

EMS 2017 – Bericht 24



Eignungstest für das Medizinstudium in der Schweiz

Bericht über Durchführung und Ergebnisse

B. Spicher und K.-D. Hänsgen

Spicher, Benjamin und Hänsgen, Klaus-Dieter (2017)
EMS Eignungstest für das Medizinstudium 2017
Berichte des Zentrums für Testentwicklung, Nr. 24,
im Auftrag der Schweizerischen Hochschulkonferenz (SHK)

Unter Mitarbeit von M. Bernasconi,
T. Cruchoad, Y. de Zordo und M. Strazzeri

In Zusammenarbeit mit der Rektorenkonferenz der Schweizer
Hochschulen (swissuniversities), die das Anmelde- und
Zuteilungsverfahren der Plätze zum Medizinstudium
durchführt und für diesen Bericht statistische Angaben zur
Verfügung stellt:

J. Bregy, Chr. Winzenried, M. Galamic

Redaktion: Yannick de Zordo



Siehe auch www.ztd.ch

Inhaltsverzeichnis

1	Zusammenfassung	7
2	Résumé	8
3	Numerus clausus (NC) und Medizinstudium	9
4	Anmeldung zum Medizinstudium und Test 2017	11
4.1	Anmeldestatistik Humanmedizin.....	11
4.2	Anmeldestatistik Veterinärmedizin	13
4.3	Anmeldestatistik Zahnmedizin	15
4.4	Grösse der Testlokale	18
4.5	Testorte und Wunschhochschulen	19
4.6	Wunschhochschule und Testort nach Wohnkanton.....	20
4.7	Teilnahmen am EMS (NC) nach Kantonen seit 1998	24
4.8	Testabsolvierung nach Geschlecht	25
4.9	Übernahme des Testergebnisses aus dem Vorjahr	27
4.10	Sprachgruppen.....	28
5	Beschreibung des verwendeten Eignungstests	29
5.1	Aufbau des Tests.....	29
5.2	Berechnung der Werte 2017	31
5.3	Mittlerer Rangplatz der Aufgabengruppen	32
6	Testanwendung in der Schweiz 2017	34
6.1	Verteilungsprüfung	34
6.2	Äquivalenz der Sprachversionen	35
6.2.1	<i>Sprachvergleich für die Aufgabengruppen</i>	<i>37</i>
6.2.2	<i>Darstellung des Korrekturverfahrens.....</i>	<i>39</i>
6.2.3	<i>Identifikation von DIF-Aufgaben und Bestimmung des Korrekturwertes</i>	<i>40</i>
6.2.4	<i>Effekte der Korrektur</i>	<i>48</i>
6.3	Vergleichbarkeit der Testlokale	51
6.4	Vergleich für die Geschlechter.....	53
6.5	Vergleiche nach Wunschhochschulen	55

7	Ergebnisse zur Testgüte	56
7.1	Zuverlässigkeit	56
7.2	Binnenstruktur	57
7.3	Konzentriertes und sorgfältiges Arbeiten	59
8	Beispielaufgaben für die Aufgabengruppen	62
8.1	Muster zuordnen	62
8.2	Medizinisch-naturwissenschaftliches Grundverständnis	62
8.3	Schlauchfiguren	64
8.4	Quantitative und formale Probleme	64
8.5	Konzentriertes und sorgfältiges Arbeiten	65
8.6	Textverständnis	65
8.7	Figuren lernen	67
8.8	Fakten lernen	67
8.9	Diagramme und Tabellen	68
9	Literatur	69
9.1	Originaltest zur Information und Vorbereitung	70
9.2	Frühere Berichte des ZTD	70

1 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2017 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Medizin ab Herbstsemester 2017. Für bestimmte Fragestellungen werden Vergleichsdaten der Testanwendungen 1998 bis 2016 herangezogen.

1. Ein Numerus clausus (NC) ist 2017 für Humanmedizin (seit 1998), Veterinärmedizin (seit 1999) und Zahnmedizin (seit 2004) notwendig für Personen, die sich an den Hochschulen in Basel, Bern, Freiburg oder Zürich angemeldet haben. In Humanmedizin auch für die neuen Studiengänge an der ETH Zürich, für die Luzerner und St. Galler Tracks der Universität Zürich sowie für den Studiengang an der USI im Tessin, der in Basel begonnen wird. An der Universität Zürich gilt der NC weiter auch für Chiropraktik (seit 2008).
2. Die gegenüber 2016 um weitere 25% gesteigerten Kapazitäten für **Humanmedizin** (inklusive Chiropraktik) wurden mit den Anmeldungen im **Februar 2017** zu **412%** (Vorjahr 473%) ausgelastet. Der Steigerung der Nachfrage um 355 Personen standen 204 neue Plätze gegenüber. In **Veterinärmedizin** wurden, bezogen auf die Anmeldungen, die Kapazitäten zu **377%** (Vorjahr 331%), in **Zahnmedizin** zu **250%** (Vorjahr 259%) ausgelastet.
3. **3930 Personen** (Vorjahr 3694) haben sich für diese Hochschulen bis zur gesetzten Frist im **Mai 2017** bei swissuniversities zum EMS-Test 2017 angemeldet. Davon bewarben sich **16 Personen** um einen Disziplinenwechsel oder Einstieg in einem höheren Semester. Weitere **110 Personen** bewarben sich mit ihrem Testergebnis aus dem Vorjahr und nahmen nicht am diesjährigen Test teil.
4. **3767 Personen** haben den EMS mit gültigem Ergebnis beendet (Vorjahr 3523). 6 Personen haben den Test vorzeitig abgebrochen (wie im Vorjahr). 163 Personen haben ihre Testanmeldung zurückgezogen bzw. sind nicht zum Test erschienen (Vorjahr 165).
5. Der Test fand am 7. Juli 2017 gleichzeitig **an 9 Testorten** und in 34 Testlokalen in drei Sprachen (Deutsch: Aarau-Suhr, Basel, Bern, Chur, Luzern, St. Gallen, Zürich; Französisch: Freiburg; Italienisch: Lugano) statt. Die Wahl des Testortes war unabhängig vom zukünftigen Studienort möglich.

2 Résumé

Le présent rapport informe sur les résultats concernant le test d'aptitudes aux études de médecine en Suisse (AMS) et les admissions aux études en question au semestre d'automne 2017. Pour certaines problématiques, des données comparatives des éditions antérieures du test (1998-2016) y figurent également.

1. En 2017, un **numerus clausus (NC)** est à nouveau nécessaire en médecine humaine (depuis 1998), en médecine vétérinaire (depuis 1999) et en médecine dentaire (depuis 2004) pour les personnes inscrites aux hautes écoles de Bâle, Berne, Fribourg ou Zurich. De nouvelles filières en médecine humaine avec NC se sont ajoutées en 2017 à savoir, à l'EPFZ, aux Universités de Lucerne et de St-Gall en coopération avec l'Université de Zurich ainsi qu'à l'Università della Svizzera italiana dont les études commenceront à Bâle. À l'Université de Zurich, le NC se poursuit pour la chiropractie (depuis 2008).
2. Par rapport à 2016, la capacité des admissions en **médecine humaine** (y compris la chiropractie) a été augmentée de 25% mais elle continue d'être dépassée à hauteur de **412%** (473% l'année précédente) suite aux inscriptions reçues en **février 2017**. L'augmentation de la demande de l'ordre de 355 personnes est compensée en partie par 204 places nouvellement mises à disposition. **En médecine vétérinaire**, le dépassement de la capacité d'accueil s'élève à **377%** pour 2017 (331% en 2016). **En médecine dentaire**, il se monte à **250%** (259% en 2016).
3. **3930 personnes** (3694 en 2016) se sont annoncées pour ces hautes écoles auprès de swissuniversities jusqu'au délai fixé en **mai 2017**, en vue de la passation du test AMS 2017. Parmi elles, **16 personnes** ont sollicité un changement de discipline ou un passage à un semestre supérieur et **110 personnes** ont fait valoir leur résultat de l'année précédente et n'ont ainsi pas pris part au test 2017.
4. **3767 personnes** ont terminé l'AMS avec des résultats valables (3523 en 2016). 6 personnes ont interrompu prématurément le test (idem en 2016). 163 personnes ont retiré leur inscription au test ou ne se sont pas présentées à celui-ci (165 en 2016).
5. Le test s'est déroulé le 7 juillet 2017, simultanément **sur 9 sites** différents comprenant au total 34 locaux de test et ce, dans trois langues (allemand : Aarau-Suhr, Bâle, Berne, Coire, Lucerne, St-Gall, Zurich / français : Fribourg / italien : Lugano). La possibilité était donnée de choisir le lieu du test indépendamment du futur lieu d'études.

3 Numerus clausus (NC) und Medizinstudium

Erneut ist die Nachfrage nach einem Humanmedizinstudium an den Hochschulen mit NC sehr stark angewachsen. Den verglichen mit dem Vorjahr 270 zusätzlichen Bewerbungen im Juli stehen 204 neu geschaffene Plätze gegenüber. Die 2017 über die üblichen schrittweisen Erhöhungen hinaus geschaffenen Studienplätze haben eine leichte prozentuale Entlastung hinsichtlich der Zulassungsquoten bewirkt, das Verhältnis der Anmeldung zu vorhandenen Studienplätzen erreicht ungefähr das Niveau von 2011.

Gegenüber dem Stand der Februar-Anmeldungen in Humanmedizin stand 2017 für knapp jede vierte Person ein Studienplatz zu Verfügung, nach dem Test betrug die Quote (inklusive Überbuchungen) 34%. Im Fach Humanmedizin steigt die Nachfrage bei den Hochschulen mit NC weiterhin besonders stark. Bei Zahnmedizin wurde einiges dafür getan, die Spezifik des Studiums hinsichtlich der Berufs- wie Studienanforderungen bekannter zu machen¹, einschliesslich der Verdeutlichung einer Nichteignung als „Sprungbrett“ in ein Humanmedizinstudium.

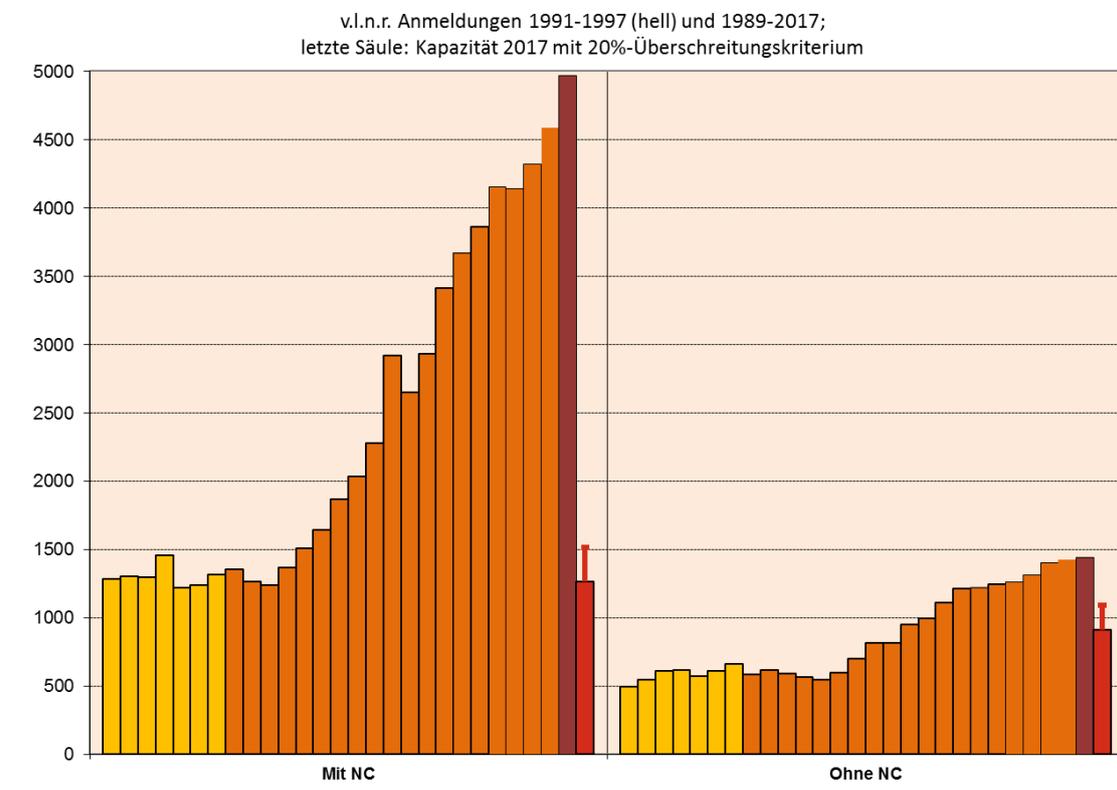


Abbildung 1: Anmeldungen zum Medizinstudium 1991 bis 2017 nach Gruppen (mit NC vs. ohne NC). Hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule: Kapazität 2017 mit 20%-Überschreitungskriterium als Grenze für Auslösung des NC.

