogen	1.24	0.27	1.19	220	.23
	Zahnmedizin	Varianzen homogen	1.79	0.18	1.21	148	.23
Französisch	Humanmedizin	Varianzen homogen	1.33	0.25	1.83	55	.07
Italienisch	Humanmedizin	Varianzen homogen	0.39	0.54	0.67	53	.51

Tabelle 37: Prüfung zur Varianzhomogenität und der Mittelwertunterschiede (t-Test) für Testwert zwischen Geschlechtern. Französische Sprachgruppe Zahnmedizin und italienische Sprachgruppe Veterinär- und Zahnmedizin wegen Fallzahl nicht berechnet.

Untertest	Geschlecht	m	s	Levene-Test (Varianz)			t-Test Mittelwert		
					F	Sig.	t	df	Sig.
Muster zuordnen	männlich	10.8	3.0	homogen	0.11	0.74	-2.75	1829	.01
	weiblich	11.2	2.9						
Med.-naturwiss. Grundverst.	männlich	10.8	3.4	homogen	2.07	0.15	3.66	1829	.00
	weiblich	10.2	3.3						
Schlauchfiguren	männlich	14.6	3.4	nicht homogen	3.41	0.06	7.50	1464.22	.00
	weiblich	13.4	3.6						
Quantitative und formale Probleme	männlich	11.9	3.7	nicht homogen	9.76	0.00	11.63	1277.59	.00
	weiblich	9.9	3.3						
Textverständnis	männlich	9.6	3.3	homogen	0.00	0.99	2.66	1829	.01
	weiblich	9.2	3.4						
Figuren lernen	männlich	10.9	3.3	homogen	0.00	0.97	-3.24	1829	.00
	weiblich	11.5	3.4						
Fakten lernen	männlich	10.6	3.8	homogen	1.08	0.30	-4.81	1829	.00
	weiblich	11.5	3.9						
Diagramme und Tabellen	männlich	11.3	3.6	nicht homogen	8.77	0.00	10.34	1289.24	.00
	weiblich	9.5	3.3						
Konzent. u. sorgfältiges Arbeiten	männlich	10.2	4.7	homogen	0.03	0.85	1.40	1829	.16
	weiblich	9.9	4.7						
Planen und Organisieren	männlich	10.3	3.6	homogen	1.03	0.31	0.42	1829	.68
	weiblich	10.2	3.5						

Tabelle 38: Geschlechtsspezifische Mittelwerte (m), Standardabweichungen (s) und Ergebnisse der Prüfungen auf Varianzhomogenität und Mittelwertsunterschiede.

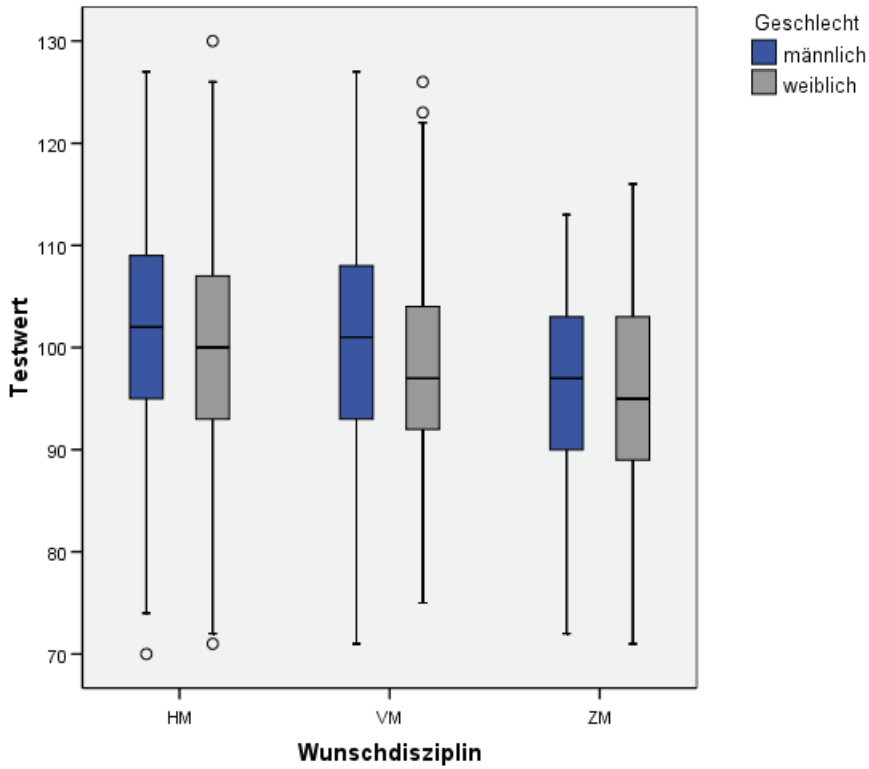


Abbildung 45: Boxplot für den Testwert nach Geschlecht und Wunschdisziplin (HM: Humanmedizin, VM: Veterinärmedizin, ZM: Zahnmedizin).

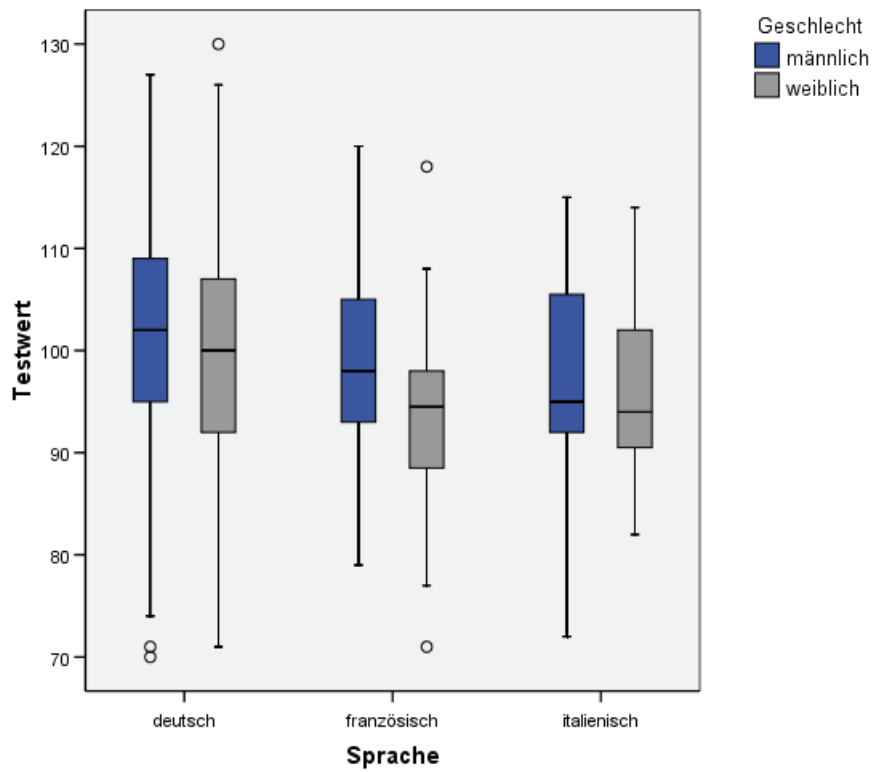


Abbildung 46: Boxplot für den Testwert nach Geschlecht, getrennt für die drei Sprachgruppen.

Die nachfolgenden Darstellungen verdeutlichen einerseits die Vergleichbarkeit der diesjährigen Ergebnisse mit früheren Jahrgängen (positive Werte bedeuten höhere Punktzahlen der männlichen Bewerber). Die Tendenz im Untertest „Konzentriertes und sorgfältiges Arbeiten“ zu höheren Punktzahlen der männlichen Testteilnehmer setzt sich 2007 (trotz der Rückkehr zu einer einfacheren Bearbeitungsanleitung) fort.

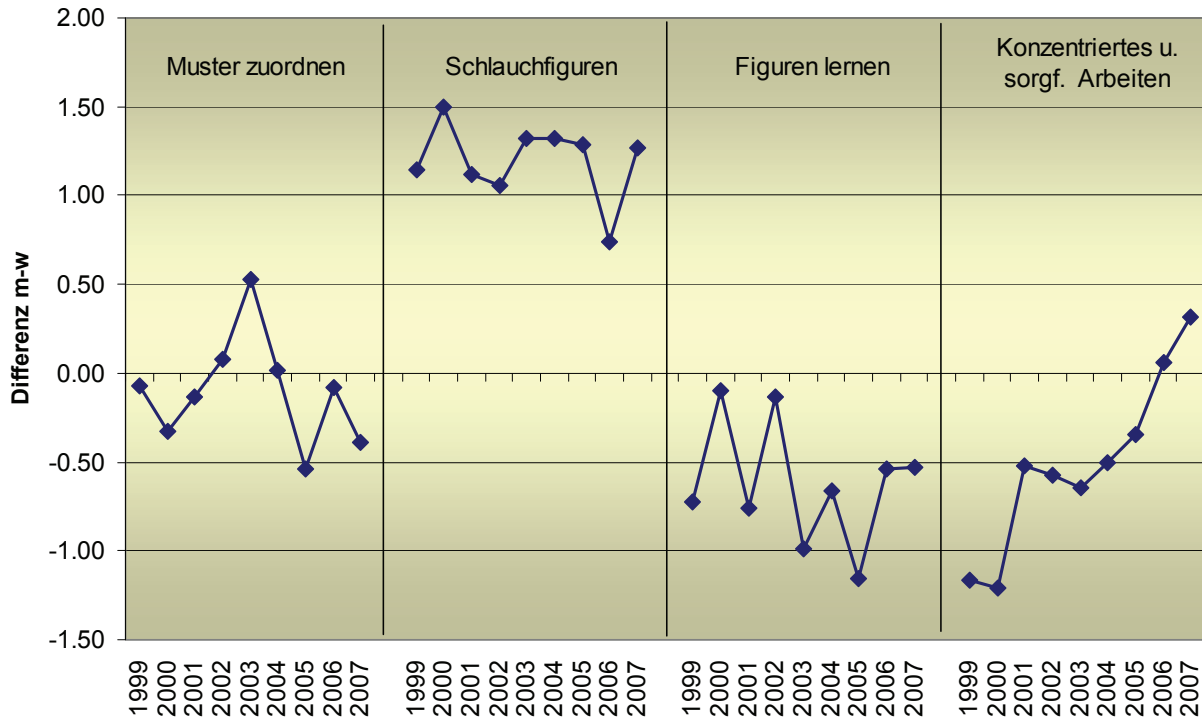


Abbildung 47: Differenzen für Punktwerte der wenig sprachabhängigen Untertests, geschlechtsspezifisch.