¹ Siehe dazu: <http://unifr.ch/ztd/self-assessment/studienberatung.html>

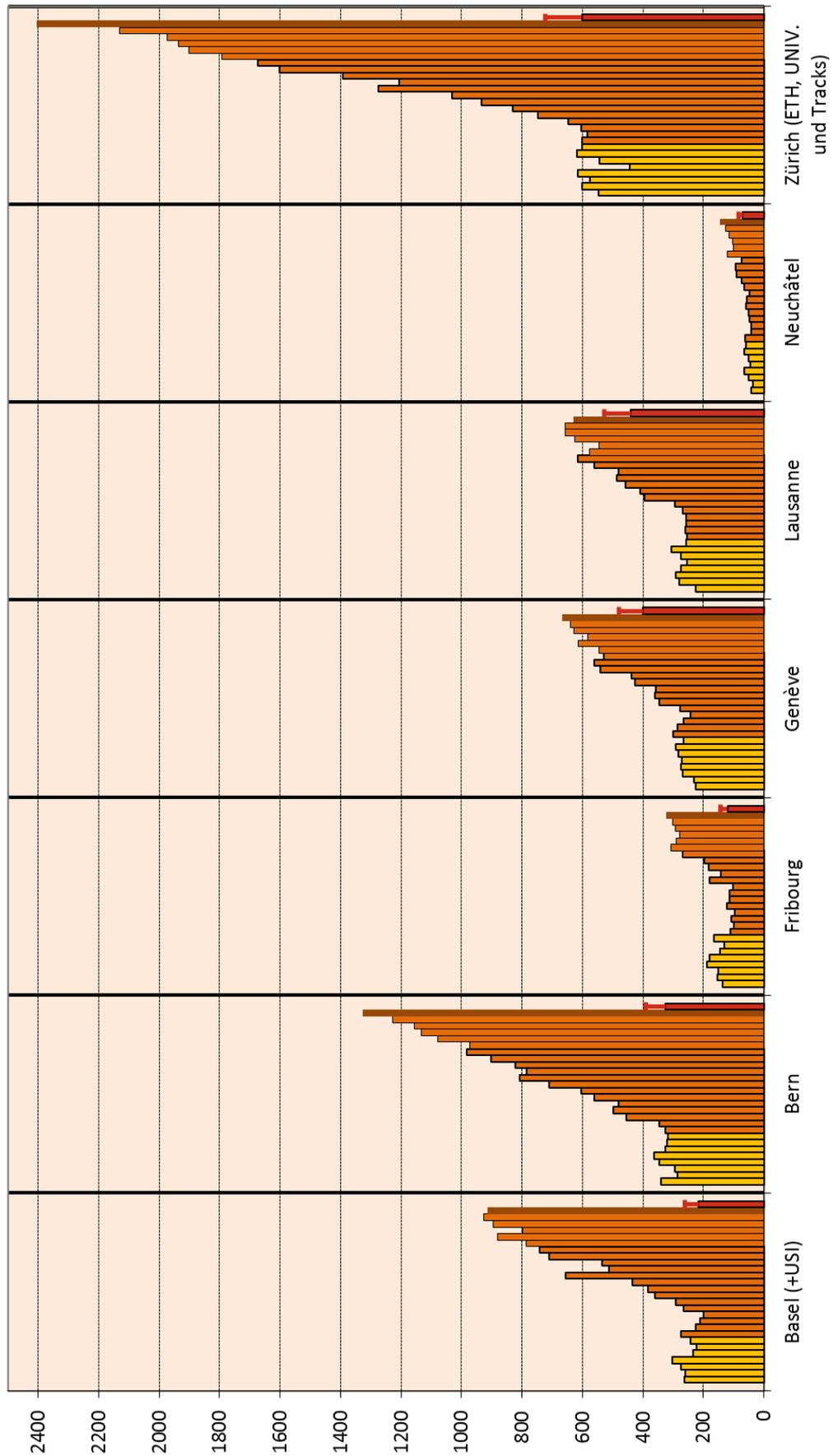


Abbildung 2: Anmeldezahlen pro Hochschule 1991 bis 2017 – hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule: Kapazität 2017 mit 20%-Überschreitungskriterium.

4 Anmeldung zum Medizinstudium und Test 2017

4.1 Anmeldestatistik Humanmedizin

Tabelle 1: Disziplinspezifische Statistiken für die am NC beteiligten Hochschulen für die Disziplin Humanmedizin.

	Humanmedizin (seit 2008 incl. Chiropraktik)												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Kapazitäten (HS mit NC)	546	546	583	603	603	653	653	653	713	793	793	793	997
Anmeldungen im Februar	1525	1712	2171	2029	2324	2651	2936	3120	3270	3310	3491	3754	4109
... in % zu Kapazität	279	314	372	336	385	406	450	478	459	417	440	473	410
Anmeldungen zum EMS Mai	1182	1356	1452	1612	1765	2109	2295	2461	2547	2653	2771	3000	3250
Absolvierung EMS	1107	1263	1374	1535	1664	2016	2172	2337	2403	2525	2659	2864	3122
Übernahme Ergebnis Vorjahr	41	45	41	52	66	54	65	54	60	42	59	70	76
Bewerbungen Juli	1143	1302	1405	1576	1707	2040	2212	2363	2426	2533	2684	2916	3186
Rückzugsquote zwischen Februar und Juli (%)	25.0	23.9	34.8	22.3	26.6	23.0	24.6	24.3	25.8	23.5	23.1	22.3	22.5
Zugeteilte Studienplätze mit „Überbuchung“	640	653	685	726	707	763	745	745	805	886	877	881	1091^m
Abgewiesene Bewerbungen	503	649	720	850	1000	1277	1467	1618	1621	1647	1807	2035	2095^m
% Bewerbungen, die Studienplatz erhalten	56	50	49	45	41	37	34	32	33	35	33	30	34^m

^m Stand Oktober gemäss Modell.

„Modell“ bedeutet, dass etwas 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 früh genau auszulasten (auch weil das Semester früh beginnt) so dass weitere Nachrücker möglichst vermieden werden können. Der Anteil der Rückzüge zwischen Februar (Anmeldung Studium) und der Testteilnahme bleibt in etwa stabil.

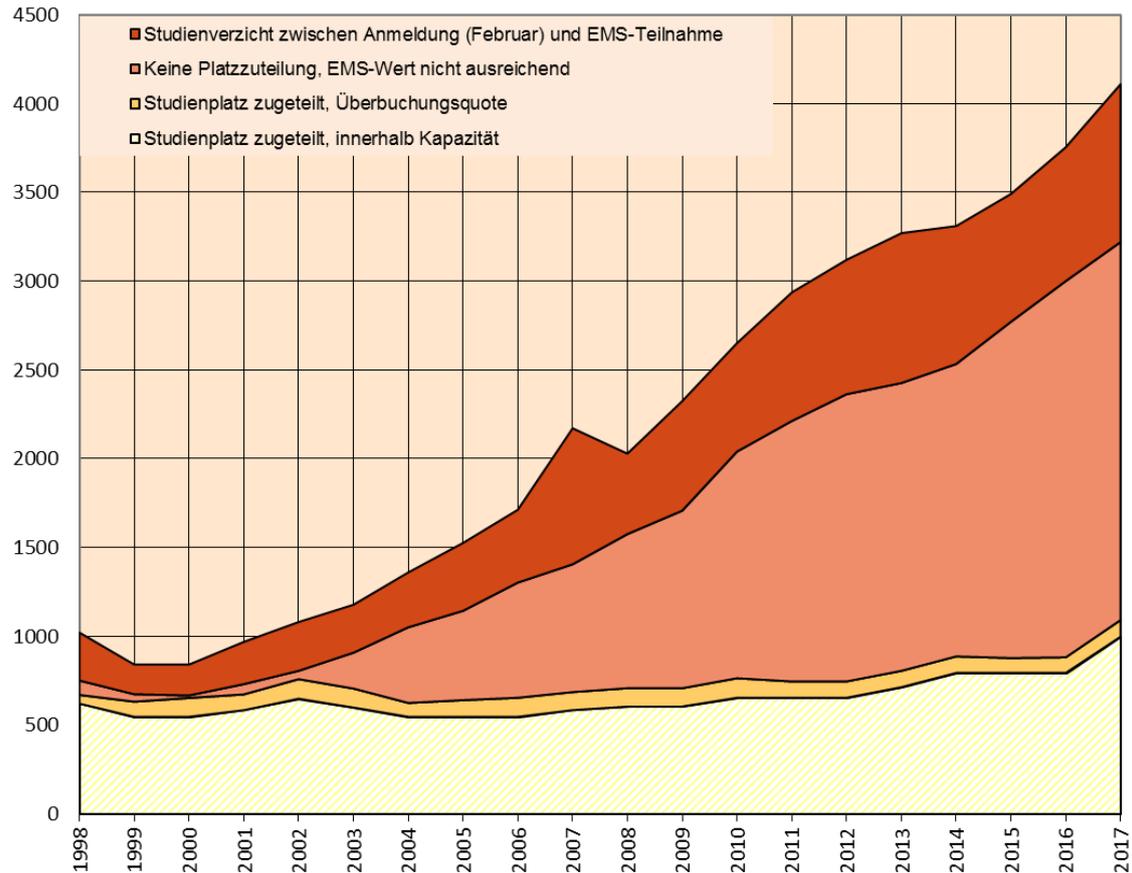


Abbildung 3: Anmeldeverlauf Humanmedizin. Seit 2002 nehmen die Anmeldezahlen und die Zahl der Testteilnahmen kontinuierlich zu. Die Zunahme der Anmeldezahlen 2007 war auf ein vereinfachtes, nur elektronisches Anmeldeverfahren zurückzuführen. Danach musste die Anmeldung zusätzlich wieder schriftlich erfolgen.

Der Kapazitätsausbau wird durch die Nachfragersteigerung nivelliert, letztere nimmt ungefähr um das Dreifache des Anstiegs der Studienplätze im ersten Jahr zu. Dennoch ist der Numerus clausus im Verhältnis zu anderen Ländern vergleichsweise mild: In Deutschland sind es 5 bzw. 12 (Winter- bzw. Sommersemester), in Österreich 6 Bewerbungen auf einen Studienplatz. Alle europäischen Länder beschränken den Zugang zum Medizinstudium. Für Schweizer ist es im Vergleich zu vielen anderen Ländern nach wie vor wahrscheinlicher, zum Medizinstudium zugelassen zu werden.

4.2 Anmeldestatistik Veterinärmedizin

Tabelle 2: Disziplinspezifische Statistiken für Veterinärmedizin und die hier am NC beteiligten Universitäten Bern und Zürich; in den anderen Hochschulen wird diese Studienrichtung nicht angeboten.

	Veterinärmedizin												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Kapazitäten (HS mit NC)	150	150	150	150	150	150	150	150	150	150	150	150	150
Anmeldungen im Februar	318	348	411	365	371	432	408	406	478	487	476	496	566
... in % zu Kapazität	212	232	274	243	247	288	272	271	319	325	317	331	377
Anmeldungen zum EMS Mai	257	279	301	290	294	360	308	327	365	403	404	427	459
Absolvierung EMS	244	268	288	272	281	342	295	317	350	384	375	408	440
Übernahme Ergebnis Vorjahr	8	10	7	15	16	11	16	8	24	17	8	8	16
Bewerbungen Juli	252	278	295	287	297	353	311	325	374	401	383	416	454
Rückzugsquote zwischen Februar und Juli (%)	20.8	20.1	28.2	21.3	19.9	18.0	23.8	20.0	21.8	17.7	19.5	16.1	19.8
Zugeteilte Studienplätze mit „Überbuchung“	180	170	173	173	173	173	173	173	173	173	173	167	167^m
Abgewiesene Bewerbungen	72	108	122	114	124	180	138	152	201	228	210	249	287^m
% Bewerbungen, die Studienplatz erhalten	71	61	59	60	58	49	56	53	46	43	45	40	37^m

^m Stand Oktober gemäss Modell.

Im Fach Veterinärmedizin nahm die Zahl der Anmeldungen per Februar 2017 deutlich zu, was sich auch in einer entsprechend erhöhten Zahl der Bewerbungen im Juli niederschlug. Die Rückzugsquote lag im Bereich der Vorjahre – die Chance auf Zuteilung eines Platzes für Personen mit absolviertem Eignungstest sank erstmals auf unter 40%.

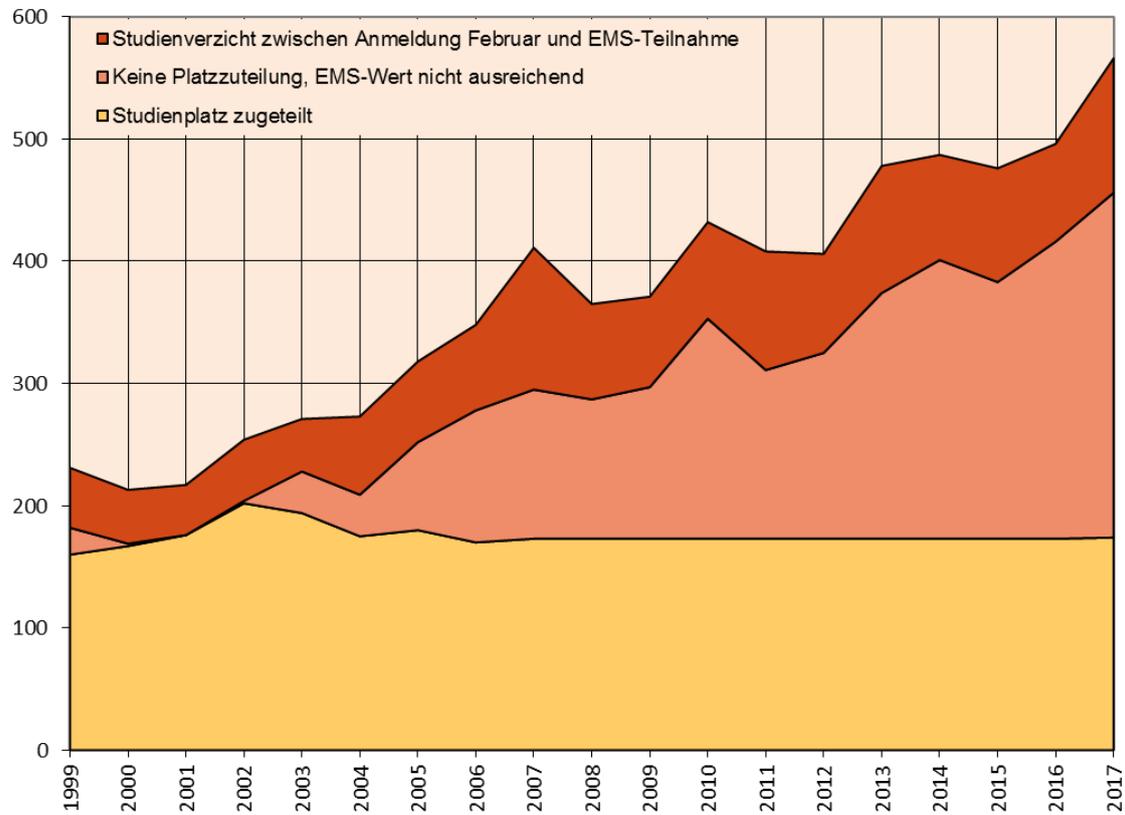


Abbildung 4: Anmeldeverlauf für Veterinärmedizin 1999 bis 2017 im Vergleich (Univ. mit NC).

4.3 Anmeldestatistik Zahnmedizin

Tabelle 3: Disziplinspezifische Statistiken für die am NC beteiligten Universitäten Basel, Bern, Freiburg und Zürich für Zahnmedizin.

	Zahnmedizin												
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Kapazitäten (Univ. mit NC)	142	142	142	142	142	142	142	139	139	131	131	131	117
Anmeldungen im Februar	190	217	267	197	239	315	324	332	396	343	352	339	292
... in % zu Kapazität	134	153	188	138	168	222	228	239	285	262	269	259	250
Anmeldungen zum EMS Mai	159	168	179	164	181	233	250	257	284	273	275	267	221
Absolvierung EMS	150	157	169	154	174	220	239	248	275	264	258	251	205
Übernahme Ergebnis Vorjahr	11	20	32	13	31	37	37	25	39	26	19	18	18
Bewerbungen Juli	161	174	201	167	205	257	276	273	313	290	277	269	223
Rückzugsquote zwischen Februar und Juli (%)	15.2	19.3	24.7	15.2	14.3	18.4	14.8	17.8	21.0	15.5	21.3	20.6	23.6
Zugeteilte Studienplätze mit „Überbuchung“	161	172	165	167	170	165	159	159	161	152	150 ^m	150^m	127^m
Abgewiesene Bewerbungen	0	2	36	0	35	92	117	114	152	138	127 ^m	119^m	96^m
% Bewerbungen, die Studienplatz erhalten	100	99	82	99	83	64	58	58	51	52	54 ^m	56^m	57^m

^m Stand Oktober gemäss Modell.

Die Kapazitäten für die Disziplin Zahnmedizin wurden 2017 von 131 auf 117 Plätze reduziert, da Freiburg sich auf Humanmedizin fokussiert und keine Plätze in Zahnmedizin (erstes Jahr) mehr bereitstellt. Dies ordnet sich in Bestrebungen ein, das Zahnmedizinstudium vom ersten Jahr an spezifischer zu gestalten und von Humanmedizin mehr abzugrenzen, was aufgrund der geringen Zahl und fehlenden spezifischen Ressourcen in Freiburg nur schwer möglich gewesen wäre. Wie bereits im Jahr 2016 nahm die Anzahl der Anmeldungen per Februar erneut ab, ebenso wie die Zahl der Testabsolventen. Die Zulassungsquote liegt mit 57% im Bereich des Vorjahres und bleibt seit 2014 leicht ansteigend. Aufgrund der nicht mehr so einfachen Wechselmöglichkeit zu Humanmedizin (bisher gleiche Studienanforderungen, aber wahrscheinlichere Zulassung) wäre dies erwünscht, wenn

sich nur noch Personen für Zahnmedizin bewerben, die später auch auf diesem Gebiet arbeiten möchten.

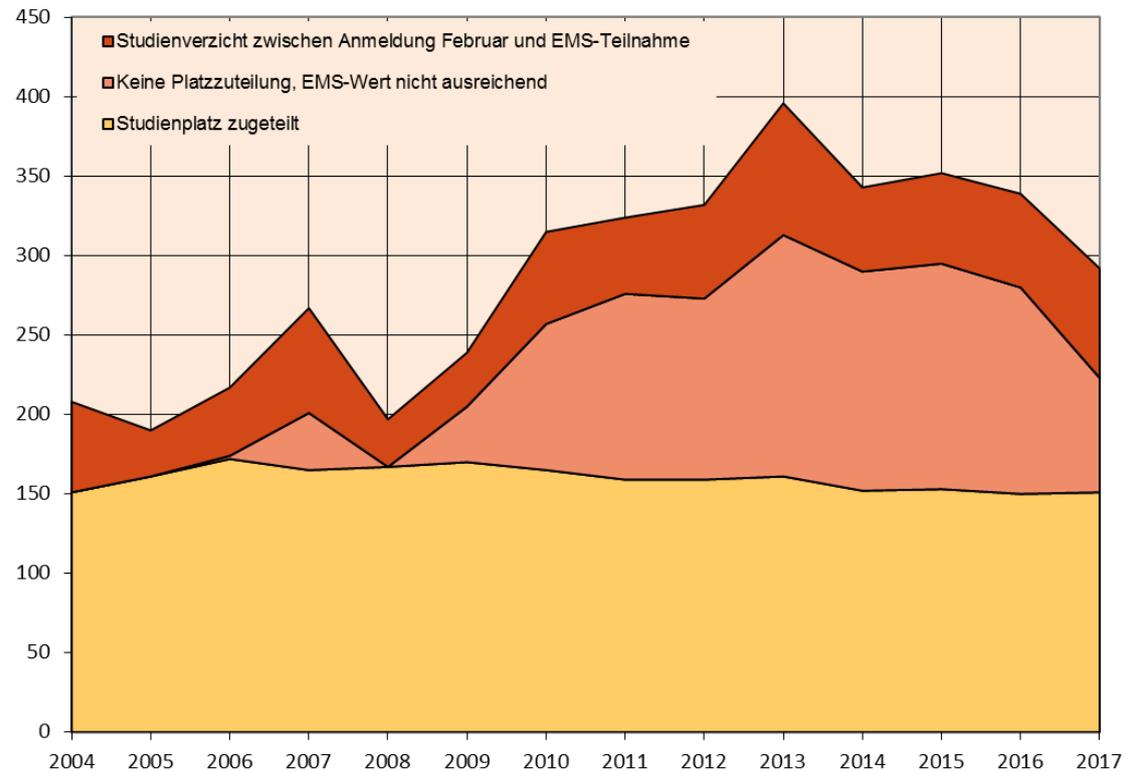


Abbildung 5: Anmeldeverlauf für Zahnmedizin 2004 bis 2017 im Vergleich (Univ. mit NC).

Durch die Eliminierung der Möglichkeit eines quasi „freien“ Disziplinenwechsels in höheren Semestern von Zahn- zu Humanmedizin sollte erreicht werden, dass sich nur Personen für Zahnmedizin bewerben, die ein wirkliches Interesse an dieser Disziplin bzw. dem Beruf haben. Es lag früher der Verdacht nahe, dass wegen des niedrigeren notwendigen Test-Prozentranges für die Zulassung erst einmal Zahnmedizin begonnen, dann aber ein Umstieg Richtung Humanmedizin versucht wurde. Die Bedingungen für einen Disziplinenwechsel ab dem 2. Studienjahr (in alle möglichen Richtungen) sind in den jeweiligen kantonalen Verordnungen klar geregelt. Der Studienwechsel ist kein „Automatismus“. Es bedarf immer zweier Voraussetzungen:

- Es muss der Zulassungswert für Humanmedizin erreicht werden, was in der Regel eine erneute Testteilnahme erfordert. Da der Zulassungswert aufgrund der Nachfrageentwicklung von Jahr zu Jahr steigt, wird die Hürde also höher als im Jahr der Erstbewerbung. Nur rund 1/3 der Kandidaten erfüllen pro Jahr diese Bedingung.
- Es muss auch Kapazität in Humanmedizin an der gewünschten Universität vorhanden sein, die Universität muss den Wechsel genehmigen (was wegen fehlender Kapazitäten nicht immer möglich ist und durch den Druck zur maximalen Auslastung der Kapazitäten auch eher die Ausnahme sein sollte).

Darauf hinzuweisen ist auch, dass ein Studium der Zahnmedizin auch spezifische Studienanforderungen stellt, unter anderem sehr gut entwickelte manipulative Fähigkeiten und Feinmotorik.

Tabelle 4: EMS-Teilnahme wegen Antrag Studienwechsel Zahn- nach Humanmedizin (in ein höheres Studienjahr).

	Wechselwunsch	davon BS	davon BE	davon FR	davon ZH
2017	12* (6)	6 (4)	1 (0)	2 (1)	3 (1)
2016	18 (8)	6 (3)	2 (0)	5 (1)	5 (4)
2015	34 (13)	19 (7)	5 (2)		10 (4)
2014	37 (20)	17 (6)	15 (11)	2 (2)	3 (1)
2013	37 (14)	16 (6)	8 (4)	3 (1)	10 (3)
2012	28 (7)	15 (3)	7 (2)		6 (2)
2011	25 (8)	14 (1)	6 (4)	1 (0)	4 (3)
2010	30 (10)	20 (5)	9 (5)		1 (0)
2009	23 (8)	16 (4)	1 (0)	2 (1)	4 (3)
2008	11 (4)	9 (3)	2 (1)		
2007	10 (5)	10 (5)			
2006	7 (2)	6 (2)		1 (0)	

* Zusätzlich 2 (1) Personen mit Wechselwunsch nach Veterinärmedizin (BE) .

In Klammern Zahl der Personen, die das Zulassungskriterium erfüllt haben – die allerdings nur dann wechseln dürfen, wenn auch Kapazitäten in Humanmedizin vorhanden sind.

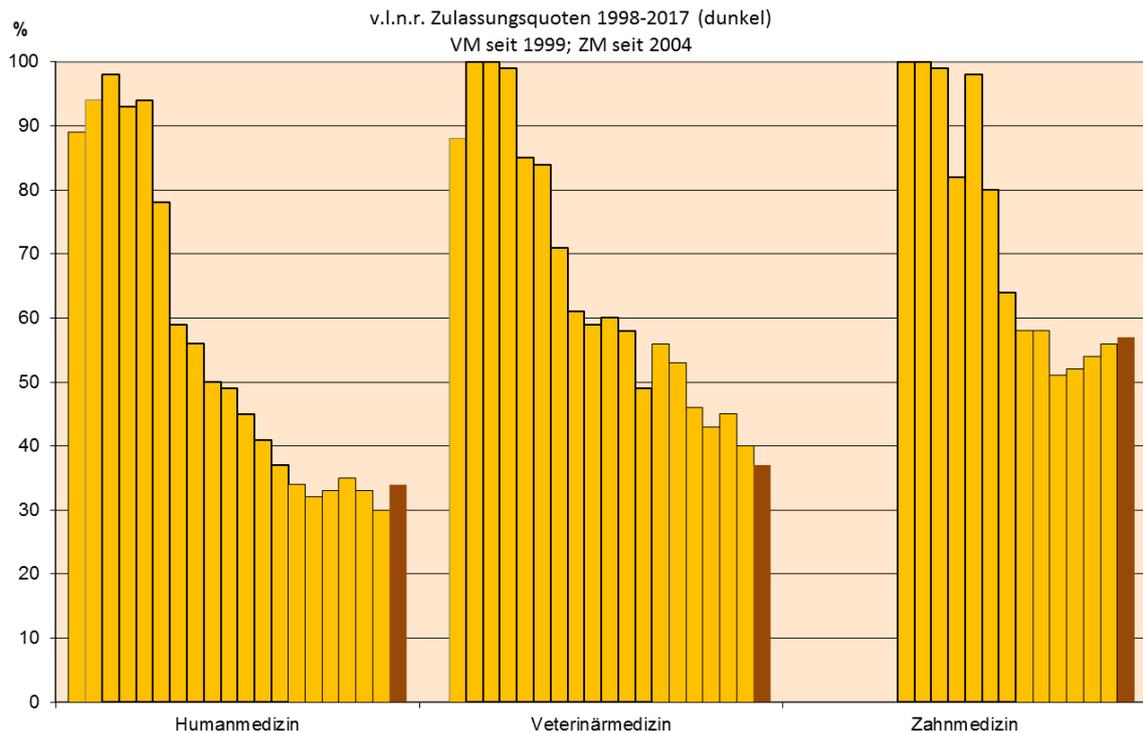


Abbildung 6: Prozent Personen mit gültigem Test-Prozentrang, denen ein Studienplatz zugewiesen werden kann.

4.4 Grösse der Testlokale

In diesem Jahr erfolgte die Testabnahme neu in 34 (+2) Testlokalen/Sektoren verteilt auf 9 Testorte. Die Grösse der Testlokale variierte im bisher üblichen Rahmen. Die Zahl der Betreuer war proportional zur Teilnehmerzahl vorgegeben. In zwei grossen Testlokalen in Zürich wurde wiederum das „Sektorkonzept“ angewendet (13/14 sowie 20/21/22/23), d.h. ein zentraler Testleiter für Zeitnahme und Instruktionsverlesung, alle anderen administrativen Aufgaben in getrennten Sektoren.

Es gibt keine Hinweise seitens des Betreuungspersonals oder der Kandidaten, dass sich unterschiedliche Raumgrössen auf die Ergebnisse auswirken. In den grossen Räumen ist der zur Verfügung stehende Platz für die KandidatInnen teilweise sogar grosszügiger (wenn es sich um grössere Hörsäle handelt).

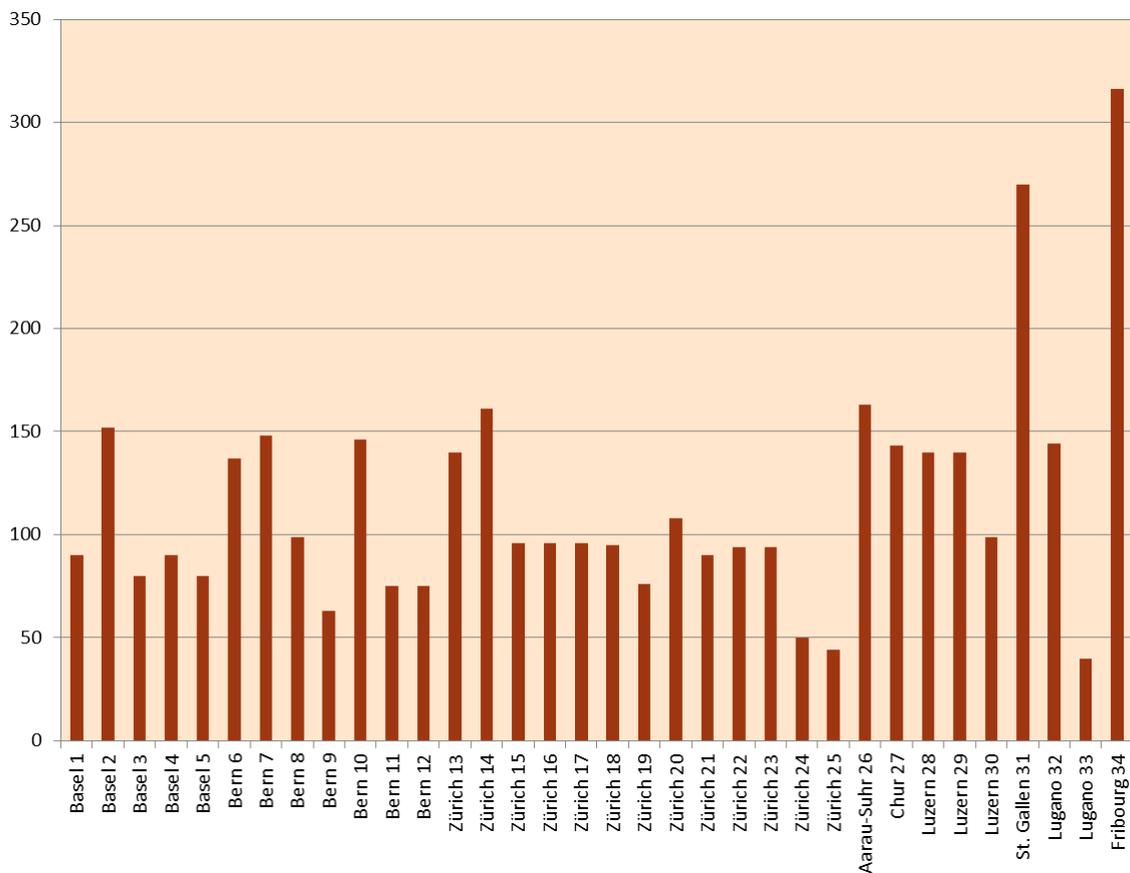


Abbildung 7: Grösse (Zahl der Personen) und Aufteilung der Testlokale 2017.

4.5 Testorte und Wunschhochschulen

Tabelle 5: Testorte und Wunschhochschulen (erste Wahl) 2017 Aufteilung pro Testort auf die Hochschulen.