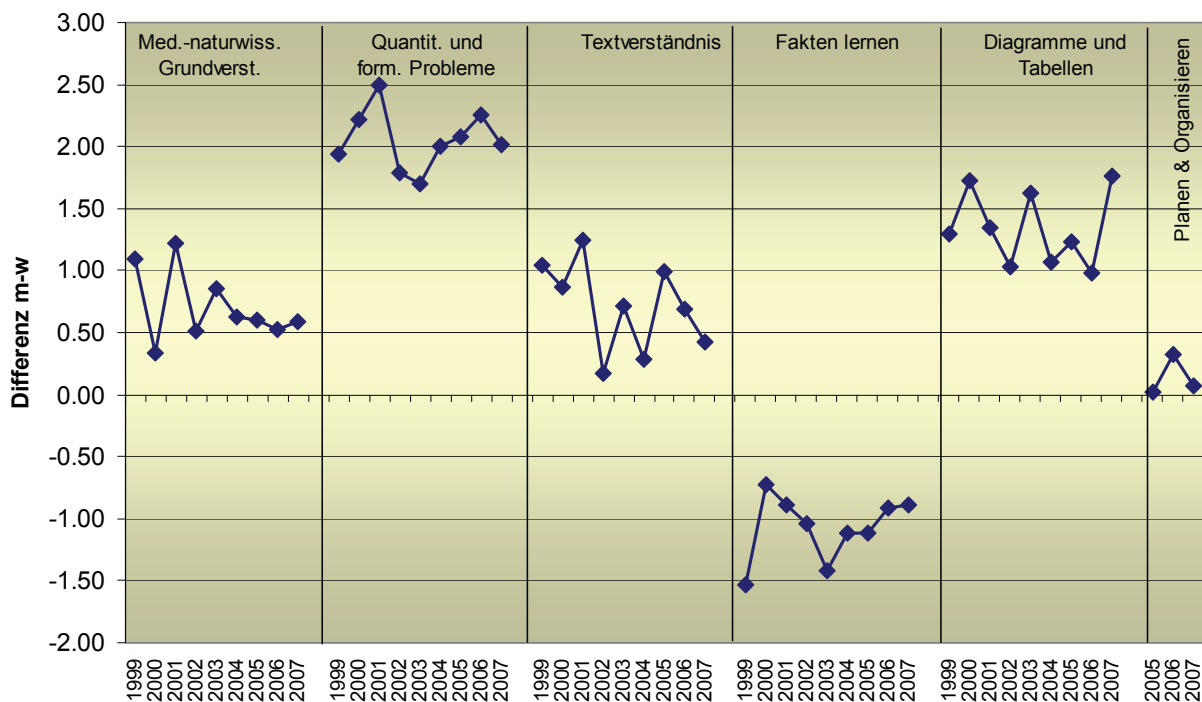


Abbildung 48: Differenzen für Punktwerte der sprachabhängigen Untertests, geschlechtsspezifisch.

7.6 Vergleiche für Altersgruppen

Für die Überprüfung der Alterseffekte wurden die Gruppen wie folgt gebildet. Die älteste Gruppe wurde nahe dem Median des Maturitätsalters geteilt.

Maturitätsjahr	Geburtsjahr			Total
	1954-1983	1984-1986	1987-1990	
1983	1			1
1986	3			3
1987	3			3
1988	2			2
1990	2			2
1991	1			1
1992	1			1
1993	2			2
1994	1			2
1995	2			3
1996	3			3
1997	5			5
1998	5			5
1999	8			8
2000	8			8
2001	8		2	10
2002	9	3	2	14
2003	15	20	14	49
2004	7	41	5	53
2005	6	107	24	137
2006	23	169	407	599
2007	46	89	785	920
	161	429	1241	1831

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

Wiederum zeigen die älteren Jahrgänge mit später Maturität tendenziell die niedrigsten Testwerte. Erneut unterscheiden sich aber die ältesten Jahrgänge mit früher Maturität nicht signifikant von den jüngsten Jahrgängen. Insbesondere in den Untertests „Quantitative und formale Probleme“ und „Diagramme und Tabellen“ erreichen diese Teilnehmer beste Ergebnisse. Das Alter der Teilnehmer ist also kein benachteiligender Faktor für das Testergebnis.

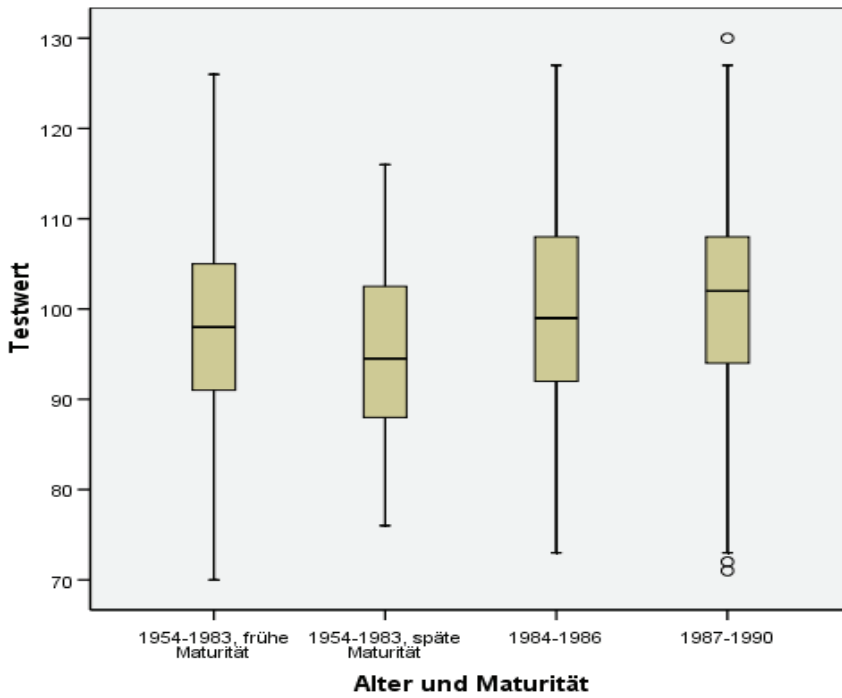


Abbildung 49: Boxplot für Testwert der Altersgruppen (Geburtsjahre) und Maturität.

7.7 Vergleiche nach Wunschuniversitäten

Die Leistungen der Kandidaten unterscheiden sich 2007 nicht signifikant hinsichtlich der bevorzugten Universitäten.

	Wunschuniversität	n	Mittelwert	Standard-abw.	Quadrat-summe	df	MQ	F	Sig.
Testwert	Bern	528	99.70	10.48	309.9	3	103.3	1.04	0.38
	Basel	383	99.71	9.63					
	Freiburg	113	99.25	9.19					
	Zürich	807	100.46	9.92					

Tabelle 41: Testwerte nach Wunschuniversitäten

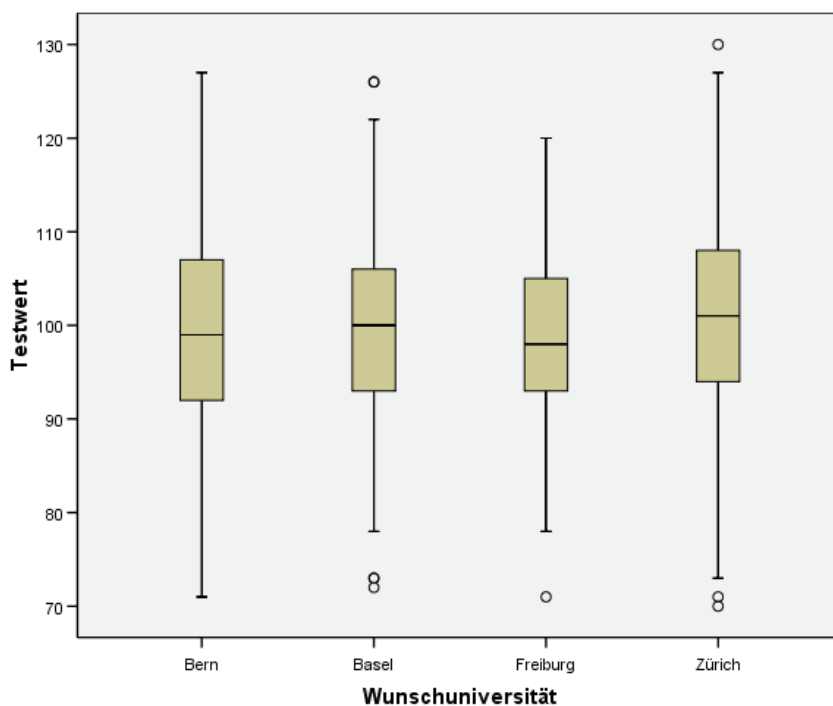


Abbildung 50: Boxplot für Testwerte nach Wunschuniversitäten.

8 Ergebnisse zur Testgüte

8.1 Zuverlässigkeit

Die Reliabilität kann anhand zweier Koeffizienten berechnet werden. Die Reliabilitätsschätzung nach der Testhalbierungsmethode (Teilung nach gerad- und ungeradzähligen Aufgaben) ist eine der gebräuchlichsten Zuverlässigkeitsschätzungen. Wegen der Stichprobengrößen werden nur die Ergebnisse des deutschsprachigen Tests der Schweiz verwendet. Die internen Konsistenzen (Cronbach Alpha) schätzen die Messgenauigkeit anhand der Korrelationen jeder Aufgabe mit allen anderen des entsprechenden Untertests.

Entscheidend sind die Kennwerte des **Punktwertes**, welcher nach der Standardisierung als Testwert für die Zulassung verwendet wird.

Die Zuverlässigkeitswerte des Punktwertes liegen 2007 mit 0.90 im gewohnt hohen Bereich der Vorjahre, in Österreich erneut sogar etwas höher (n=3 940). Auch die Konsistenzen der Testprofile liegen mit 0.80/0.84 in dem aus den Vorjahren gewohnten Bereich. Hierbei ist zu beachten, dass zu hohe Werte für sehr gleichartige Untertests (mit der Frage, ob man einzelne weglassen kann) sprechen würden, sehr niedrige Werte für eine heterogene Testbatterie, die nicht ohne weiteres zu einem Testwert zusammengefasst werden dürfte. Der Bereich um 0.80 scheint deshalb optimal, weil vergleichbare Werte auch in den Jahren mit einer erfolgreichen Evaluation des Zusammenhanges von Studienerfolg und Eignungstest gefunden worden sind. Dieser Wertebereich wird in den Folgejahren eingehalten. Es bestehen somit keine Einwände seitens der Zuverlässigkeit der Messung, den Punktwert bzw. Testwert für die Eignungsmessung zu verwenden.

	Reliabilität nach Testhalbierungsmethode										
	DE	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
DE / CH	.91 - .93	.92	.91	.90	.91	.91	.92	.91	.92	.90	.90
AT										.92	.93

Tabelle 42: Reliabilität des Punktwertes (Split Half) für Deutschland, Schweiz und Österreich.

	Konsistenz des Testprofiles									
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
CH	.83	.81	.81	.80	.80	.81	.78	.82	.79	.80
AT									.83	.84

Tabelle 43: Konsistenz des Testprofils für Schweiz und Österreich.

Die Zuverlässigkeitswerte der Untertestwerte wurden untersucht, um Bewertungen der Untertests vorzunehmen – sie sind für die Güte des Testwertes selbst nur indirekt bedeutsam. Hier müssten lediglich Eingriffe erfolgen, wenn der Testwert selbst eine zu geringe Zuverlässigkeit aufweisen würde (was aber ausgeschlossen werden kann).

Insgesamt werden für Schweizer Verhältnisse der Vorjahre vergleichbare Kennwerte erreicht. Der Untertest „Planen und Organisieren“ wurde aufgrund der Erfahrungen aus dem Jahr 2006 vereinfacht und entspricht wieder den Schwierigkeitsgraden der übrigen Untertests. Auch die Reliabilität dieses Untertests konnte dadurch gesteigert werden (Vermeidung von „Rateantworten“).