Disziplin	Testort	Wunschhochschulen								
		BE	BS	ETH	FR	USI	ZH	ZH (LU)	ZH (SG)	Gesamt
Human- medizin	Basel	8	356	12	2	0	15	1	1	395
	Bern	489	19	13	42	0	32	0	0	595
	Chur	24	22	8	1	0	51	0	3	109
	Freiburg	7	7	4	148	0	10	0	0	176
	Lugano	14	38	8	25	6	60	1	0	152
	Luzern	94	74	19	8	0	81	30	0	306
	St. Gallen	34	43	18	3	0	73	0	49	220
	Suhr	29	35	11	2	0	42	2	0	121
	Zürich	42	49	91	4	0	842	6	2	1036
	Gesamt		741	643	184	235	6	1206	40	55
Veterinär- medizin	Basel	21					13			34
	Bern	74					0			74
	Chur	2					17			19
	Freiburg	117					5			122
	Lugano	12					15			27
	Luzern	24					16			40
	St. Gallen	4					19			23
	Suhr	8					7			15
	Zürich	4					80			84
	Gesamt		266					172		
Zahnmedizin	Basel	0	27				0			27
	Bern	38	7				3			48
	Chur	1	2				8			11
	Freiburg	0	1				0			1
	Lugano	1	0				1			2
	Luzern	5	4				11			20
	St. Gallen	0	6				12			18
	Suhr	4	7				8			19
	Zürich	2	0				57			59
	Gesamt		51	54				100		
Gesamt	Basel	29	383	12	2	0	28	1	1	456
	Bern	601	26	13	42	0	35	0	0	717
	Chur	27	24	8	1	0	76	0	3	139
	Freiburg	124	8	4	148	0	15	0	0	299
	Lugano	27	38	8	25	6	76	1	0	181
	Luzern	123	78	19	8	0	108	30	0	366
	St. Gallen	38	49	18	3	0	104	0	49	261
	Suhr	41	42	11	2	0	57	2	0	155
	Zürich	48	49	91	4	0	979	6	2	1179
Gesamt		1058	697	184	235	6	1478	40	55	3753

Die Testorte entsprechen in der Mehrheit dem Standort der gewünschten Hochschule. Für die Zuteilung der Studienplätze ist der Ort der Testabsolvierung allerdings nicht relevant.

4.6 Wunschhochschule und Testort nach Wohnkanton

Die folgenden Tabellen zeigen die Wunschhochschulen nach Wohnkanton für die einzelnen Disziplinen. Angegeben sind auch Vergleichswerte (Total nach Kanton) seit 2011.

Tabelle 6: Gewünschte Studienorte Humanmedizin nach Wohnkantonen 2017 und Vergleich mit 2011 bis 2016 (1998 bis 2010 siehe Bericht 17 des ZTD).

Wohnkanton/ Wohnort	Humanmedizin 2017 gewünschter Studienort									Total					
	BE	BS	ETH	FR	USI	ZH	ZH (LU)	ZH (SG)	Total	2016	2015	2014	2013	2012	2011
AG	46	84	21	3	0	101	1	1	257	252	226	226	198	185	172
AI	2	2	0	0	0	3	0	0	7	9	10	8	12	12	8
AR	6	6	2	0	0	11	0	4	29	25	24	23	26	19	23
BE	400	14	12	26	0	23	0	0	475	389	372	370	309	352	323
BL	3	151	6	1	0	3	0	2	166	169	173	129	142	138	113
BS	0	126	4	1	0	7	1	0	139	150	148	126	131	132	115
FR	11	0	1	93	0	6	0	0	111	101	94	106	78	94	96
GE	0	0	3	5	0	4	0	0	12	14	12	3	3	3	4
GL	1	1	2	0	0	11	0	0	15	11	7	12	15	8	11
GR	18	18	6	3	0	36	0	0	81	70	60	71	72	68	64
JU	4	5	0	7	0	2	0	0	18	16	8	5	5	3	10
LU	81	50	8	3	0	35	19	0	196	193	192	162	147	134	133
NE	2	0	1	4	0	1	0	0	8	7	7	7	7	5	4
NW	8	5	1	1	0	1	4	0	20	22	8	12	15	13	7
OW	4	2	1	1	0	4	2	0	14	12	13	14	7	10	9
SG	30	35	15	0	0	78	0	41	199	186	155	167	137	150	139
SH	0	2	3	0	0	19	0	0	24	13	26	26	30	32	21
SO	48	41	6	2	0	8	0	0	105	102	85	98	86	72	75
SZ	6	5	10	2	0	39	7	0	69	60	62	53	41	36	29
TG	9	10	5	2	0	46	0	6	78	81	68	57	70	69	66
TI	14	37	7	24	5	54	0	0	141	116	94	93	108	97	86
UR	3	3	2	0	0	7	0	0	15	10	5	9	10	8	9
VD	0	2	0	29	0	2	0	0	33	21	26	20	15	15	16
VS	21	6	1	24	0	4	0	1	57	61	43	30	39	39	43
ZG	6	13	7	1	0	39	2	0	68	54	55	51	54	49	50
ZH	12	13	58	1	0	637	3	0	724	657	630	586	599	538	481
FL*	0	3	0	0	0	6	0	0	9	10	11	14	16	9	4
Übriges Ausland	6	9	2	2	1	19	1	0	40	53	45	47	31	47	61
Total	741	643	184	235	6	1206	40	55	3110	2864	2659	2525	2403	2337	2172

*Das Fürstentum Liechtenstein (FL) wird als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Tabelle 7: Gewünschte Studienorte Veterinärmedizin nach Wohnkantonen 2017 und Vergleich mit 2007 bis 2016 (1999 bis 2006 siehe z.B. Bericht 17 des ZTD).

Wohnkanton/ Wohnort	Veterinärmedizin 2017 gewünschter Studienort			Total									
	BE	ZH	Total	2016	2015	2014	2013	2012	2011	2010	2009	2008	2007
AG	10	16	26	27	23	24	24	21	21	24	11	14	16
AI	0	1	1	2	1	1	0	1	0	0	1	1	1
AR	1	2	3	3	4	4	5	2	2	1	1	1	1
BE	61	0	61	59	59	55	51	45	41	62	32	40	35
BL	8	7	15	12	20	17	15	17	10	14	4	10	15
BS	8	4	12	9	10	7	5	5	4	7	8	7	7
FR	19	0	19	17	12	17	26	10	17	17	17	19	10
GE	14	2	16	17	21	18	12	12	12	16	15	6	12
GL	0	4	4	7	4	5	4	2	3	2	1	3	2
GR	2	8	10	11	15	11	12	12	14	10	12	9	9
JU	5	0	5	2	3	4	4	2	1	1	2	5	8
LU	20	6	26	24	16	16	17	19	20	28	20	14	16
NE	15	0	15	17	6	11	8	7	3	3	13	8	6
NW	1	1	2	2	2	2	2	1	0	2	0	1	3
OW	1	0	1	1	2	4	2	1	2	1	1	1	1
SG	4	19	23	24	23	15	16	16	14	13	18	17	20
SH	1	7	8	2	6	5	2	2	4	2	1	2	6
SO	12	2	14	10	7	13	12	7	9	8	10	9	6
SZ	0	2	2	2	4	11	12	12	10	3	4	4	5
TG	1	10	11	12	14	10	10	7	4	7	8	10	12
TI	12	15	27	17	9	24	12	17	17	17	8	17	11
UR	1	0	1	1	3	0	0	4	0	0	1	0	1
VD	52	3	55	39	29	32	27	26	27	29	27	22	19
VS	12	0	12	25	20	16	10	12	7	17	11	8	5
ZG	2	6	8	4	4	4	5	5	2	3	2	6	5
ZH	1	53	54	57	53	54	48	45	44	50	49	32	50
FL*	0	3	3	3	0	1	2	1	2	1	1	1	0
Übriges Ausland	3	1	4	2	5	3	7	6	5	4	3	5	6
Total	266	172	438	408	375	384	350	317	295	342	281	272	288

*Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

Tabelle 8: Gewünschte Studienorte Zahnmedizin nach Wohnkantonen 2017 und Vergleich mit 2007 bis 2016. (2004 bis 2006 siehe z.B. Bericht 17 des ZTD).

Wohnkanton/ Wohnort	Zahnmedizin 2017 gewünschter Studienort				Total									
	BE	BS	ZH	Total	2016	2015	2014	2013	2012	2011	2010	2009	2008	2007
AG	6	7	15	28	27	22	23	26	33	23	16	13	11	11
AI	0	0	1	1	0	0	1	2	0	1	0	0	2	1
AR	32	5	1	38	1	4	2	2	1	3	5	3	0	1
BE	0	7	0	7	55	43	43	48	38	36	31	41	22	31
BL	0	14	0	14	14	14	17	17	12	19	10	10	12	11
BS	6	7	15	28	11	15	14	23	22	10	10	7	11	5
FR	1	0	0	1	7	2	4	4	8	7	7	5	4	6
GE	0	0	3	3	0	0	0	0	0	0	0	0	0	0
GL	0	0	9	9	2	2	2	2	3	3	0	0	0	2
GR	0	1	0	1	7	3	3	6	6	5	6	6	4	6
JU	3	2	7	12	0	1	1	0	0	0	2	0	2	2
LU	1	0	0	1	14	14	22	14	13	15	19	7	11	9
NE	0	1	0	1	1	0	0	0	0	1	0	0	0	1
NW	1	6	13	20	1	1	2	4	1	1	2	0	2	0
OW	0	1	1	2	3	1	2	1	0	1	1	0	1	0
SG	3	7	3	13	19	27	21	24	24	12	13	18	17	17
SH	1	0	5	6	4	3	4	1	2	0	0	1	3	0
SO	1	0	4	5	8	7	12	15	15	15	14	10	10	13
SZ	1	0	1	2	3	6	8	5	3	4	8	2	4	3
TG	0	0	5	5	9	8	5	5	5	2	9	5	4	2
TI	1	0	31	32	4	4	8	4	9	8	7	2	5	10
UR	1	0	0	1	0	2	1	1	3	2	2	0	1	3
VD	0	0	3	3	1	2	2	1	0	2	2	0	1	0
VS	0	0	9	9	1	0	2	3	0	0	2	2	1	4
ZG	0	1	0	1	7	8	7	3	2	2	1	3	3	2
ZH	3	2	7	12	50	61	53	58	43	56	44	36	15	25
FL*	0	1	0	1	1	2	2	4	0	3	1	1	4	1
Übriges Ausland	0	2	1	3	1	6	3	2	5	8	8	2	4	3
Total	51	54	100	205	251	258	264	275	248	239	220	174	154	169

*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. Sie dient u.a. der künftigen Kapazitätsplanung für die Testorte.

Tabelle 9: Testorte nach Wohnkantonen 2017 (alle Teilnehmenden).

Wohnkanton/ Wohnort	Testort									
	Basel	Bern	Chur	Freiburg	Lugano	Luzern	St. Gallen	Suhr	Zürich	Total
AG	30	11	0	0	1	14	2	136	118	312
AI	0	1	0	0	0	0	7	0	0	8
AR	1	1	0	0	0	0	29	0	2	33
BE	2	543	0	16	0	1	1	0	11	574
BL	182	1	1	0	0	0	1	1	2	188
BS	162	1	0	0	1	1	0	0	2	167
FR	0	38	0	91	0	0	0	0	2	131
GE	0	2	0	26	0	0	0	0	1	29
GL	1	0	6	0	0	1	0	0	14	22
GR	3	3	77	1	6	3	1	0	7	101
JU	3	0	0	21	0	0	0	0	0	24
LU	1	8	0	0	0	220	0	2	3	234
NE	1	1	0	20	0	0	0	0	1	23
NW	0	2	0	0	0	19	0	0	2	23
OW	2	0	0	0	0	12	0	0	2	16
SG	3	4	37	1	0	2	160	0	37	244
SH	0	0	0	0	0	0	3	0	32	35
SO	46	60	0	0	0	1	0	15	11	133
SZ	0	0	1	0	0	27	0	0	49	77
TG	3	2	0	0	0	0	49	0	41	95
TI	1	1	1	0	167	0	0	0	0	170
UR	0	0	0	0	0	12	1	0	3	16
VD	0	1	0	87	0	0	0	0	0	88
VS	2	28	0	33	0	1	1	0	4	69
ZG	1	0	0	0	0	49	0	0	31	81
ZH	3	3	4	1	2	1	4	0	794	812
FL*	1	0	12	0	0	0	0	0	0	13
Übriges Ausland	13	7	1	3	4	2	4	1	14	49
Total	461	718	140	300	181	366	263	155	1183	3767

*Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

4.7 Teilnahmen am EMS (NC) nach Kantonen seit 1998

Seit 1998 wurden in der Schweiz 40'789 gültige Ergebnisse im Rahmen des NC erzielt. Seit 1998 unterliegt Humanmedizin, seit 1999 Veterinärmedizin und erst seit 2004 Zahnmedizin einem NC.

Tabelle 10: Gesamtzahl der Teilnehmer am EMS 1998 bis 2017 – nur Hochschulen mit NC.

Wohnkanton/ Wohnort	Disziplin			Geschlecht		Sprache			Total
	Humanmedizin	Veterinärmedizin	Zahnmedizin	männlich	weiblich	deutsch	französisch	italienisch	
AG	2669	354	271	1185	2109	3283	3	8	3294
AI	93	14	8	56	59	115	0	0	115
AR	285	37	31	130	223	352	1	0	353
BE	4855	784	516	2171	3984	6045	109	1	6155
BL	1826	207	169	781	1421	2195	7	0	2202
BS	1696	112	159	741	1226	1958	2	7	1967
FR	1272	244	63	573	1006	563	1013	3	1579
GE	69	222	0	59	232	11	280	0	291
GL	140	47	23	60	150	209	1	0	210
GR	969	184	77	441	789	1169	1	60	1230
JU	91	57	9	54	103	12	144	1	157
LU	2066	294	167	938	1589	2522	1	4	2527
NE	61	143	4	40	168	19	189	0	208
NW	171	27	17	87	128	215	0	0	215
OW	150	18	13	51	130	181	0	0	181
SG	2135	287	241	945	1718	2662	1	0	2663
SH	312	51	28	142	249	391	0	0	391
SO	1153	161	153	569	898	1465	2	0	1467
SZ	606	91	59	229	527	754	1	1	756
TG	925	162	78	364	801	1165	0	0	1165
TI	1232	237	78	636	911	35	1	1511	1547
UR	140	18	19	69	108	177	0	0	177
VD	208	478	11	152	545	42	655	0	697
VS	529	190	28	265	482	435	312	0	747
ZG	623	84	47	306	448	752	1	1	754
ZH	7518	804	567	3107	5782	8837	20	32	8889
FL	140	20	25	64	121	185	0	0	185
Ausland	543	70	54	264	403	597	48	22	667
Total	32477	5397	2915	14479	26310	36346	2792	1651	40789

Das Fürstentum Liechtenstein (FL) ist als Mitglied der Interkantonalen Universitätsvereinbarung aufgeführt.

4.8 Testabsolvierung nach Geschlecht

In **Humanmedizin** beträgt das Geschlechterverhältnis der Teilnahmen 65 zu 35 zugunsten der Frauen (vergleiche Abbildung 9). Erstmals seit 5 Jahren hat der Anteil männlicher Kandidaten wieder leicht zugenommen.

In **Veterinärmedizin** bleibt der Anteil der Frauen sehr hoch – seit drei Jahren auf konstantem Niveau.

Für **Zahnmedizin** sind die Verhältnisse mit Humanmedizin vergleichbar.

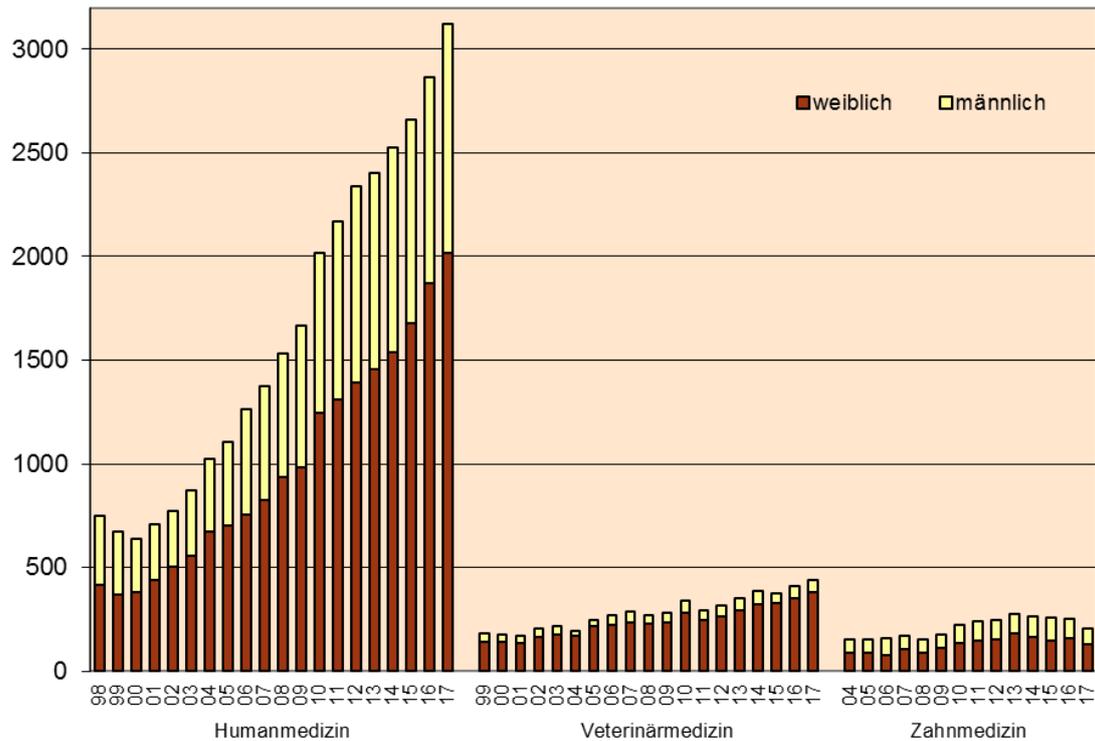


Abbildung 8: Anzahl männlicher und weiblicher Bewerber für Humanmedizin (ab 2008 inklusive Chiropraktik), Veterinärmedizin und Zahnmedizin, Jahre mit NC.

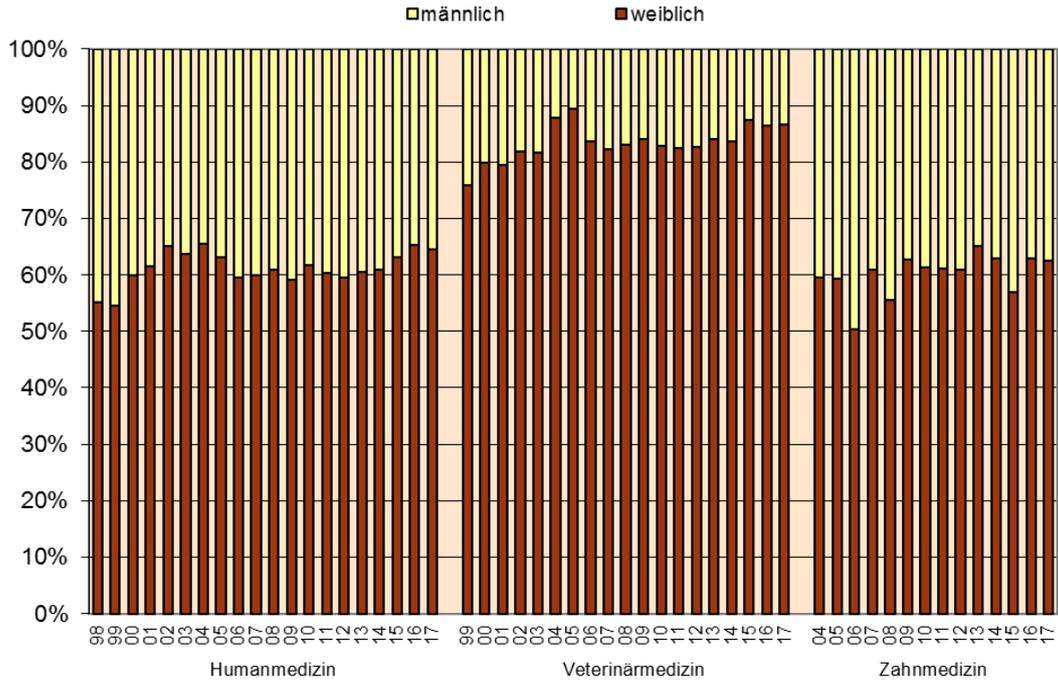


Abbildung 9: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (ab 2008 inklusive Chiropraktik), Veterinärmedizin und Zahnmedizin, Jahre mit NC, bezogen auf 100% pro Jahr und Disziplin.

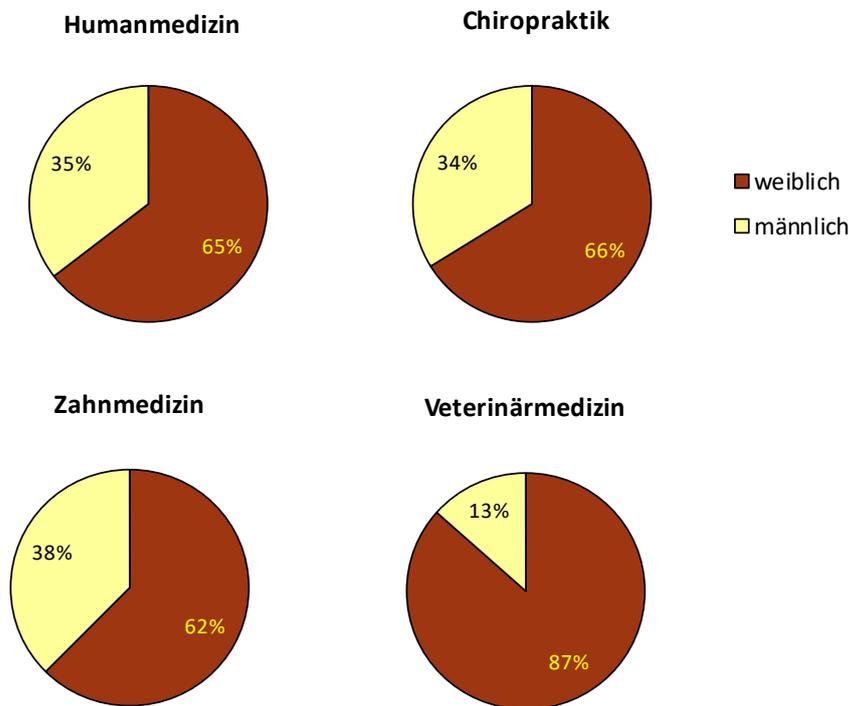


Abbildung 10: Verhältnisse der Geschlechter nach Disziplin im Vergleich 2017.

4.9 Übernahme des Testergebnisses aus dem Vorjahr

Als Alternative zu einem Neuantritt kann ein vorhandenes Ergebnis des Vorjahres übertragen werden, wobei dann immer das neue Testergebnis zählt. Weil der NC bisher von Jahr zu Jahr strenger wird, kann dabei ein Test-Prozentrang im Folgejahr nicht mehr für die Zulassung ausreichen, obwohl er in diesem Jahr noch genügt. Bei der Entscheidung für diese Variante sollte daher eine gute Reserve vorhanden sein, wenn man sich dafür entscheidet.

Deshalb treten viele Personen lieber erneut zum EMS an – vielleicht auch, weil die Wahrscheinlichkeit für eine Verbesserung beim zweiten Testantritt bisher immer hoch war. Dafür sind verschiedene Ursachen denkbar, eine scheint zu sein, dass man sich auf die erste Teilnahme nicht ausreichend vorbereitet hatte und dann bei der zweiten Testteilnahme besser vorbereitet an den Start geht.

In dem Zusammenhang sei noch einmal darauf hingewiesen, dass sich der auf dem Testbescheid mitgeteilte Prozentrang auf alle Teilnehmenden eines Jahrganges, also alle Disziplinen bezieht. Deshalb können diese nur bedingt die Zulassungs-/Ablehnungsquoten pro Disziplin erklären, da zwischen den Disziplinen beträchtliche Unterschiede vorhanden sind und auch die Test-Prozentrangüberträge aus dem Vorjahr noch einzurechnen sind.

Tabelle 11: Test-Prozentrangstatistiken für Personen mit Übernahme des Testergebnisses aus dem Vorjahr.

Disziplin	Jahr	N	Minimum	Maximum	Mittelwert	Stand.- abw.	Zugelassen 1. Runde
Human- medizin	2014	42	69	99	87.9	8.4	39 (3 nicht)
	2015	59	11	99	85.2	14.7	54 (5 nicht)
	2016	70	50	100	86.1	10.1	60 (10 nicht)
	2017	76	21	100	86.2	13.1	69 (7 nicht)
Veterinär- medizin	2014	17	37	97	70.8	15.9	16 (1 nicht)
	2015	8	51	98	67.1	14.6	8 (0 nicht)
	2016	8	65	95	77.7	11.4	8 (0 nicht)
	2017	16	56	100	75.7	15.7	16 (0 nicht)
Zahn- medizin	2014	26	27	79	60.9	11.2	25 (1 nicht)
	2015	19	39	91	64.1	13.8	19 (0 nicht)
	2016	18	10	78	55.6	17.2	16 (2 nicht)
	2017	18	42	99	63.2	15.8	18 (0 nicht)

Zugelassen in der 1. Runde sind Personen, die sofort einen Platz zugeteilt erhielten (ohne Nachrücker).

4.10 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 zu absolvieren. Der Anteil der französisch- und italienischsprachigen Kandidaten bleibt in Human- und Veterinärmedizin über die Jahre vergleichbar.

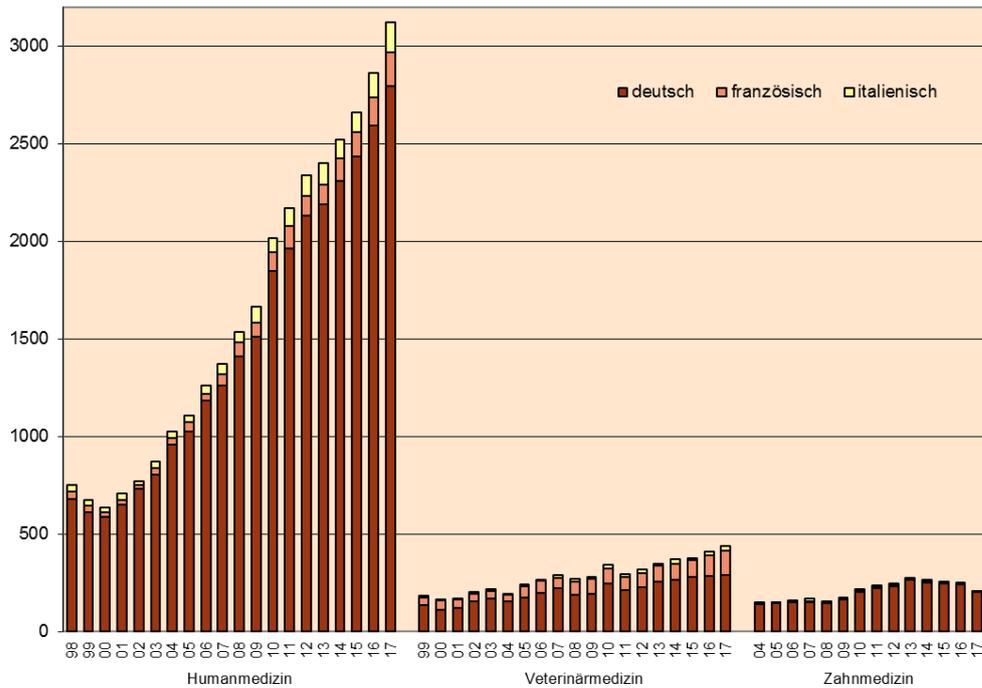


Abbildung 11: Vergleich der Sprachgruppen 1998 bis 2017, Absolutzahlen.

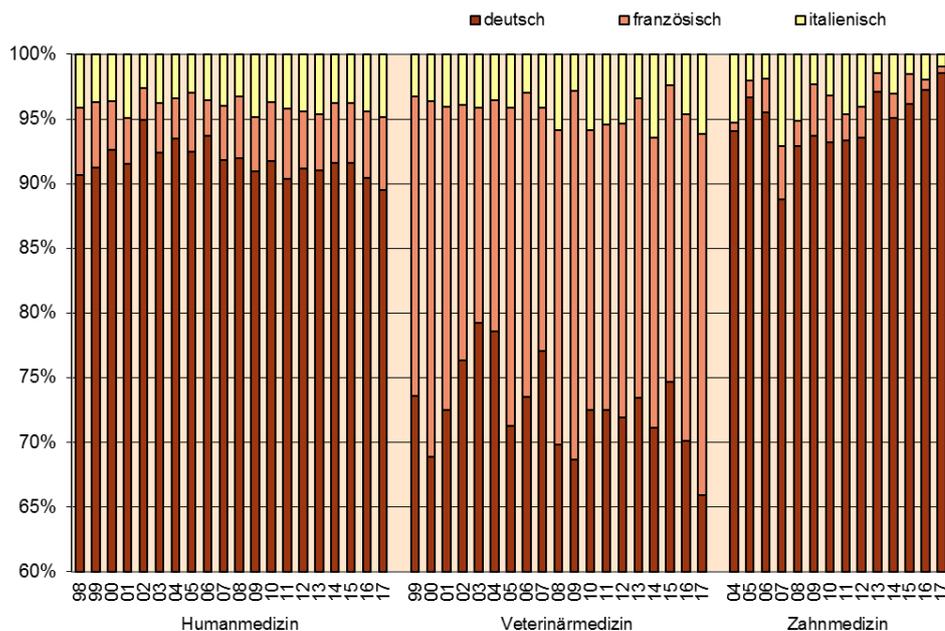


Abbildung 12: Vergleich der Sprachgruppen 1998 bis 2017. Bewerber pro Disziplin und Jahr auf 100% bezogen (dargestellt nur Prozentband zwischen 60 und 100%).