Untertest	Reliabilität nach Testhalbierungsmethode										
	Deutschland	Schweiz (d)									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Muster zuordnen	.67 - .73	.72	.61	.62	.59	.56	.55	.57	.70	.65	.62
Med.-nat. Grundverständnis	.65 - .72	.72	.71	.60	.69	.57	.73	.67	.73	.74	.65
Schlauchfiguren	.75 - .82	.75	.74	.74	.72	.67	.76	.74	.73	.71	.75
Quant. u. form. Probleme	.70 - .75	.78	.75	.74	.74	.71	.68	.72	.76	.71	.72
Textverständnis	.69 - .74	.75	.70	.75	.75	.71	.78	.70	.75	.68	.69
Figuren lernen	.72 - .75	.72	.70	.62	.62	.81	.78	.73	.74	.68	.64
Fakten lernen	.68 - .73	.70	.70	.61	.72	.69	.64	.72	.73	.75	.72
Diagramme und Tabellen	.68 - .72	.66	.62	.61	.63	.64	.62	.68	.71	.61	.70
Planen und Org.									.63	.42	.68

Tabelle 44: Zuverlässigkeit 1998 bis 2007; Vergleich der Reliabilitäten nach der Testhalbierungsmethode für die Untertests, Ergebnisse Deutschland aus Trost et al. (1997), S. 26 ff.

Untertest	Innere Konsistenz (Cronbachs Alpha)										
	Deutschland	Schweiz (d)									
		1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Muster zuordnen	.64 - .69	.69	.60	.53	.56	.57	.53	.54	.67	.60	.55
Med.-nat. Grundverständnis	.64 - .73	.70	.70	.61	.66	.57	.72	.66	.72	.71	.65
Schlauchfiguren	.76 - .80	.71	.74	.71	.73	.68	.74	.72	.72	.70	.73
Quant. u. form. Probleme	.69 - .71	.76	.75	.76	.74	.71	.67	.71	.76	.71	.70
Textverständnis	.68 - .71	.73	.75	.74	.70	.68	.78	.67	.73	.66	.69
Figuren lernen	.69 - .74	.72	.70	.62	.64	.82	.77	.71	.74	.66	.63
Fakten lernen	.62 - .70	.68	.72	.64	.69	.70	.73	.72	.72	.73	.73
Diagramme und Tabellen	.67 - .71	.67	.75	.63	.61	.62	.70	.66	.70	.61	.69
Planen und Org.									.62	.42	.63

Tabelle 45: Zuverlässigkeit 1998 bis 2007; Vergleich der Skalenskonsistenzen für die Untertests, Ergebnisse für Deutschland aus Trost et al. (1997), S. 26 ff.

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.

	Muster zuordnen	Med.-naturwis. Grundv.	Schlauchfiguren	Quant. u. formale Probl.	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arbeiten	Planen und Organisieren	Punkt-wert CH	Punkt-wert AT
Muster zuordnen		0.17	0.37	0.15	0.19	0.27	0.27	0.16	0.31	0.09	0.48	0.54
Med.-naturw. Grundverst.	0.22		0.28	0.52	0.59	0.20	0.21	0.58	0.17	0.43	0.67	0.70
Schlauchfiguren	0.40	0.33		0.35	0.28	0.35	0.32	0.32	0.32	0.20	0.63	0.68
Quant. u. formale Probl.	0.25	0.55	0.40		0.48	0.16	0.18	0.63	0.18	0.47	0.68	0.72
Textverständnis	0.23	0.63	0.36	0.51		0.22	0.24	0.52	0.16	0.40	0.66	0.69
Figuren lernen	0.36	0.25	0.42	0.25	0.27		0.40	0.19	0.29	0.15	0.54	0.59
Fakten lernen	0.28	0.30	0.34	0.28	0.31	0.41		0.19	0.24	0.19	0.55	0.58
Diagramme und Tabellen	0.25	0.60	0.39	0.64	0.54	0.26	0.27		0.17	0.47	0.69	0.73
Konzent. u. sorgfält. Arbeiten	0.35	0.20	0.38	0.27	0.21	0.33	0.25	0.24		0.15	0.54	0.57
Planen und Organisieren	0.20	0.48	0.30	0.49	0.45	0.23	0.27	0.53	0.22		0.58	0.64

Tabelle 46: Korrelationen zwischen Punktwerten der Untertests CH (Schweiz, über der Diagonale) und AT (Österreich, unter der Diagonale) sowie mit dem Gesamtwert.

Die Korrelationen der einzelnen Untertests mit der Gesamtpunktzahl bleiben über die Jahre recht stabil. Der Scree-Plot stellt den Eigenwerteverlauf der Faktoren grafisch dar. Die Ein-Faktorenlösung bleibt deutlich hervorgehoben:

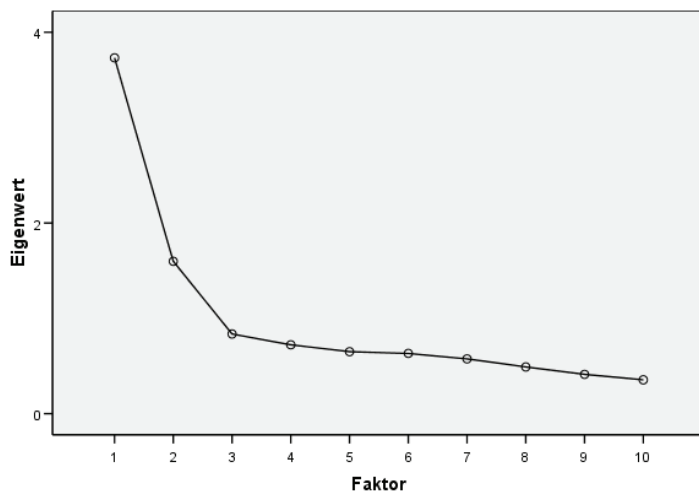


Abbildung 51: Scree-Plot (Eigenwerteverlauf) der Faktoren 1 bis 10 für die Untertests des EMS 2007.

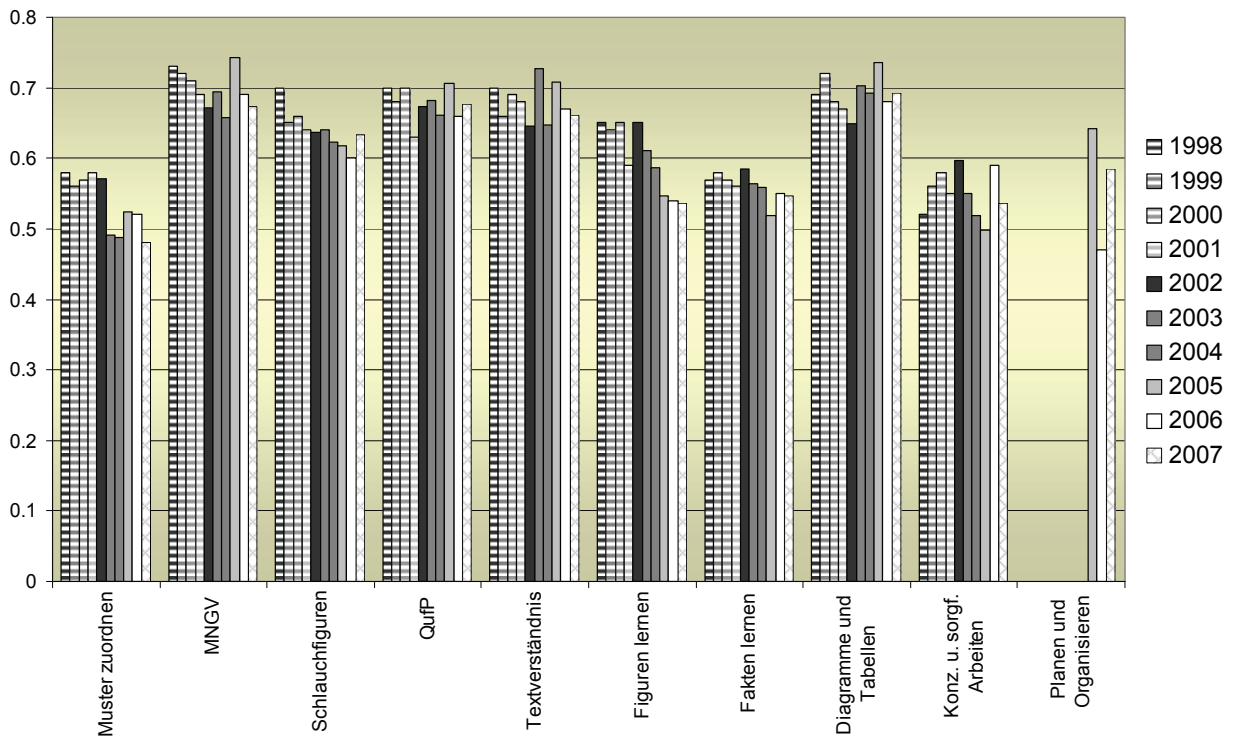


Abbildung 52: Korrelationen zwischen Punktwerten und Untertests nach Jahren.

Unrotierte Lösung															
Faktor	Eigenwert					% Varianz					Kumuliert %				
	2007	2006	2005	2004	2003	2007	2006	2005	2004	2003	2007	2006	2005	2004	2003
1	3.73	3.65	4.04	3.36	3.64	37.32	36.50	40.38	37.3	40.5	37.32	36.50	40.38	37.3	40.5
2	1.60	1.61	1.48	1.39	1.6	15.98	16.10	14.77	15.4	17.8	53.29	52.61	55.14	52.7	58.3
3	0.84	0.85	0.91	0.9	0.92	8.35	8.54	9.15	10.1	10.3	61.65	61.15	64.29	62.8	68.5
4	0.72	0.77	0.71	0.82	0.67	7.22	7.72	7.09	9.06	7.5	68.87	68.87	71.38	71.9	76
5	0.65	0.65	0.61	0.62	0.54	6.50	6.49	6.06	6.9	5.95	75.37	75.37	77.44	78.8	82
6	0.63	0.61	0.58	0.52	0.5	6.32	6.08	5.82	6.36	5.54	81.69	81.45	83.26	85.1	87.5
7	0.57	0.56	0.53	0.54	0.48	5.75	5.55	5.26	6.03	5.38	87.43	87.00	88.52	91.1	92.9
8	0.49	0.52	0.47	0.47	0.35	4.90	5.19	4.69	4.62	3.93	92.33	92.19	93.21	95.8	96.8
9	0.41	0.41	0.37	0.38	0.29	4.12	4.11	3.74	4.24	3.17	96.45	96.30	96.95	100	100
10	0.35	0.37	0.30			3.55	3.70	3.05			100	100	100		

Tabelle 47: Faktorenanalyse: Verlauf der unrotierten Lösung (Schweiz).

Varimax-rotierte Lösungen															
Eigenwert					% Varianz					Kumuliert %					
2007	2006	2005	2004	2003	2007	2006	2005	2004	2003	2007	2006	2005	2004	2003	
Zwei-Faktorenlösung															
3.05	2.91	3.26	2.63	2.87	30.54	29.12	32.58	29.26	31.9	30.54	29.12	32.58	29.26	31.9	
2.28	2.35	2.27	2.11	2.37	22.75	23.49	22.70	23.48	26.36	53.29	52.61	55.28	52.74	58.26	
Drei-Faktorenlösung															
3.05	2.91	3.27	2.56	2.84	30.46	29.05	32.74	28.39	31.57	30.46	29.05	32.74	28.39	31.57	
1.68	1.81	1.73	1.93	1.83	16.80	18.13	17.31	21.43	20.34	47.26	47.18	50.05	49.81	51.91	
1.44	1.40	1.43	1.17	1.5	14.39	13.97	14.33	12.97	16.62	61.65	61.15	64.38	62.79	68.54	

Tabelle 48: Varianzanteile der einzelnen Faktorenlösungen Schweiz 2003-2007 (rotierte Lösungen).