5 Beschreibung des verwendeten Eignungstests

5.1 Aufbau des Tests

Der EMS besteht aus neun Aufgabengruppen, die zu einem Gesamtwert verrechnet werden. Die Tabelle 12 gibt einen Überblick über den Testaufbau, die Aufgabenzahl und die Dauer der einzelnen Aufgabengruppen. Aufgabenbeispiele finden Sie im Anhang.

Tabelle 12: Struktur und Ablauf des EMS

Aufgabengruppe	Aufgaben	Max. Punktzahl	Bearbeitungszeit
Muster zuordnen	20	20	18 min
Med. naturwiss. Grundverständnis	20	20	50 min
Schlauchfiguren	20	20	12 min
Quantitative und formale Probleme	20	20	50 min
Konzentriertes und sorgfältiges Arbeiten	Blatt mit 1600 Zeichen	20	8 min
1 Stunde Pause (bis Wiederbeginn)			
<i>Einprägephase</i> Figuren lernen Fakten lernen	Lernheft mit Vorlagen für beide Aufgabengruppen		4 min
			6 min
Textverständnis	18	18	45 min
<i>Reproduktionsphase</i> Figuren lernen Fakten lernen	20	20	5 min
	20	20	7 min
Diagramme und Tabellen	20	20	50 min
Gesamttest	158 + 1600 Zeichen	178	
Gesamtdauer (inkl. Pause) 9.00 bis ca. 16.00 Uhr			

Beim EMS handelt es sich um eine Adaptation des deutschen **Tests für Medizinische Studiengänge (TMS)** mit Weiterentwicklungen für die Schweiz. Der TMS mit 9 Aufgabengruppen hat sich in Deutschland bewährt und wird seit 1986 eingesetzt (1986 bis 1996 deutschlandweit, seit 2007 erneut für ausgewählte Universitäten). In der Schweiz wird der Test seit 1998 erfolgreich angewendet – bis 2003 mit identischer Struktur wie beim deutschen TMS.

Seit 2004 wird beim Konzentrationstest jährlich eine Version verwendet, die vorher nicht bekannt ist (Zeichen und Durchstreichregel) – nur der Typ der Anforderung und die Zeitdauer bleiben jeweils gleich. Dadurch werden Lerneffekte durch exzessives Üben vermieden.

Im Jahr 2005 wurde die Teststruktur in der Schweiz revidiert. Aufgrund einer Anforderungsanalyse der neuen Studienbedingungen wurde eine neue Aufgabengruppe „Planen und Organisieren“ aufgenommen, um die Planungskompetenz mit zu berücksichtigen. Damit die Testlänge vergleichbar bleibt, wird seitdem in der Schweiz auf nicht gewertete Einstreuaufgaben verzichtet (die ursprünglich der Entwicklung neuer Aufgaben dienen). Ab 2016 wurde zur ursprünglichen Teststruktur zurückgekehrt, weil nur noch bisher unbenutzte Aufgaben verwendet werden sollen. „Planen und Organisieren“ wäre wegen des exklusiven Einsatzes in der Schweiz nur mit grossem Aufwand jeweils jährlich neu zu konstruieren, sodass darauf verzichtet wird. Der Test ist dadurch etwas kürzer geworden, da weiterhin keine Einstreuaufgaben verwendet werden.

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.

Ausgangspunkt der Aufgabenentwicklung 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. 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 Eingang in den TMS gefunden.

Jedes Jahr werden neue Aufgaben für die Aufgabengruppen 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 erfolgt in Deutschland mittels „Einstreuaufgaben“, die mit bearbeitet, aber nicht gewertet werden. 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.

5.2 Berechnung der Werte 2017

Die 9 Aufgabengruppen des Eignungstests für das Medizinstudium (EMS) liefern Punkte als Summe richtig gelöster Aufgaben – zwischen 0 und 20 bzw. 18 bei Textverständnis. Aufgrund der Kennwerteanalysen (Schwierigkeiten und Trennschärfen) können alle Aufgaben gewertet werden.

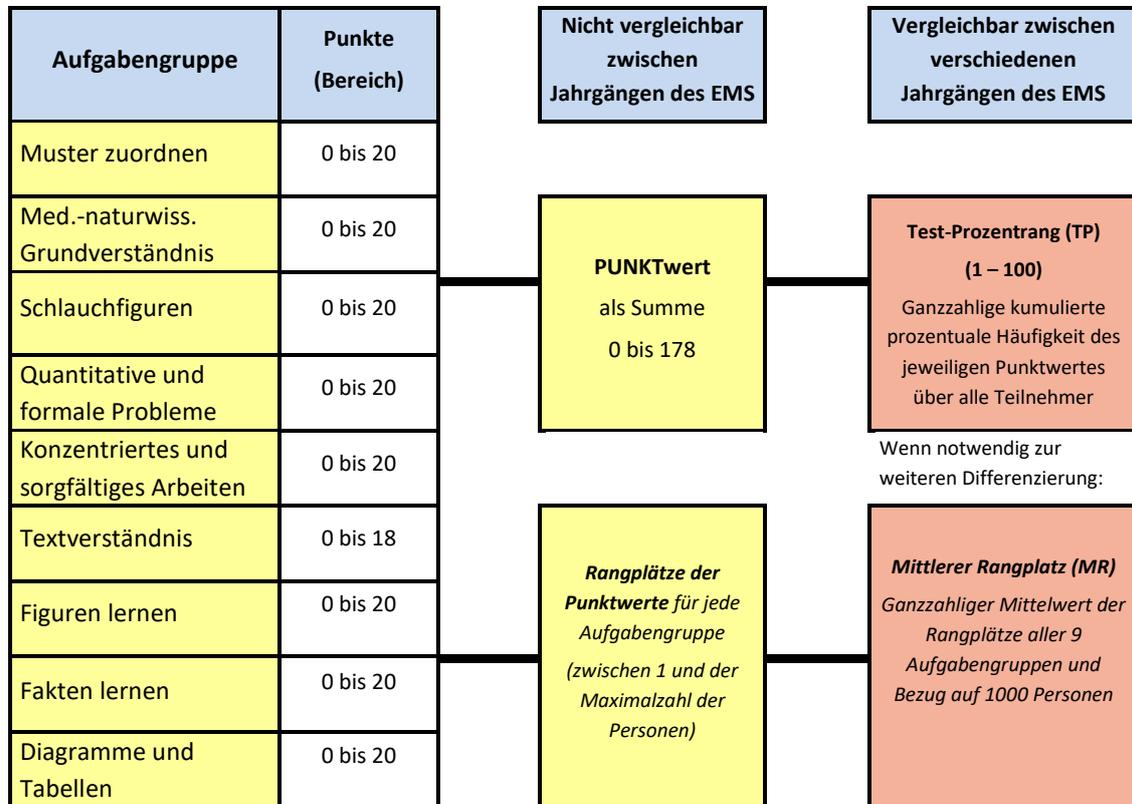


Abbildung 13: Punktwerte der einzelnen Aufgabengruppen 2017 und ihre Zusammenführung über den Punktwert zum Test-Prozentrang und mittlerer Rangplatz.

Für die Aufgabengruppe „Konzentriertes und sorgfältiges Arbeiten“ müssen insgesamt 1600 Zeichen der Reihe nach bearbeitet werden – davon sind 400 Zielzeichen zu markieren. Es können in der zur Verfügung stehenden Zeit in der Regel nicht alle Zeichen bearbeitet werden. Die Position des letzten markierten Zeichens bestimmt, wie viele Zeichen als bearbeitet gewertet werden. Alle übersehenen und fälschlich markierten Zeichen vor diesem letzten bearbeiteten Zeichen zählen als Fehler und werden von der Menge der korrekt markierten Zeichen abgezogen. Der resultierende Wert ist der „Rohwert“, der in eine Skala zwischen 0 und 20 transformiert wird, um mit den anderen Aufgabengruppen gleichgewichtig zum Punktwert addiert zu werden. Die schlechtesten 2.5% erhalten den Punktwert 0, die besten 2.5% erhalten den Punktwert 20. Der Wertebereich des Rohwertes für die verbleibenden 95% wird in 19 äquidistante Abschnitte eingeteilt und den Punktwerten 1 bis 19 zugeordnet. Obwohl sich die Testanforderung zwischen den Jahren ändert, resultieren so vergleichbare und annähernd normalverteilte Punktwerteverteilungen. Alle Punkte der Aufgabengruppen werden zu einer Summe addiert (**Punktwert**). Dieser Wert hat den Nachteil, dass er nicht zwischen Tests verschiedener Jahre vergleichbar ist (diese Vergleichbarkeit ist eine Voraussetzung, dass Ergebnisse ins Folgejahr übertragen werden können).

Aus der Gesamtpunktzahl (Punktwert) werden zwei Werte berechnet:

- **der Test-Prozentrang (TP) und**
- **der mittlere Rangplatz der Aufgabengruppen (MR).**

Der Test-Prozentrang (TP) zeigt an, wie viele Prozent aller teilnehmenden Personen ein schlechteres oder gleich gutes Ergebnis erreicht haben. Die Differenz zu 100 zeigt an, wie viele Prozent der Personen ein besseres Testergebnis erreicht haben. Er wird direkt aus der kumulierten Häufigkeitsverteilung der Punktwerte aller Personen berechnet, die in einem Jahr am EMS-Test teilnahmen und ist ganzzahlig. Werte unter 10 werden auf 10 gerundet.

Es kann vorkommen, dass durch die bisher kontinuierlich zunehmenden Anmeldezahlen ein für die Zulassung ausreichender Wert im Folgejahr nicht mehr ausreichend ist. Dabei ist schwer vorherzusagen, wie sich die Kapazitäten und Anmeldezahlen für das Folgejahr entwickeln werden. Es sollte eine ausreichende Reserve gegenüber dem für die Zulassung im Jahr der Testabsolvierung ausreichenden Wert einkalkuliert werden, wenn man das Ergebnis übertragen möchte.

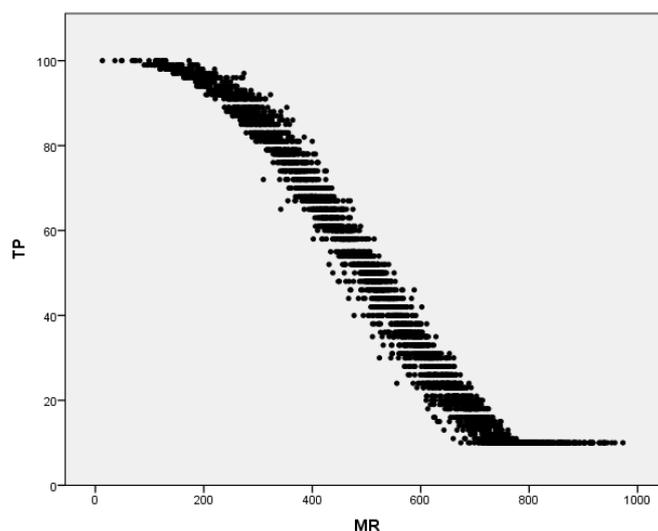
5.3 Mittlerer Rangplatz der Aufgabengruppen

Nur bei gleichem Test-Prozentrang werden die Personen in der Reihenfolge des mittleren Rangplatzes aller Aufgabengruppen (MR) berücksichtigt, wenn nicht mehr alle zugelassen werden können. Diese Rangplätze werden dann über alle 9 Aufgabengruppen gemittelt. Damit diese mittleren Rangplätze ebenfalls zwischen den Jahren vergleichbar sind, werden sie auf 1000 Personen bezogen umgerechnet. Niedrige mittlere Rangplätze entsprechen jeweils den besseren Leistungen. 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 einer schwierigen Aufgabengruppe stärker, indem dort niedrigere Rangplätze für die besten Leistungen resultieren.

Werden in 2 Aufgabengruppen maximal 20 Punkte erzielt, gehen immer 40 Punkte in den Punktwert ein. Haben in einer Aufgabengruppe 100 Personen diesen Wert erreicht, erhalten diese Personen den mittleren Rangplatz 50.5. Haben in einer anderen Aufgabengruppe nur 10 Personen diesen Wert erreicht, ist der Rangplatz 5.5. Dies wird besser bewertet, weil die 20 Punkte der zweiten Aufgabengruppe schwerer zu erzielen waren, deshalb wertvoller sind.

Zur Herstellung der Vergleichbarkeit der Rangplätze zwischen den Jahren wird der mittlere Rangplatz auf die Basis von 1000 Personen umgerechnet.

Abbildung 14: Beziehung zwischen Test-Prozentrang und mittlerem Rangplatz für 2017 (Werte kleiner 10 bei TP auf 10 angehoben).



- Der Mittlere Rangplatz MR bewegt sich zwischen 13 und 973. Niedrige Werte stehen für bessere Leistungen.

- Er korreliert mit dem Test-Prozentrang TP mit 0.98 (trotz der Anhebung aller TP-Werte unter 10 auf 10).

Beispielrechnung

100 Personen nehmen an einem Test mit 2 Aufgabengruppen teil, in denen man je maximal 20 Punkte erreichen kann.

Aufgabengruppe 1:

- Hier erzielen 10 Kandidaten die maximale Punktzahl 20. Bei gleicher Punktzahl erhält jede Person den durchschnittlichen Rang dieser Personen (die 1 bis 10 belegen).
- Kandidat A hat diese maximale Punktzahl erreicht und belegt damit in dieser Aufgabengruppe zusammen mit den anderen 9 Kandidaten Rang 5.5 (Durchschnitt der Ränge 1 bis 10, also $(1+10)/2=5.5$).
- Kandidat B hat, zusammen mit 14 weiteren Kandidaten, 19 Punkte erzielt. Diese Kandidaten erhalten in dieser Aufgabengruppe den Rang 18 (Durchschnitt der Ränge 11 bis 25, also $(11+25)/2=18$).

Aufgabengruppe 2:

- In dieser Aufgabengruppe erreicht nur Kandidat B die volle Punktzahl von 20 Punkten. Er erhält als einziger den Rang 1.
- Kandidat A erreicht hier 19 Punkte, zusammen mit nur 3 weiteren Kandidaten. Sie alle erhalten in dieser Aufgabengruppe den durchschnittlichen Rang $(2+5)/2=3.5$.

Beide Kandidaten erreichten damit die gleiche Punktzahl von 39 Punkten (und den gleichen Test-Prozentrang).

Der Mittlere Rangplatz für die beiden Kandidaten berechnet sich wie folgt:

- Kandidat A: $(5.5 + 3.5)/2=4.5$
- Kandidat B: $(18 + 1)/2=9.5$

Und auf 1000 Personen:

- Kandidat A: $4.5/100 \times 1000=45$
- Kandidat B: $9.5/100 \times 1000=95$

Hätten 1000 Personen am Test teilgenommen, resultiert für Kandidat A im Mittel über beide Aufgabengruppen der 45. Rang, für Kandidat B der 95. Rang. Kandidat A wird dafür belohnt, in der ersten Aufgabengruppe die maximale, aber vor allem auch in der zweiten Aufgabengruppe zusammen mit nur sehr wenigen Kandidaten eine hohe Punktzahl erreicht zu haben.

Kandidat B erhält durch die geringere Punktzahl in der ersten Aufgabengruppe (in dem viele Personen das gleiche oder höhere Ergebnis erzielt haben) einen höheren (schlechteren) Rang.

Der Unterschied zwischen den Kandidaten ist dadurch gerechtfertigt, dass Kandidat A trotz gleicher Punktzahl über beide Aufgabengruppen mehr Kandidaten übertroffen hat als Kandidat B.

Müsste man zwischen beiden Personen unterscheiden, wäre Kandidat A der bessere.

6 Testanwendung in der Schweiz 2017

6.1 Verteilungsprüfung

Die Verteilung des Punktwerts entspricht annähernd einer Normalverteilung. Metrische Prüfverfahren sind bedingt aussagefähig. Die Punktwerte der einzelnen Aufgabengruppen sind wie in allen Vorjahren statistisch nicht als normalverteilt anzusehen (Tabelle 13).

Tabelle 13: Kennwerte der Punktwertskalen Gesamt und Aufgabengruppen

	Punktwert	Quantitative und formale Probleme	Schlauchfiguren	Textverständnis	Med.-naturwiss. Grundv.	Figuren lernen	Fakten lernen	Muster zuordnen	Diagramme und Tabellen	Konzentr. und sorgf. Arbeiten
Mittelwert	97.9	8.7	15.7	10.1	9.7	14.1	9.8	11.3	7.9	10.6
Median	98	8	16	10	10	14	10	11	8	11
Modalwert	95	7	18	10	8	15	10	12	8	12
Stand.-abweichung	19.4	3.2	3.3	2.8	3.5	3.5	3.6	3.4	2.8	4.6
Spannweite	145	18	17	17	20	18	20	19	18	20
25. Perzentil	85	6	14	8	7	12	7	9	6	8
75. Perzentil	112	11	18	12	12	17	12	14	10	14
Schiefe	-0.13	0.32	-0.85	-0.11	0.12	-0.37	0.31	-0.23	0.19	-0.11
Kurtosis	-0.23	-0.18	0.22	-0.19	-0.36	-0.33	-0.28	-0.33	-0.27	-0.41

Die Aufgabengruppe „Figuren lernen“ konnte 2017 durch verschiedene Massnahmen leicht erschwert werden, so dass sie nun wieder besser zwischen den Personen differenziert. Einen sehr hohen Mittelwert (sowie Modalwert) erreicht weiterhin die Aufgabengruppe „Schlauchfiguren“. Die geringsten Punktzahlen wurden 2017 in den Aufgabengruppen „Quantitative und formale Probleme“ sowie „Diagramme und Tabellen“ erreicht.

6.2 Äquivalenz der Sprachversionen

Sprachrelevante Aufgaben wurden 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 (vergleichbare Ergebnisse bei vergleichbaren Fähigkeiten) zwischen den Sprachversionen zu schaffen.

Die Chancengleichheit als Gleichbehandlung der Sprachgruppen muss gewährleistet sein, die Lösungswahrscheinlichkeit einer Aufgabe muss bei gleicher Fähigkeit tatsächlich auch vergleichbar sein. Hervorzuheben ist, dass in französischer und italienischer Sprache die Schlussredaktion des Tests von nativ französisch- bzw. italienischsprachigen Personen durchgeführt wird. Diese wurden dahingehend geschult, dass es vor allem auf eine Übertragung der Schwierigkeit in die jeweilige Sprache ankommt.

Die Gleichheit ist nicht allein durch Übersetzung zu beeinflussen, sondern hängt auch von kulturellen und weiteren Faktoren ab. Bei allen Adaptationen muss mit Unterschieden hinsichtlich der Aufgabenschwierigkeiten zwischen den Sprachversionen gerechnet werden.

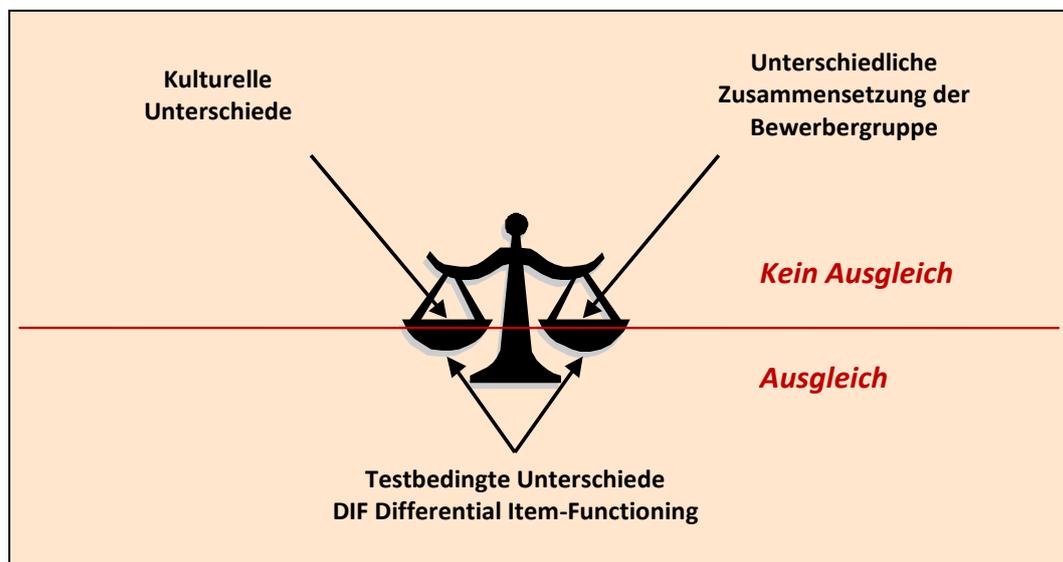


Abbildung 15: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen.

Zwei Ursachen (kulturelle Unterschiede und die unterschiedliche Zusammensetzung der Bewerbergruppe) führen potentiell zu „echten“ Personen-Unterschieden, die nicht auf den Test oder seine Adaptation zurückgeführt werden können. Sie müssen durch die Ergebnisse entsprechend abgebildet werden – hier erfolgt kein Ausgleich.

Testbedingte Unterschiede hingegen werden durch die Adaptation verursacht und sollen möglichst vollständig ausgeglichen werden. Mögliche Ursache wäre die 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 Leistungsvoraussetzungen (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 einer Aufgabe 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 fünf von neun Aufgabengruppen. Der Sprachausgleich beschränkt sich demnach auf die stärker sprachabhängigen Aufgabengruppen „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ sowie „Diagramme und Tabellen“.
2. Den Ausgleich nur vorzunehmen, wenn signifikante Mittelwertunterschiede in der entsprechenden Aufgabengruppe vorhanden sind. Geprüft werden dazu die Abweichungen der jeweiligen Zielsprache von der deutschen Sprachgruppe.

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

Spricht man von DIF, so muss dieses von systematischen Mittelwertsdifferenzen, welche alle Aufgaben einer Gruppe gleichermaßen betreffen, abgegrenzt werden. Die Ergebnisse der letzten Jahre haben gezeigt, dass sich die Sprachgruppen in manchen Aufgabengruppen signifikant unterscheiden. Differenzen kommen nicht nur durch einzelne Aufgaben zustande, sondern auch durch eine Verschiebung der Schwierigkeiten aller Aufgaben einer Gruppe. Es ist unwahrscheinlich, dass diese systematischen Unterschiede testbedingt sind. Testbedingte Unterschiede als sprachliche Besonderheiten zeigen sich in spezifischen Abweichungen einzelner Aufgaben. 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.

6.2.1 Sprachvergleich für die Aufgabengruppen

Auch 2017 treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Aufgabengruppen auf. Die unkorrigierten Punktzahlen für die Sprachgruppen sind nachfolgend abgebildet.

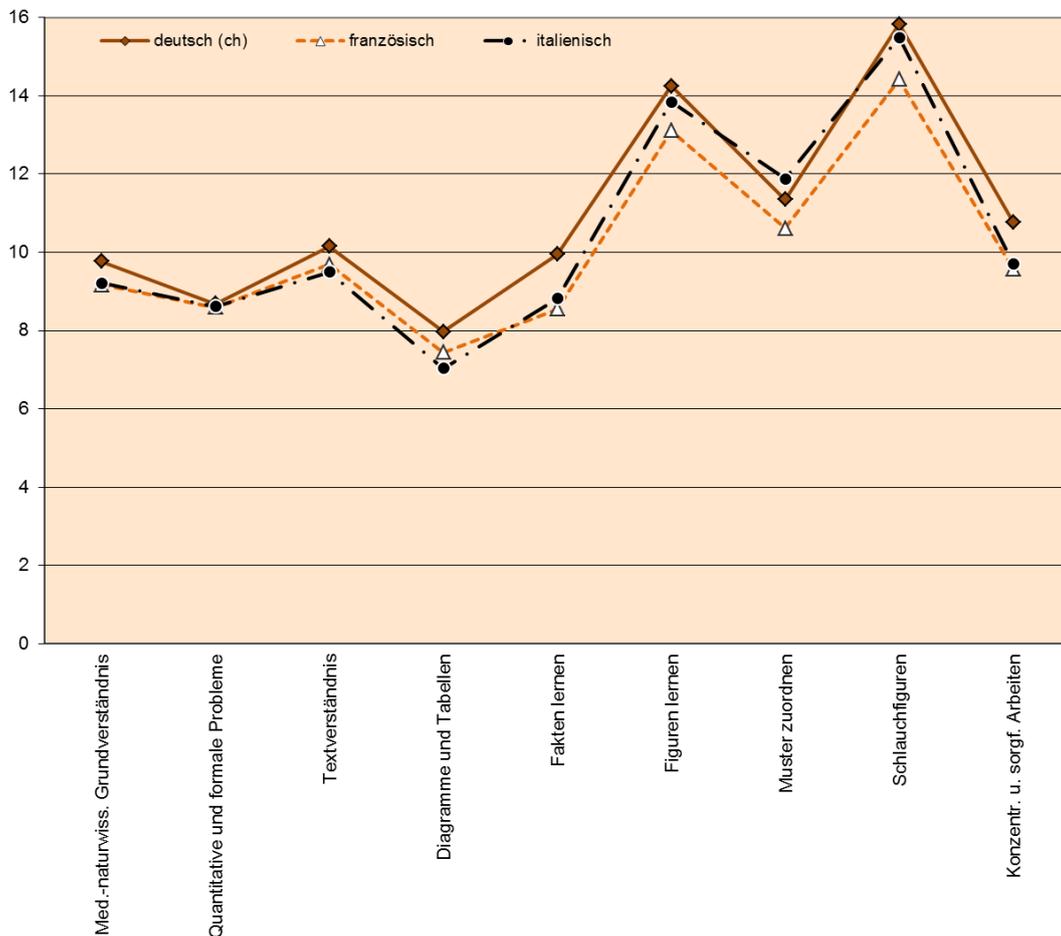


Abbildung 16: Mittelwerte der Punktwerte für die Aufgabengruppen nach Sprachen (unkorrigiert).

Abbildung 17 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmern für das Jahr 2017, in Abbildung 18 ist der Vergleich zwischen deutsch- und italienischsprachigen Personen dargestellt. Die senkrechten Linien bezeichnen Mittelwert und Standardabweichung der Differenzen über die Jahrgänge seit 1998. Höhere Differenzwerte entsprechen besseren Ergebnissen der deutschen Sprachgruppe. Für beide Sprachgruppen lagen die Differenzen 2017 überwiegend im unteren Bereich der bekannten Unterschiede, teilweise sind sie sogar deutlich geringer ausgefallen.

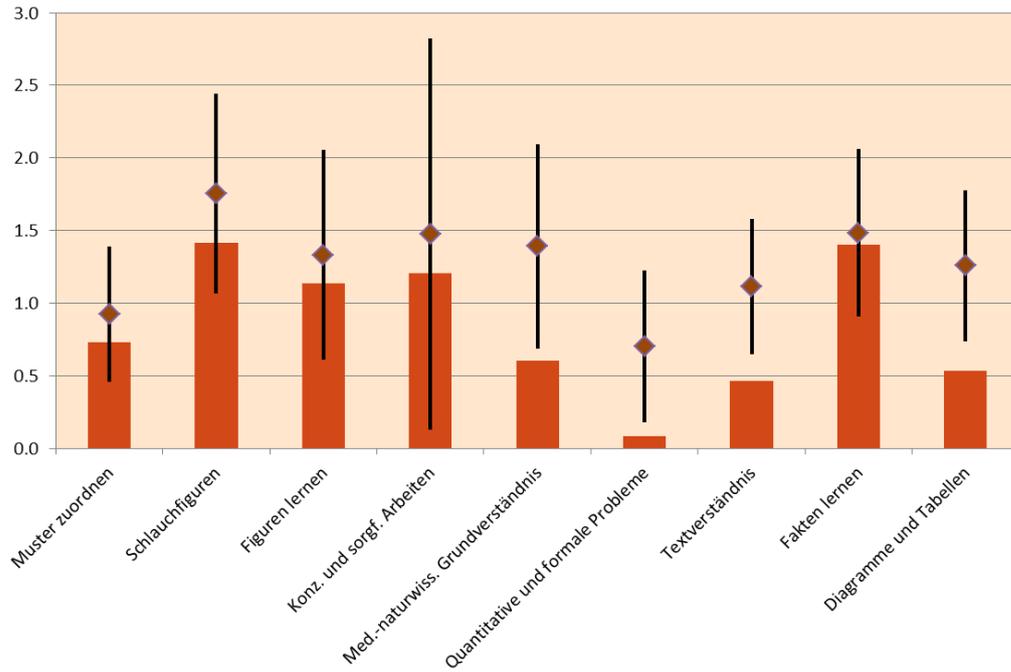


Abbildung 17: Differenz für Aufgabengruppen deutschsprachig-französischsprachig 2017 (unkorrigiert) und Mittelwert mit Standardabweichung aller vorhergehenden Testdurchführungen 1998 – 2016.