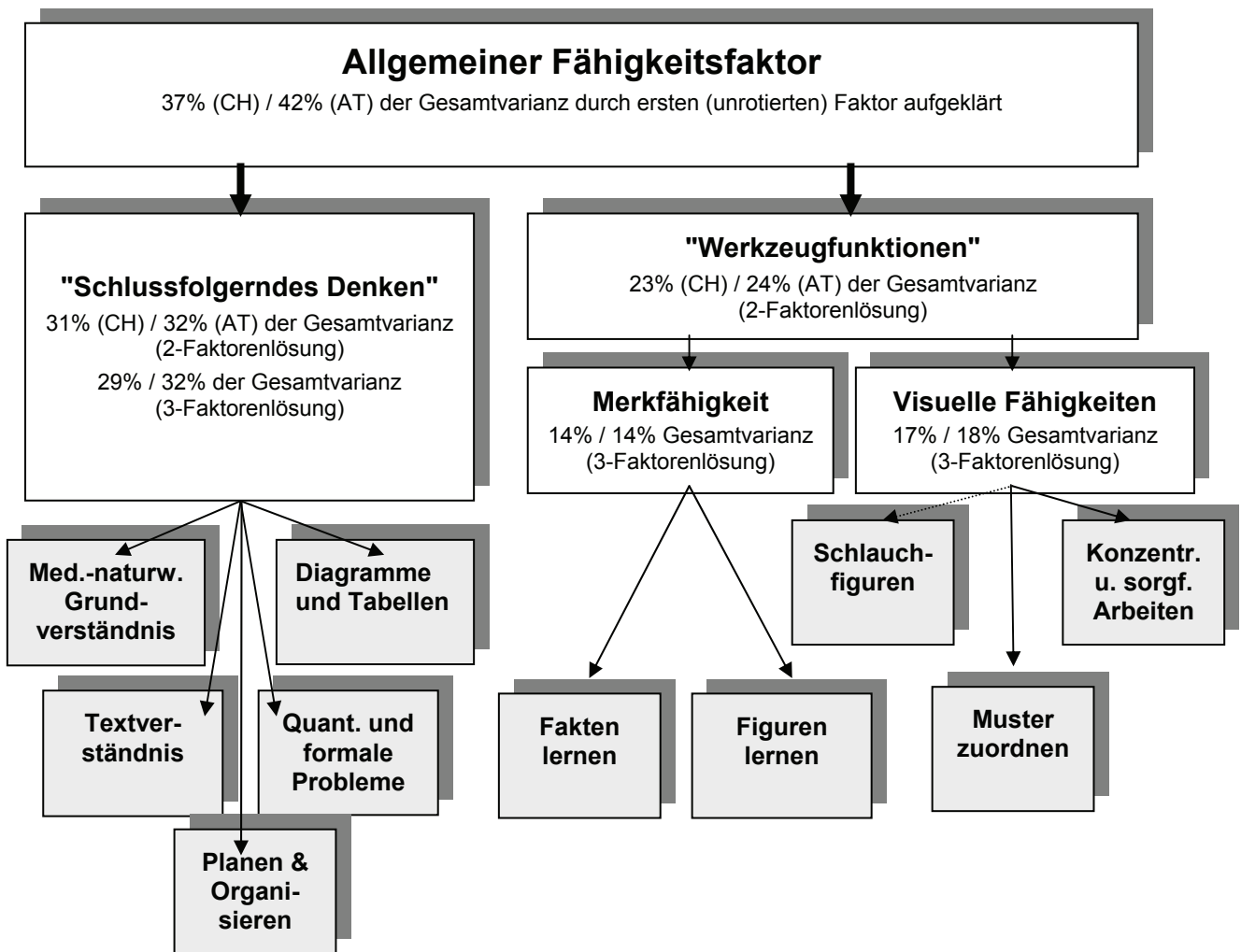


Abbildung 53: Struktur der Untertests des EMS, hierarchische Faktorenstruktur 2007 CH (Schweiz) und AT (Österreich).

Die Struktur des EMS bleibt weiterhin stabil – trotz Einführung eines neuen Untertests („Planen und Organisieren“, 2005) und der laufenden Neufassung der Instruktionen zum Untertest „Konzentriertes und sorgfältiges Arbeiten“.

Die Leistungen im gesamten Test werden in Form eines „allgemeinen Fähigkeitsfaktors“ am stärksten durch die Untertests „medizinisch-naturwissenschaftliches Grundverständnis“, „Diagramme und Tabellen“, „Textverständnis“ sowie „Quantitative und formale Probleme“ charakterisiert.

In der Zweifaktorenlösung stechen die Untertests „Diagramme und Tabellen“, „medizinisch-naturwissenschaftliches Grundverständnis“ sowie „Quantitative und formale Probleme“ für den Faktor „Schlussfolgerndes Denken“ und der Untertest „Figuren lernen“ als typischer Repräsentant für den Faktor „Werkzeugfunktionen“ hervor.

Die Dreifaktorenlösung teilt den Faktor „Werkzeugfunktionen“ in „Merkfähigkeit“ (repräsentiert durch „Fakten lernen“ und „Figuren lernen“) und „Visuelle Fähigkeiten“ („Konzentriertes und sorgfältiges Arbeiten“, „Muster zuordnen“, „Schlauchfiguren“).

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Muster zuordnen	.56	.53	.57	.57	.71	.48	.45	.46	.47	.43
Med.-naturwiss. Grundverständnis	.76	.76	.74	.74	.7	.74	.71	.79	.74	.74
Schlauchfiguren	.69	.64	.65	.62	.69	.61	.61	.58	.57	.60
Quant. und formale Probleme	.7	.7	.72	.66	.68	.73	.7	.75	.71	.73
Textverständnis	.72	.7	.71	.72	.62	.77	.7	.76	.72	.72
Figuren lernen	.62	.59	.53	.56	.62	.55	.54	.48	.47	.48
Fakten lernen	.53	.52	.55	.52	.57	.51	.51	.47	.48	.48
Diagramme und Tabellen	.72	.75	.71	.72	.56	.76	.74	.78	.74	.76
Konzentr. u. sorgf. Arbeiten	.47	.51	.55	.49	.54	.48	.44	.41	.52	.43
Planen und Organisieren								.68	.51	.62

Tabelle 49: Ladungen im ersten Faktor der unrotierten Lösung 1998 bis 2007 („Generalfaktor“ bzw. allgemeiner Fähigkeitsfaktor).

	Faktor 1						Faktor 2						Kommunalitäten (h ²)						
	02	03	04	05	06	07	02	03	04	05	06	07	02	03	04	05	06	07	
Diagramme und Tabellen	.77	.83	.79	.81	.80	.82	-.17	-.17					.63	.73	.66	.70	.66	.69	
Med.-naturwiss. Grundverständnis	.77	.86	.80	.82	.77	.79	.21	-.11					.63	.76	.65	.71	.64	.64	
Textverständnis	.82	.83	.70	.81	.77	.74	-.11	-.18					.68	.74	.63	.68	.62	.58	
Quant. und formale Probleme	.79	.77	.70	.78	.79	.79	-.18	.20					.65	.65	.60	.64	.64	.64	
Planen und Organisieren				.72	.61	.69										.54	.37	.49	
Konzentr. u. sorgf. Arbeiten							.82	.79	.33	.51	.62	.64	.68	.62	.20	.28	.41	.41	
Figuren lernen	.20						.70	.78	.74	.74	.73	.69	.52	.65	.56	.56	.54	.49	
Muster zuordnen	.20						.63	.58	.70	.74	.66	.67	.43	.62	.47	.56	.45	.46	
Fakten lernen		.26					.67	.50	.67	.62	.64	.64	.47	.78	.47	.40	.42	.44	
Schlauchfiguren	.29	.26		.29		.29	.60	.66	.67	.62	.68	.65	.45	.60	.52	.47	.50	.50	

Tabelle 50: Ladungen und Kommunalitäten der Zwei-Faktorenlösung 2002 bis 2007, varimaxrotiert.

Faktor	Schweiz								Deutschland		
	1		2		3		h ²		1	2	3
	2006	2007	2006	2007	2006	2007	2006	2007			
Diagramme und Tabellen	.80	.82					.67	.67	.82		
Medizin.-naturwiss. Grundverständnis	.79	.79			.25		.68	.69	.81		.2
Quantitat. u. formale Probleme	.79	.79					.64	.64	.8	.18	
Textverständnis	.77	.73					.64	.66	.79		.2
Planen und Organisieren	.57	.69	.30		-.30		.51	.42	-	-	-
Muster zuordnen			.71	.76			.54	.69		.81	
Konzentr. u. sorgfält. Arbeiten			.72	.74			.56	.63		.7	.4
Schlauchfiguren			.69	.61	.23	.28	.56	.56	.35	.71	
Fakten lernen					.77	.83	.66	.73	.21		.87
Figuren lernen			.36		.71	.76	.63	.63	.13	.47	.64

Tabelle 51: Faktorenanalyse: Varimaxrotierte Drei-Faktorenlösung, Schweiz 2006 und 2007 sowie Vergleichswerte Deutschland (deutsche Daten nach Blum, 1996, in Trost et al., 1998, S. 42).

8.3 Item-Trennschärfen

Die Item-Trennschärfen sind die Korrelationen des Punktwertes für den jeweiligen Untertest mit den zugeordneten Items. Positive Korrelationen weisen darauf hin, dass die Leistungsbesten im jeweiligen Untertest auch beim entsprechenden Item die richtige Lösung bevorzugt gewählt haben. Zu beachten ist, dass bei sehr leichten und sehr schwierigen Items wegen der geringeren Antwortvarianz auch die Trennschärfe in der Regel geringer ausfallen wird. Negative Trennschärfen würden auf Items hinweisen, die missverständlich formuliert sind oder keine eindeutige Lösung haben – die Leistungsbesten eine andere als die vorgegebene Lösung gewählt haben. Entsprechend der festgelegten Auswertedirektive des EMS werden solche Items von der Auswertung nachträglich ausgeschlossen und nicht gewertet, um Risiken der Fehlbewertung auszuschließen. Dies war 2007 nicht notwendig.

Die nachfolgenden Abbildungen zeigen die Trennschärfen. Die Item-Trennschärfen (korrigierte Item-Total-Korrelation) folgen weitgehend den deutschen Kennwerten.

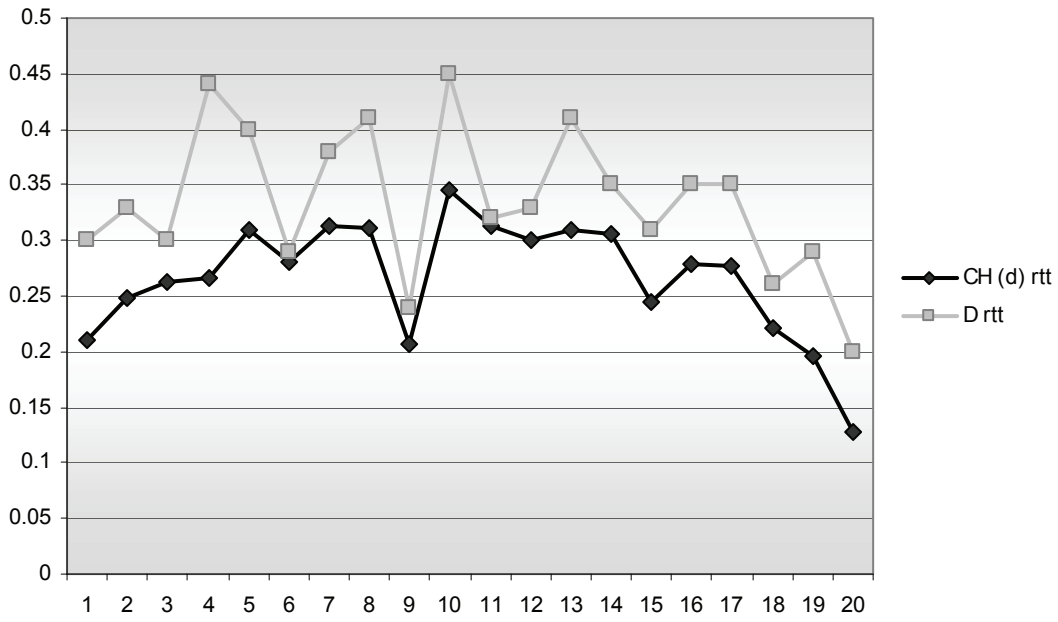


Abbildung 54: Trennschärfen für den Untertest „Quantitative und formale Probleme“.

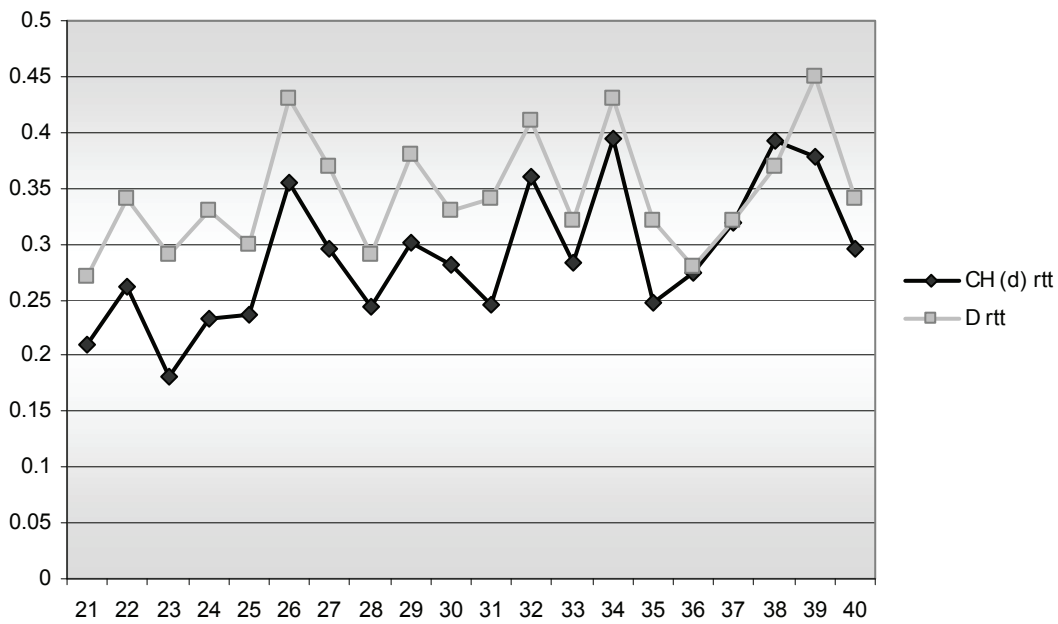


Abbildung 55: Trennschärfen für den Untertest „Schlauchfiguren“.

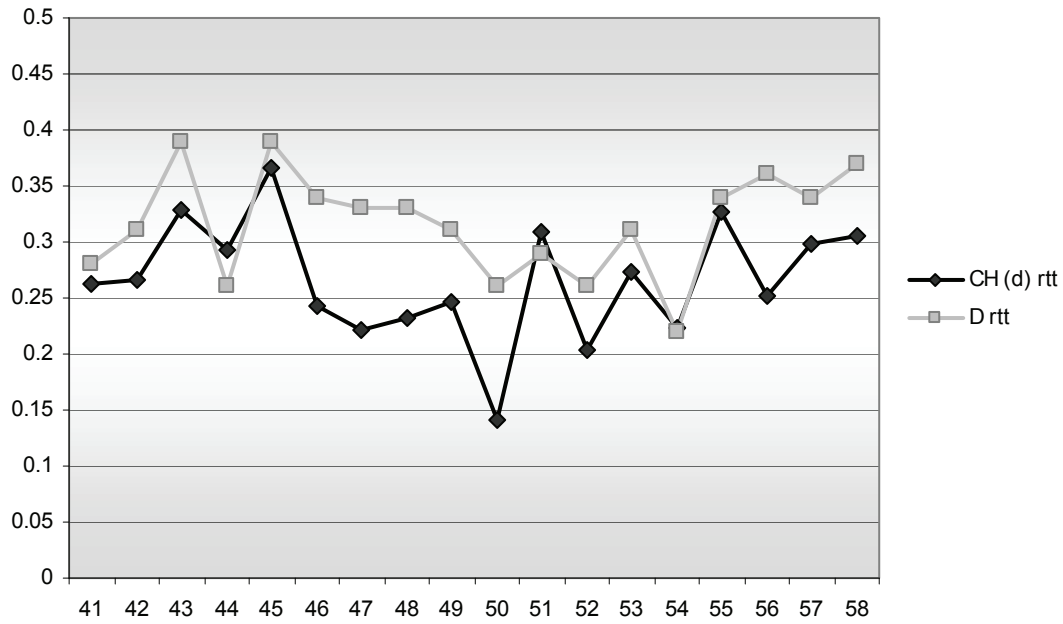


Abbildung 56: Trennschärfen für den Untertest „Textverständnis“.

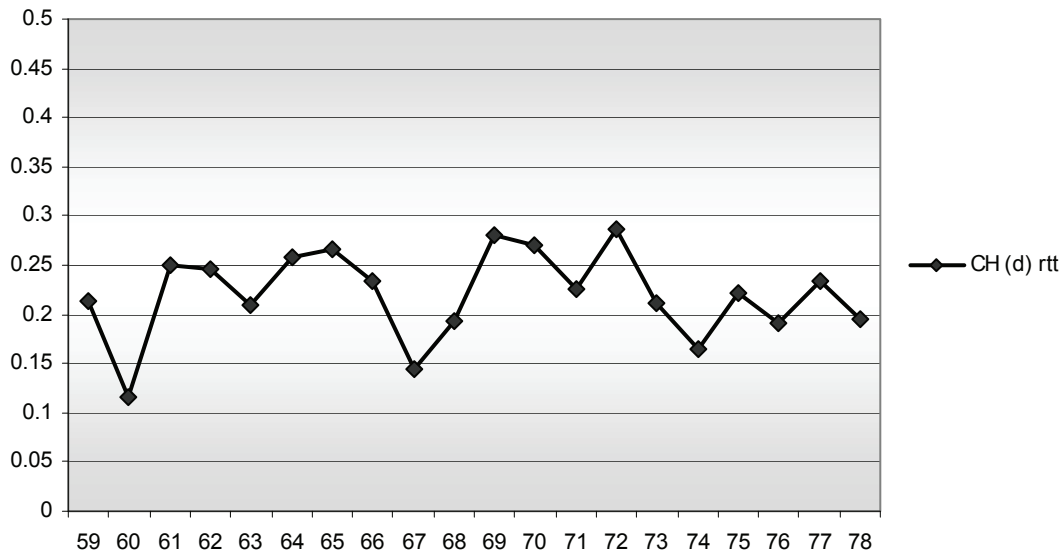


Abbildung 57: Trennschärfen für den Untertest „Planen und Organisieren“ (keine deutschen Vergleichswerte).

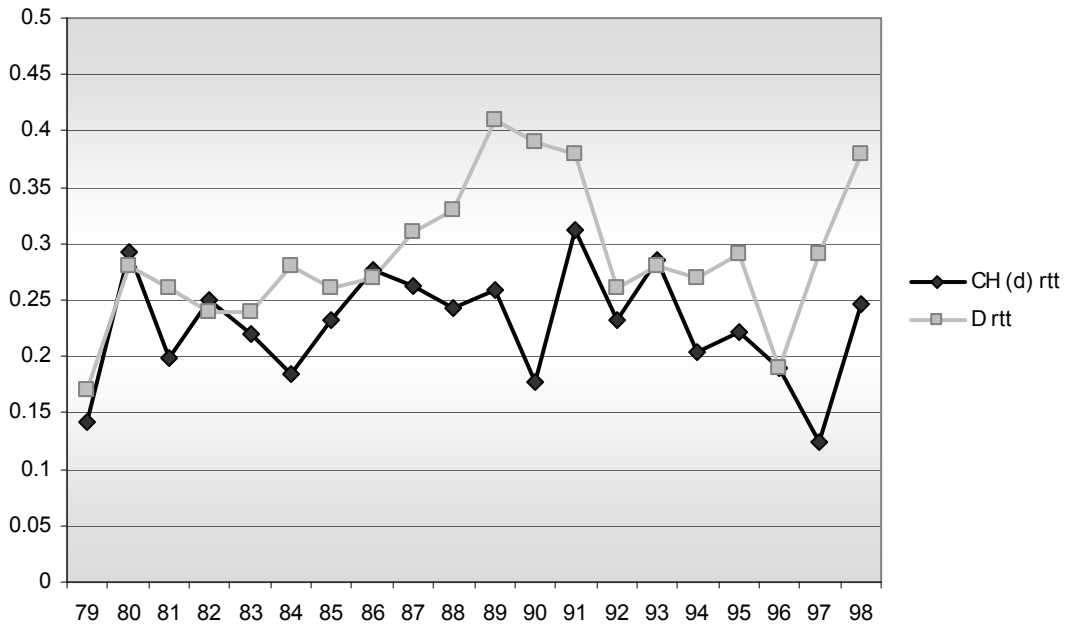


Abbildung 58: Trennschärpen für den Untertest „Medizinisch-naturwissenschaftliches Grundverständnis“.

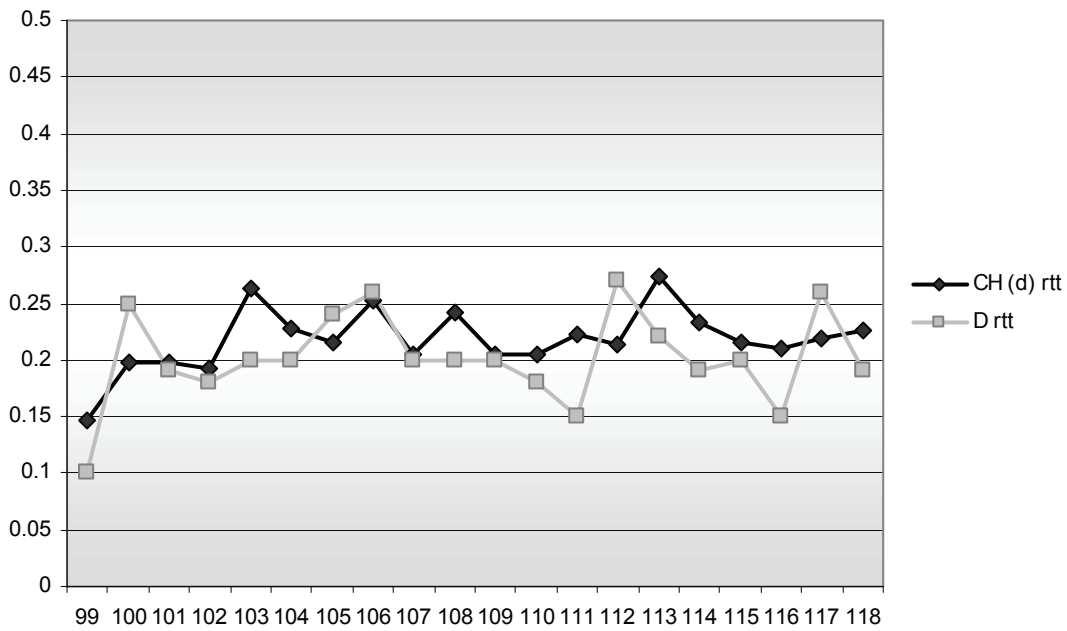


Abbildung 59: Trennschärpen für den Untertest „Figuren lernen“.

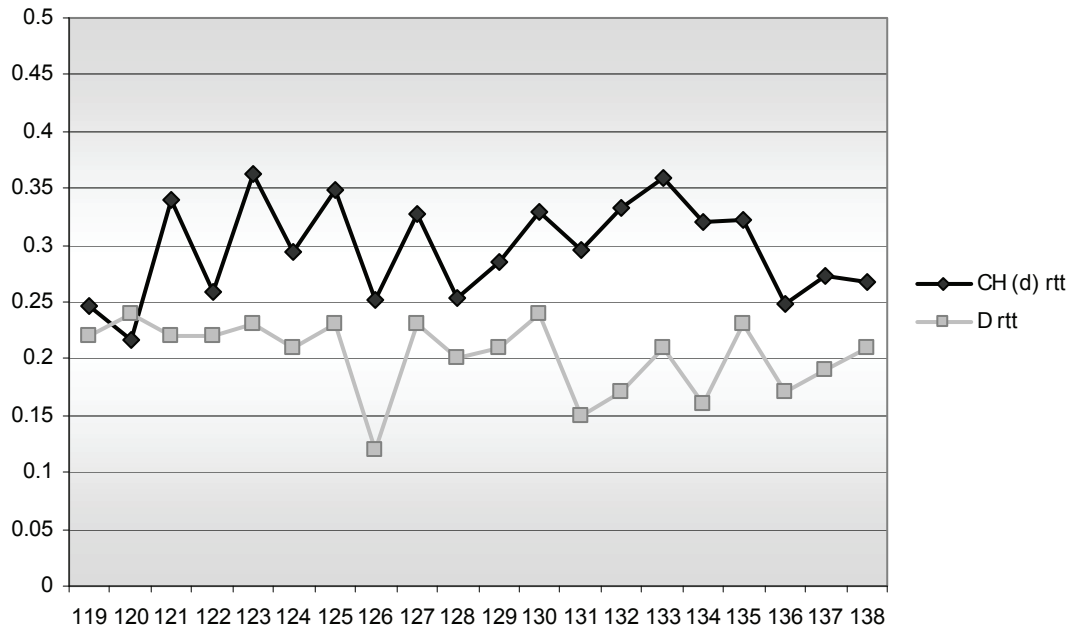


Abbildung 60: Trennschärpen für den Untertest „Fakten lernen“.

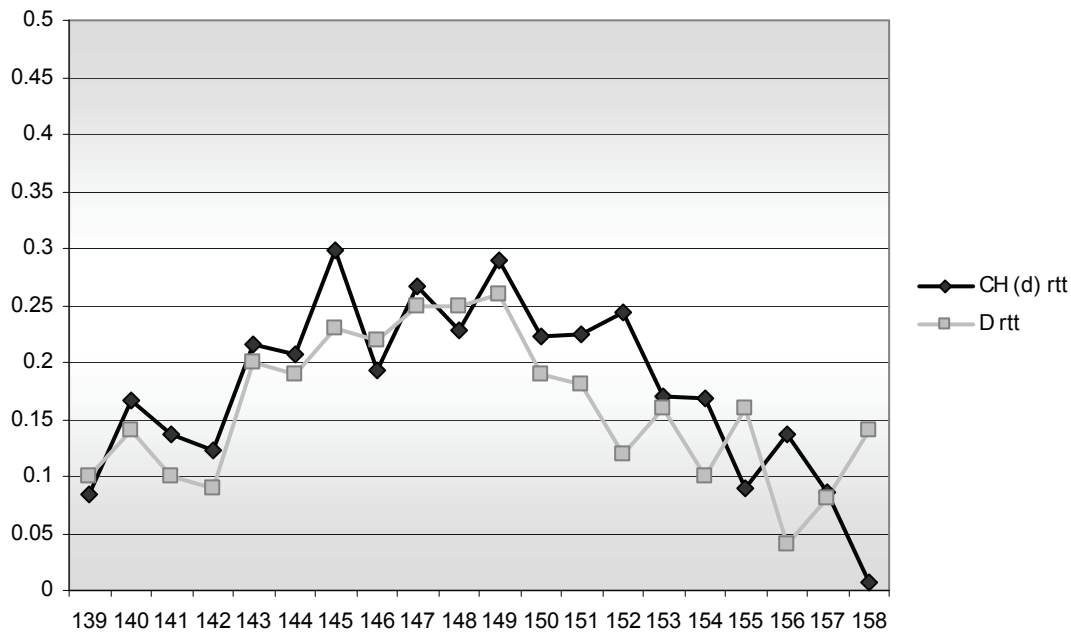


Abbildung 61: Trennschärpen für den Untertest „Muster zuordnen“.

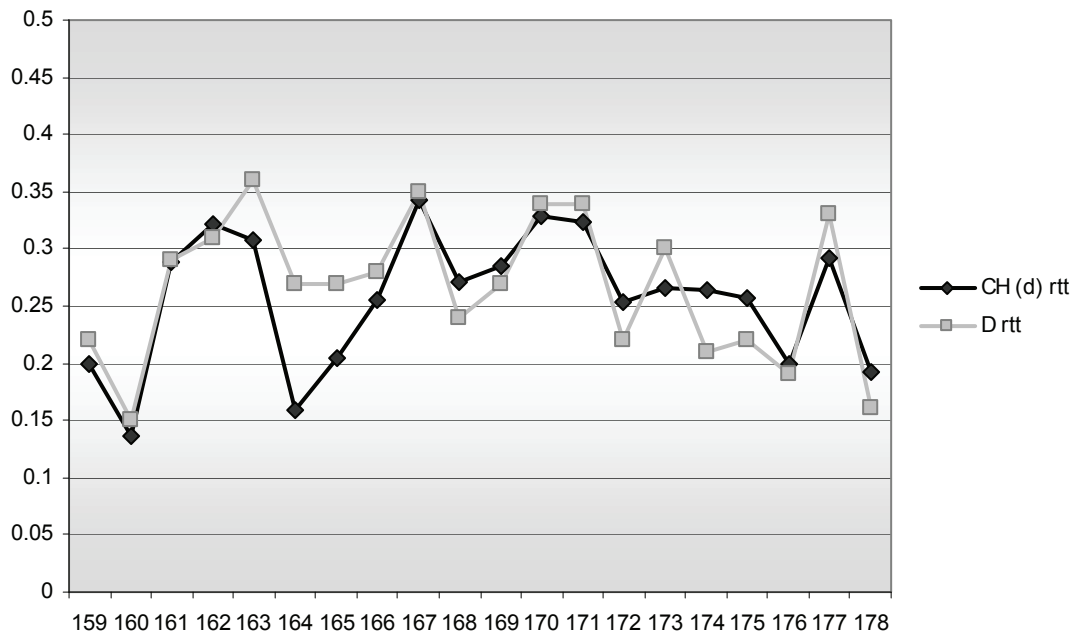


Abbildung 62: Trennschärfen für den Untertest „Diagramme und Tabellen“.

8.4 Konzentriertes und sorgfältiges Arbeiten

Dieser Untertest wurde 2004 revidiert und jährlich mit einer variierenden Anforderung (Zeichen sowie Durchstreichregel) durchgeführt, die vorher nicht bekannt ist. Es hatte sich früher gezeigt, dass durch exzessives Üben (40 bis teilweise 80mal) bei diesem Test ein so hoher Automatisierungsgrad erreicht wurde, dass nahezu alle Zeichen bearbeitet werden konnten (vgl. Bericht 9 des ZTD und Vorbereitungsreport 2003). Dieses Ergebnis ist nicht im Sinne der Messintention dieses Tests: Es soll kein „Fleisstest“ sein, sondern tatsächlich Konzentration und Sorgfalt gemessen werden.

Aus der Literatur ist bekannt, dass hochspezialisierte 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 deutlich aus. In den Jahren 2004 bis 2006 findet sich eine glockenförmige Verteilung des Punktwertes – im Unterschied zu stark rechtsschiefen Verteilungen in den Vorjahren.

Ab 2006 wurde auch in der Test-Info eine Form bereitgestellt, die formal der neuen Version entspricht. Die trotzdem weiter empfohlene Vorbereitung mit der veröffentlichten Originalversion soll dem Zweck dienen, Erfahrungen zu sammeln, wie man 8 Minuten zusammenhängend konzentriert arbeiten kann und wie man die Markierungen vornehmen muss, damit 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 über den gesamten Leistungsbereich gut. Die Sorgfalt beim Markieren (nicht in Nachbarzeichen markieren, deutlich markieren) gehört bei diesem Test mit zur Anforderung – entsprechende Regelverstöße werden als Fehler gewertet.

2007 bestand der Untertest aus zwei graphischen, sehr ähnlichen Zeichen (Kreise mit Öffnungen bei „6 Uhr“, beziehungsweise „7.30 Uhr“) – nennen wir sie „a“ und „b“. Es sollte jedes „b“ markiert werden, wenn direkt danach ein „a“ folgte.

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

Der Punktwert (PW) berechnet sich wie folgt:

$$\text{PW} = \text{Richtige} - \text{Fehler I} - \text{Fehler IIa} - \text{Fehler IIb}$$

Der Rohwert Richtige minus Fehler beträgt im Mittel 278.2 mit einer Standardabweichung von 67.34. Werden mehr Fehler als Richtige gemacht, entstehen negative Werte, die auf Null gesetzt werden.

Die Standardisierung zum Punktwert von 0 bis 20 erfolgt nach folgender Regel: Aus der Verteilung der Rohwerte erhalten die unteren 2.5% der Personen den Punktwert 0, die oberen 2.5% den Punktwert 20. Der Wertebereich für die mittleren 95% der Personen wird in 19 gleiche Abschnitte geteilt, sodass sich die nachfolgende Umrechnungstabelle ergibt. Dieses Verfahren gewährleistet, dass auch unterschiedlich schwierige Konzentrationstests zu einer vergleichbaren Bepunktung über die einzelnen Jahre führen und der differenzierte Wertebereich nicht durch Ausreisser nach oben und unten eingeschränkt wird. Die Leistung der mittleren 95% der Personen entspricht einer Normalverteilung.

Punktwert 0-20	Von Rohwert	Bis Rohwert
0	0	173
1	174	184
2	185	196
3	197	207
4	208	218
5	219	229
6	230	241
7	242	252
8	253	263
9	264	274
10	275	286
11	287	297
12	298	308
13	309	319
14	320	331
15	332	342
16	343	353
17	354	364
18	365	376
19	377	387
20	388	400

Tabelle 52: Umrechnung Rohwert in Punktwert „Konzentriertes und Sorgfältiges Arbeiten“ 2007

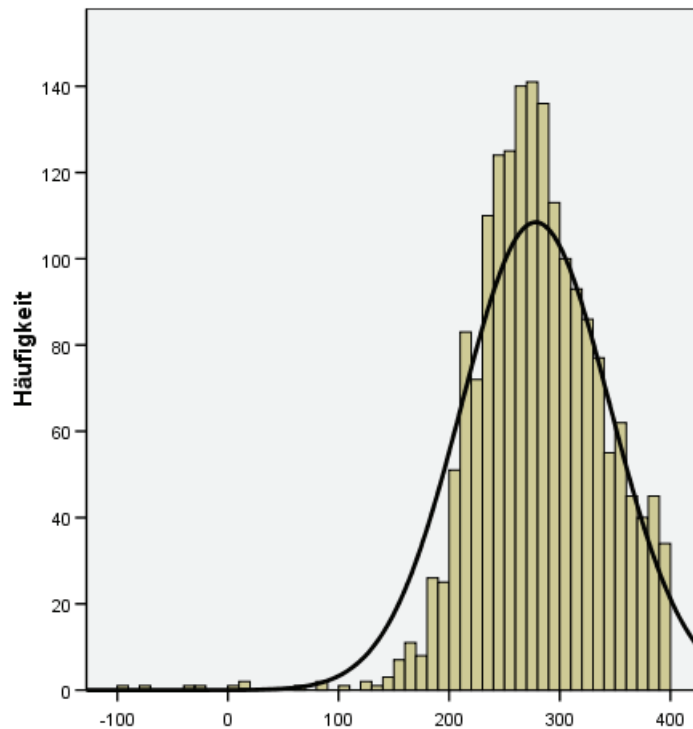


Abbildung 63: Verteilung des Rohwertes für „Konzentriertes und Sorgfältiges Arbeiten 2007“, negative Werte nicht Null gesetzt. 5 Ausreisser liegen unter -100 (nicht dargestellt).

		Mittelwert	Standardabweichung	Minimum	Maximum
Deutsch	Richtige	296.34	54.32	4	400
	Fehler	14.52	30.36	0	638
	... Auslassungen (I)	7.22	17.21	0	315
	... Fehlmarkierungen (II)	7.30	17.03	0	378
	letztes bearbeitetes Zeichen	1212.76	219.02	547	1599
Französisch	Richtige	259.76	46.11	157	399
	Fehler	11.52	16.13	0	151
	... Auslassungen (I)	5.30	7.06	0	37
	... Fehlmarkierungen (II)	6.22	11.24	0	114
	letztes bearbeitetes Zeichen	1058.15	183.38	631	1599
Italienisch	Richtige	265.22	52.68	5	399
	Fehler	17.24	56.72	0	506
	... Auslassungen (I)	7.87	28.74	0	253
	... Fehlmarkierungen (II)	9.37	28.29	0	253
	letztes bearbeitetes Zeichen	1090.73	178.21	675	1599
Gesamt	Richtige	292.64	54.78	4	400
	Fehler	14.44	31.26	0	638
	... Auslassungen (I)	7.12	17.40	0	315
	... Fehlmarkierungen (II)	7.32	17.36	0	378
	letztes bearbeitetes Zeichen	1197.53	219.68	547	1599

Tabelle 53: Sprachspezifische Statistiken für Parameter des Untertests „Konzentriertes und sorgfältiges Arbeiten“ und Gesamt.

	Fehler gesamt	Auslassungen (I)	Fehlmarkierungen (II)
Richtige	-.162(**)	-.155(**)	-.137(**)
Fehler gesamt		.900(**)	.899(**)
Auslassungen (I)	.900(**)		.618(**)

Tabelle 54: Korrelationen der Parameter im Konzentrationstest; **Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

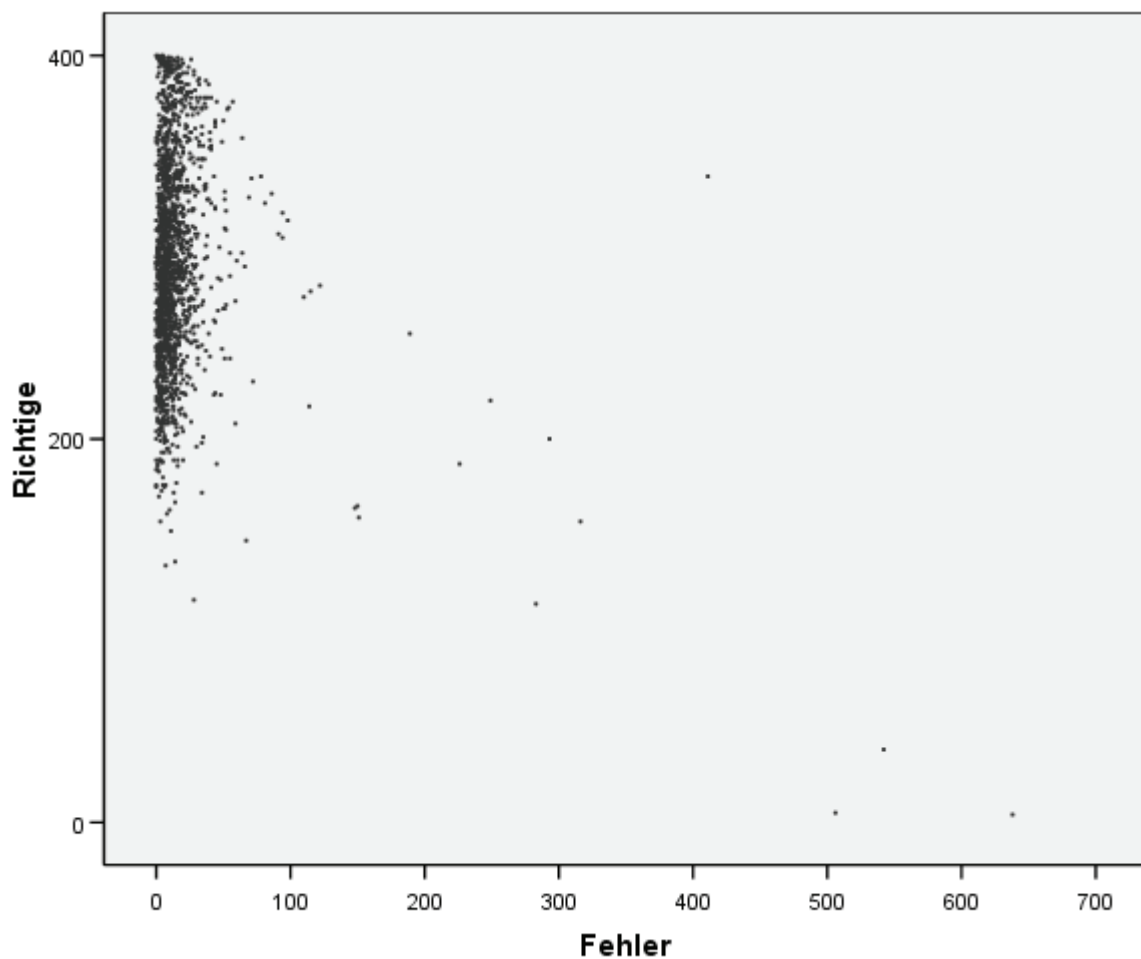


Abbildung 64: Scatterplot für alle Personen – Mengenleistung (Richtige) und Genauigkeit (niedrige Fehlerzahl).

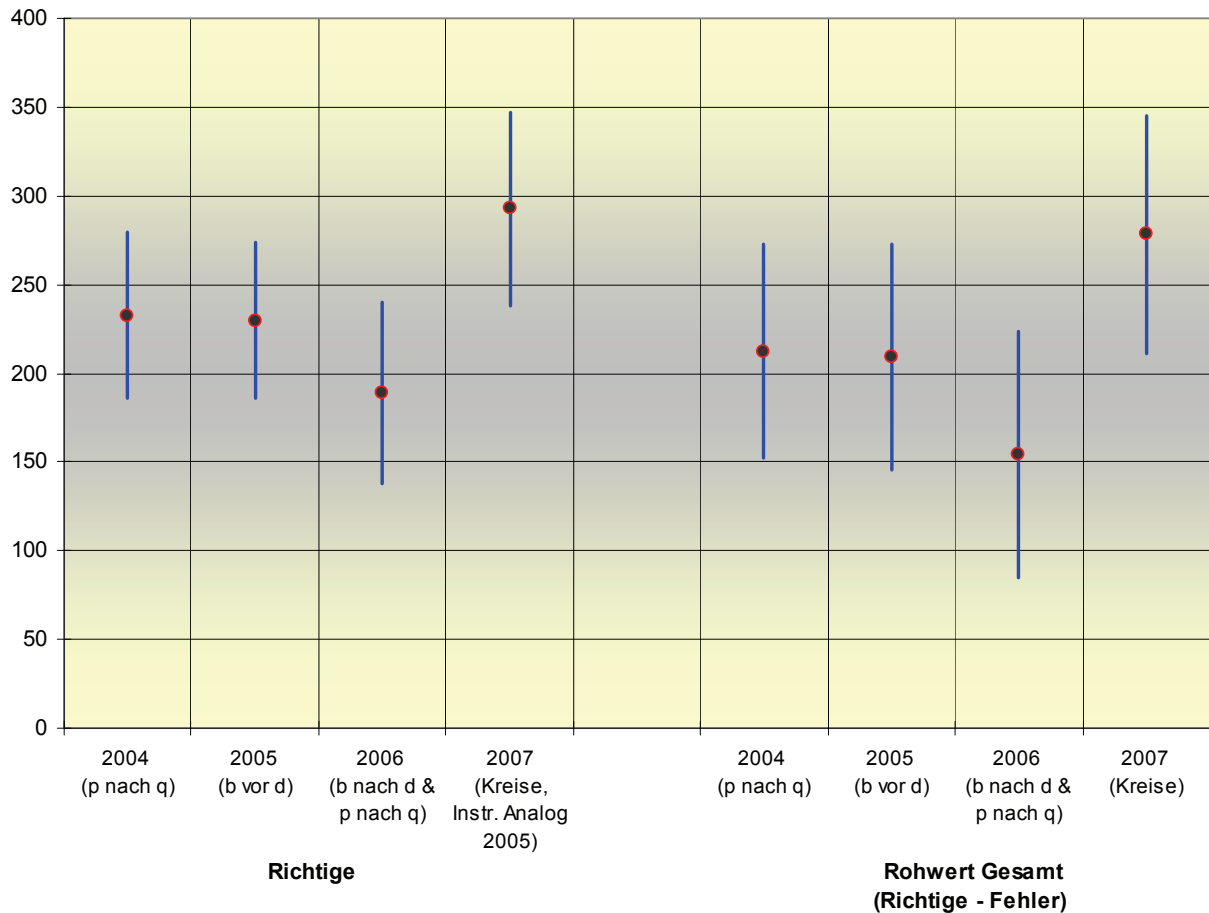


Abbildung 65: Leistungen „Konzentriertes und sorgfältiges Arbeiten“ seit Einführung variierender Instruktionen und Zeichen (2005)

Trotz der Ähnlichkeit der beiden verwendeten grafischen Zeichen handelte es sich bei der Ausgabe 2007 des Untertests „Konzentriertes und sorgfältiges Arbeiten“ um eine vergleichsweise leichte Variante. Die Streuung der Ergebnisse ist allerdings mit den Vorjahren vergleichbar und tendenziell sogar etwas grösser – damit ist die Differenzierung zwischen den Kandidaten gewährleistet. Auch ein Deckeneffekt war trotz der insgesamt höheren Punktzahlen nicht zu beobachten. Nur eine Person in der Schweiz und drei in Österreich haben den Maximalwert von 400 Rohwertpunkten erreicht, die Differenzierung ist also auch im obersten Leistungssegment erfüllt.

9 Teilnahmen am EMS nach Kantonen seit 1998

Seit 1998 haben bisher 11 740 Personen am EMS teilgenommen. Zu beachten ist, dass seit 1998 Humanmedizin, seit 1999 Veterinärmedizin und erst seit 2004 Zahnmedizin einem NC an den Universitäten Basel, Bern, Freiburg und Zürich unterliegen¹.

Wohnkanton/ Wohnort	Disziplin			Geschlecht		Sprache			Total
	Human- medizin	Veterinär- medizin	Zahn- medizin	männlich	weiblich	deutsch	franzö- sisch	italie- nisch	
AG	735	139	49	346	577	923	0	0	923
AI	20	6	2	9	19	28	0	0	28
AR	65	11	9	36	49	85	0	0	85
BE	1461	279	121	658	1203	1832	28	1	1861
BL	517	73	37	234	393	623	4	0	627
BS	475	38	22	203	332	534	0	1	535
FR	389	73	14	185	291	176	300	0	476
GE	7	76	0	18	65	6	77	0	83
GL	37	12	6	15	40	55	0	0	55
GR	310	68	22	145	255	384	0	16	400
JU	11	28	2	10	31	2	39	0	41
LU	548	94	26	228	440	667	1	0	668
NE	10	52	2	13	51	4	60	0	64
NW	50	13	2	20	45	65	0	0	65
OW	47	2	2	16	35	51	0	0	51
SG	638	108	46	271	521	792	0	0	792
SH	92	17	8	38	79	117	0	0	117
SO	331	62	34	169	258	426	1	0	427
SZ	170	27	10	62	145	206	0	1	207
TG	282	69	21	115	257	372	0	0	372
TI	308	72	25	170	235	7	1	397	405
UR	38	8	7	16	37	53	0	0	53
VD	31	165	0	58	138	11	185	0	196
VS	134	52	17	69	134	144	59	0	203
ZG	157	41	6	81	123	202	1	1	204
ZH	2146	318	119	914	1669	2561	8	14	2583
FL	46	5	6	17	40	57	0	0	57
Ausland	125	25	12	65	97	146	13	3	162
Total	9180	1933	627	4181	7559	10529	777	434	11740

Tabelle 55: Gesamtzahl der Teilnehmer am EMS 1998 bis 2007. Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

¹Aufgrund eines Beschlusses der Schweizerischen Universitätskonferenz von 2007 werden keine Testergebnisse mehr nach Kantonen aufgeschlüsselt veröffentlicht, weil dies vereinzelt für kantonale Bildungsvergleiche benutzt worden ist, obwohl die Repräsentativität für diesen Zweck nicht gegeben ist und darauf ausdrücklich hingewiesen wurde.

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: 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 BE sind dies täglich?

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

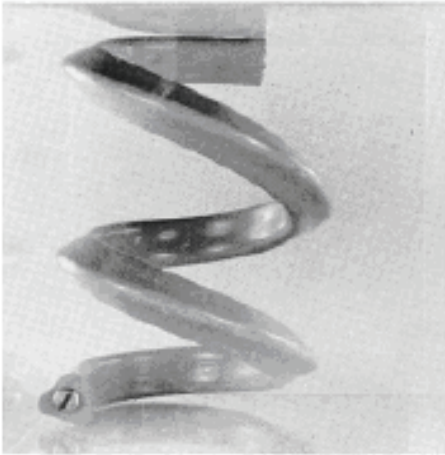
Bei solchen Fragen werden die Kenntnisse der Mittelstufen-Mathematik, nicht jedoch Lerninhalte vorausgesetzt. Der Patient soll ein Fünftel von 4800 kJ in Kohlehydraten zu sich nehmen, das sind also 960 kJ. Dividiert man diese Zahl durch 16, so erhält man die Anzahl g Kohlehydrate, nämlich 60 g, die es 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.

10.2 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.

9)



Hier sehen Sie den Würfel von vorne!



Hier sehen Sie den Würfel von ...?
(hinten!)

- (A): r
- (B): l
- (C): u
- (D): o
- (E): h

10.3 Untertest: Textverständnis

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

Ein Beispiel:

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

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

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

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

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

Eine hohe Konzentration von T_3 und T_4 im Blut hemmt die TSH- und die TRH-Sekretion, eine niedrige Konzentration stimuliert sie. Bei den an der Steuerung der Schilddrüsenhormon-Sekretion beteiligten Arealen von Hirnanhangsdrüse und Hypothalamus können krankheitsbedingte Störungen auftreten, die zu einer Über- oder Unterfunktion der Schilddrüse führen.

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

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

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

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

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

10.4 Untertest: Planen und Organisieren

Der Untertest „Planen und Organisieren“ ist seit 1995 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 „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.

10.5 Untertest: Konzentriertes und Sorgfältiges Arbeiten

Nachfolgend eine Beispielinstruktion aus der Test-Info:

Mit diesem Test soll Ihre Fähigkeit, rasch, sorgfältig und konzentriert zu arbeiten erfasst werden. Sie sehen nachfolgend ein Blatt mit 40 Zeilen, die aus je 40 Buchstaben u und m gebildet werden.

**Ihre Aufgabe ist es, zeilenweise jedes u zu markieren,
VOR dem in der Zeile unmittelbar ein m steht:**

m ✗

Sie dürfen kein u markieren, vor dem kein m steht. Sie dürfen natürlich auch kein m markieren. Beides wären Fehler. Nachfolgend sehen Sie ein richtig bearbeitetes Beispiel:

m ✗ m ✗ m ✗ u u m m ✗ m ✗ u u u m ✗ m m

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.

10.6 Untertest: Medizinisch-naturwissenschaftliches Grundverständnis

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

Im Kindesalter kann das Zentrum für Sprache, Spracherwerb und Sprachverständnis noch in der linken oder in der rechten Hälfte (Hemisphäre) des Gehirns in einem umschriebenen Hirnrindengebiet

(sog. Sprachregion) angelegt werden. Spätestens im zwölften Lebensjahr sind die sprachlichen Fähigkeiten jedoch fest in einer der beiden Hemisphären verankert, und zwar bei den Rechtshändern in der Regel links, bei den Linkshändern in der Mehrzahl ebenfalls links, zum Teil aber auch rechts; die korrespondierende Region der Gegenseite hat zu diesem Zeitpunkt bereits andere Funktionen fest übernommen. Welche der nachfolgenden Aussagen lässt bzw. lassen sich aus diesen Informationen ableiten?

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

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

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

- I. Da bei der Mehrzahl der Linkshänder die Sprachregion in der linken Hemisphäre liegt, müssen sie also mit einer Sprachstörung rechnen, weshalb Aussage I falsch ist.
- II. Da es im Kindesalter noch offen ist, in welcher Hälfte des Gehirns die Sprachregion angelegt wird, besteht für ein Vorschulkind immer noch die Möglichkeit, die Muttersprache wieder zu erlernen. Die Sprachregion wird dann in der rechten Hälfte der Hemisphäre angelegt. Somit ist Aussage II ebenfalls falsch.
- III. Da spätestens im zwölften Lebensjahr die Sprachregion bei Rechtshändern in der Regel fest in der linken Hälfte des Gehirns liegt, ist bei zwanzigjährigen Rechtshändern zu erwarten, dass sie die Fähigkeit, eine Sprache zu erlernen, verloren haben. Die Aussage III ist darum richtig.

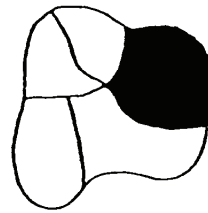
Nach dieser Analyse des Textes ist es offensichtlich, dass die Antwort (C) richtig ist.

10.7 Untertest: Figuren lernen

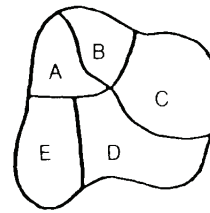
Für beide nachfolgenden Gedächtnistests wird nach der Mittagspause das Material zum Einprägen ausgeteilt. Vor der Abfrage des Gelernten wird der Untertest „Textverständnis“ bearbeitet, damit liegt die Zeit des Behaltens der gelernten Inhalte über einer Stunde. Gedächtnisleistungen sind wichtige Voraussetzungen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

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

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 angegeben 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 ...

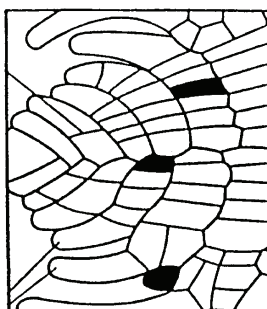
- (A) Installateur
- (B) Lehrer
- (C) Dachdecker
- (D) Handelsvertreter
- (E) Physiker

10.9 Untertest: Muster zuordnen

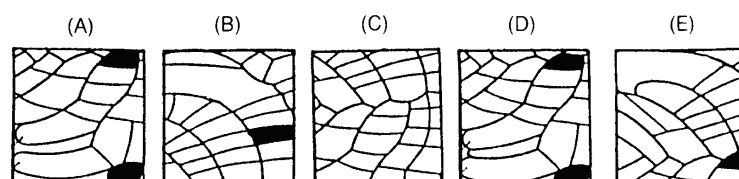
In diesem Untertest wird die Fähigkeit geprüft, Ausschnitte in einem komplexen Bild wieder zu erkennen. Dazu werden pro Aufgabe ein Muster und je fünf Musterausschnitte (A) bis (E) vorgegeben. Die Testteilnehmerin oder der Testteilnehmer soll herausfinden, welcher dieser fünf Musterausschnitte an irgendeiner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

10.10 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 die richtige 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: 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

Bericht 12 (2006)
EMS Eignungstest für das Medizinstudium in der Schweiz 2006 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher
Bericht 11 (2005)
EMS Eignungstest für das Medizinstudium in der Schweiz 2005 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> <i>Inkl. einer Befragung über die Studienmotivation; K.-D. Hänsgen und B. Spicher</i>
Bericht 10 (2004)
EMS Eignungstest für das Medizinstudium in der Schweiz 2004 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher
Bericht 9 (2003)
EMS Eignungstest für das Medizinstudium in der Schweiz 2003 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher
Bericht 8 (2002)
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; K.-D. Hänsgen und B. Spicher</i>
Bericht 7 (2001)
EMS Eignungstest für das Medizinstudium in der Schweiz 2001 <i>Bericht über die Durchführung und Ergebnisse, Vorhersage von Prüfungserfolg</i> <i>Zusammenfassender Vergleich 1998 bis 2001 - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher
Bericht 6 (2000)
EMS Eignungstest für das Medizinstudium in der Schweiz 2000 <i>Bericht über die Durchführung und Ergebnisse, Evaluation der Vorhersage von</i> <i>Prüfungserfolg. Zusammenfassender Vergleich 1998 bis 2000 - avec un résumé en</i> <i>français K.-D. Hänsgen und B. Spicher</i>
Bericht 5 (1999)
EMS Eignungstest für das Medizinstudium in der Schweiz 1999 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher
Bericht 4 (1998)
EMS Eignungstest für das Medizinstudium in der Schweiz 1998 <i>Bericht über die Durchführung und Ergebnisse - avec un résumé en français</i> K.-D. Hänsgen und B. Spicher

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>

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