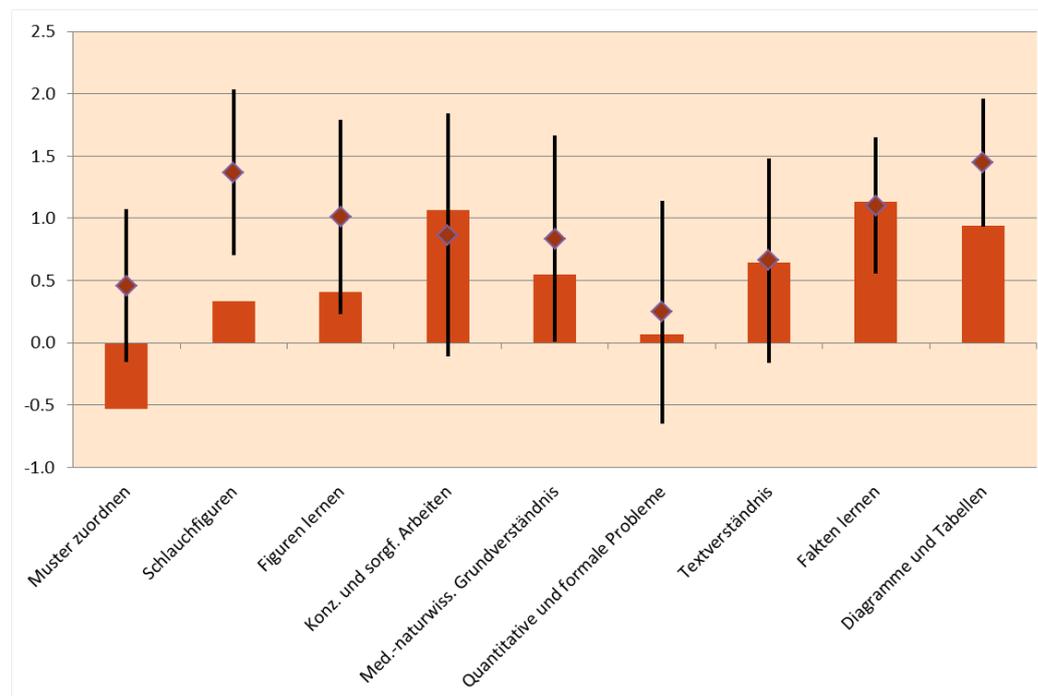


Abbildung 18: Differenz für Aufgabengruppen deutschsprachig-italienischsprachig 2017 (unkorrigiert) und Mittelwert mit Standardabweichung aller vorhergehenden Testdurchführungen 1998 – 2016.

6.2.2 Darstellung des Korrekturverfahrens

6.2.2.1 Identifikation und Ausgleich testbedingter Unterschiede geschehen wie folgt:

- Sprachausgleiche werden nur in den sprachabhängigen Aufgabengruppen „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ sowie „Diagramme und Tabellen“ vorgenommen. Dass Unterschiede in den nichtsprachlichen Tests testbedingt sind (und ausgeglichen werden müssen), ist unwahrscheinlich. Die Aufgaben selber erfordern praktisch keine Sprachkompetenzen und die sprachspezifischen Anforderungen der Instruktionen scheinen so einfach, dass Verständnisunterschiede unwahrscheinlich sind.
- Damit ein Sprachausgleich in einer bestimmten sprachabhängigen Aufgabengruppe 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 die Aufgabengruppe auch nicht zu Mittelwertunterschieden hinsichtlich des Test-Prozentrangs bei. Eine Korrektur allfälliger DIF-Aufgaben (die vorhanden sein können) würde dann eventuell zusätzliche Unterschiede produzieren.
- Es werden jene Aufgaben ausgeglichen, welche über eine systematische Differenz hinaus DIF aufweisen, die also eine andere Position in der Rangreihe der Itemschwierigkeiten einnehmen. Für diese Aufgaben 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.
- Als Sprachausgleich wird den französisch- und italienischsprachigen Probanden die Differenz zur entsprechenden Aufgabenschwierigkeit in der deutschsprachigen Version gutgeschrieben, falls sie die Aufgabe 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-Aufgaben mitbestimmt. Man kann das so interpretieren, dass für DIF-Aufgaben bei falscher Antwort zumindest der Schwierigkeitsunterschied zur Referenzgruppe ausgeglichen wird.

Eine generelle Gutschrift des Mittelwertunterschiedes pro Aufgabengruppe für die französisch- und italienischsprachige Gruppe könnte für einzelne Personen zu mehr Punkten pro Aufgabengruppe führen, als theoretisch maximal vergeben werden und sich auf die Gewichtung der Aufgabengruppen negativ auswirken. Aus diesem Grund wird lediglich die Differenz der Aufgabenschwierigkeit gutgeschrieben. Nur Personen, welche die betreffende Aufgabe nicht korrekt gelöst haben, erhalten diesen Bonus.

6.2.2.2 Welche Aufgabengruppen ausgleichen?

Die folgenden Tabellen zeigen die Sprachvergleiche für die sprachabhängigen Aufgabengruppen. Referenz ist die deutschsprachige Testversion. Die französisch- (Tabelle 14) wie auch die italienischsprachigen (Tabelle 15) Teilnehmer erzielten in (denselben) vier der fünf sprachabhängigen Aufgabengruppen signifikant geringere Punktwerte als die deutschsprachigen Teilnehmer.

Tabelle 14: Signifikanzprüfung deutsche vs. französische Sprachform (sprachabhängige Aufgabengruppen).

Aufgabengruppe	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	3286	9.77	3.49	.004	Ja
	f	181	9.22	3.23		
Quantitative und form. Probleme	d	3286	8.68	3.21	.642	Nein
	f	181	8.61	3.05		
Textverständnis	d	3286	10.15	2.76	.006	Ja
	f	181	9.50	2.88		
Fakten lernen	d	3286	9.95	3.58	.000	Ja
	f	181	8.82	3.61		
Diagramme und Tabellen	d	3286	7.98	2.87	.000	Ja
	f	181	7.03	2.73		

Tabelle 15: Signifikanzprüfung deutsche vs. italienische Sprachform (sprachabhängige Aufgabengruppen).

Aufgabengruppe	Sprache	n	m	s	Sig.	Ausgleich
Med.-naturwiss. Grundverständnis	d	3286	9.77	3.492	.038	Ja
	i	181	9.22	3.235		
Quantitative und form. Probleme	d	3286	8.68	3.213	.773	Nein
	i	181	8.61	3.052		
Textverständnis	d	3286	10.15	2.763	.002	Ja
	i	181	9.50	2.878		
Fakten lernen	d	3286	9.95	3.581	.000	Ja
	i	181	8.82	3.614		
Diagramme und Tabellen	d	3286	7.98	2.868	.000	Ja
	i	181	7.03	2.726		

Mit der Bestimmung der auszugleichenden Aufgabengruppen ist noch nicht eruiert, ob es sich bei den Differenzen um zwischen den Stichproben „real“ vorhandene Differenzen oder um mögliche Test-adaptationsbedingte Unterschiede handelt. Zu diesem Zweck ist die nachfolgend dargestellte Analyse der Aufgaben der betreffenden Aufgabengruppen notwendig.

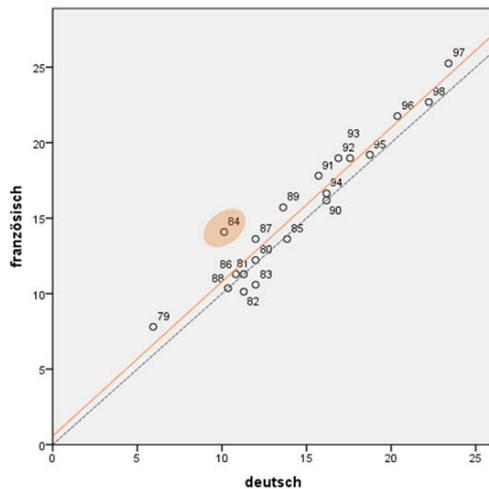
6.2.3 Identifikation von DIF-Aufgaben und Bestimmung des Korrekturwertes

Eine Möglichkeit zur Identifikation von DIF, auch bei relativ kleinen Stichprobengrößen, ist die Methode „Delta-Plot“. Bei diesem Verfahren werden die zu vergleichenden Aufgabenschwierigkeiten z-standardisiert und anschliessend in „Delta-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 wenigen Probanden gelöste) Aufgaben.

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 Aufgabenschwierigkeiten wie auch deren Reihenfolge in beiden Sprachversionen vergleichbar sind.



Die Diagonale wird in den folgenden Abbildungen durch eine unterbrochene schwarze 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 Aufgabengruppen, deren Ursachen eher nicht testbedingt (Übersetzung), sondern in realen Gegebenheiten zu vermuten sind. Um die Regressionsgerade gruppierte Aufgaben folgen also diesen Gegebenheiten und sind demzufolge nicht auszugleichen.

Abbildung 19: Beispiel für einen Delta-Plot.

Auffällige Abweichungen einzelner Punkte von dieser Regressionsgeraden hingegen weisen auf Aufgaben hin, welche zusätzlich zu systematischen Verschiebungen 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 Aufgaben stimmen hingegen in der untersuchten Sprachgruppe bezüglich der „Schwierigkeitshierarchie“ nicht mit der Referenzgruppe überein (sie sind im Vergleich zu den anderen Aufgaben zu schwer oder zu leicht ausgefallen). Die Ursache solcher Differenzen kann eher testbedingt, also beispielsweise in der Übersetzung vermutet werden. Betroffene Aufgaben 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 Aufgabengruppen die kritischen Aufgaben 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 Aufgaben 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 Aufgabenschwierigkeit zur deutschen Aufgabenschwierigkeit gutgeschrieben (also die Differenz der Lösungswahrscheinlichkeit). Dies betrifft nur jene Personen, welche die fragliche Aufgabe nicht korrekt beantwortet haben.

6.2.3.1 Medizinisch-naturwissenschaftliches Grundverständnis

Einzig Item 31 liegt für die französische Sprachgruppe über der kritischen Differenz, es werden den Personen, welche diese Aufgabe nicht korrekt gelöst haben 0.2 Punkte gut geschrieben. Für die italienische Sprachgruppe übertrifft kein Item die kritische Grenze.

Tabelle 16: DIF-Analyse „Medizinisch-naturwissenschaftliches Grundverständnis“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
21	0.79	0.77	0.80	8.36	8.75	8.16		
22	0.67	0.65	0.57	10.70	11.09	12.65		
23	0.59	0.62	0.53	12.26	11.68	13.43		
24	0.65	0.64	0.66	11.09	11.29	10.90		
25	0.56	0.59	0.43	12.85	12.26	15.38		
26	0.60	0.57	0.54	12.07	12.65	13.24		
27	0.62	0.55	0.61	11.68	13.04	11.87		
28	0.60	0.56	0.60	12.07	12.85	12.07		
29	0.54	0.59	0.55	13.24	12.26	13.04		
30	0.58	0.48	0.52	12.46	14.41	13.63		
31	0.53	0.33	0.49	13.43	17.34	14.21	0.20	
32	0.48	0.40	0.45	14.41	15.97	14.99		
33	0.39	0.38	0.43	16.16	16.36	15.38		
34	0.38	0.40	0.36	16.36	15.97	16.75		
35	0.38	0.30	0.31	16.36	17.92	17.73		
36	0.35	0.31	0.24	16.95	17.73	19.09		
37	0.29	0.28	0.26	18.12	18.31	18.70		
38	0.25	0.25	0.27	18.90	18.90	18.51		
39	0.24	0.23	0.27	19.09	19.29	18.51		
40	0.27	0.27	0.33	18.51	18.51	17.34		

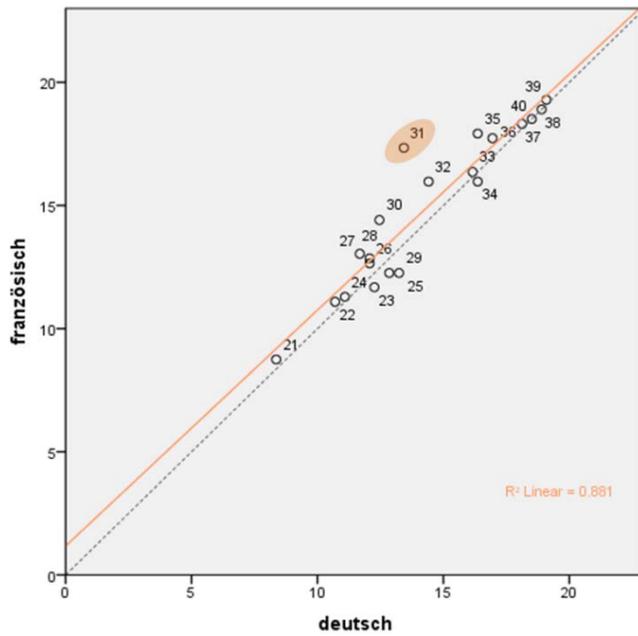


Abbildung 20: Delta-Plot für „Medizinisch-naturwissenschaftliches Grundverständnis“ (deutsch-französisch).

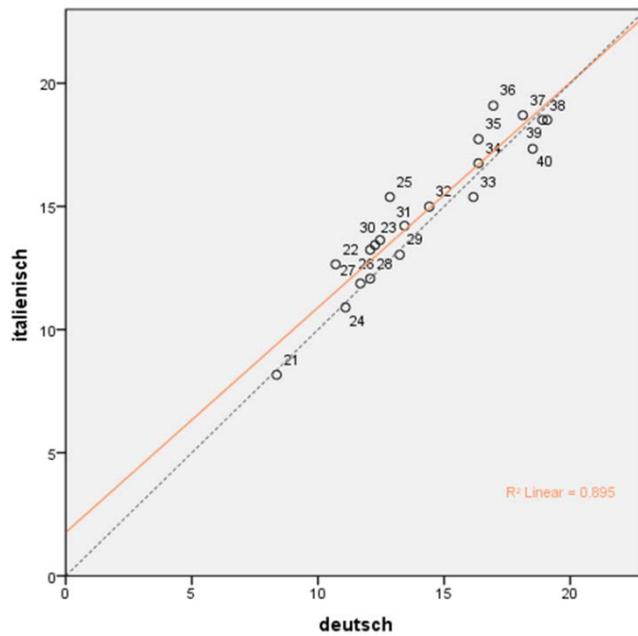


Abbildung 21: Delta-Plot für „Medizinisch-naturwissenschaftliches Grundverständnis“ (deutsch-italienisch).

6.2.3.2 Textverständnis

Der umgekehrte Fall zu „MNGV“ tritt für die Aufgabengruppe „Textverständnis“ ein. Die Aufgaben der französischsprachigen Version verteilen sich eng um die Regressionsgerade (Abbildung 22). In der italienischsprachigen Version fällt lediglich eine Aufgabe auf. Die Differenz beträgt dort 0.14 Punkte.

Tabelle 17: DIF-Analyse „Textverständnis“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
81	0.91	0.89	0.88	6.02	6.41	6.60		
82	0.83	0.76	0.77	7.58	8.94	8.75		
83	0.67	0.63	0.61	10.70	11.48	11.87		
84	0.54	0.57	0.40	13.24	12.65	15.97		0.14
85	0.54	0.48	0.60	13.24	14.41	12.07		
86	0.38	0.36	0.39	16.36	16.75	16.16		
87	0.89	0.85	0.81	6.41	7.19	7.97		
88	0.86	0.82	0.71	6.99	7.77	9.92		
89	0.54	0.51	0.51	13.24	13.82	13.82		
90	0.56	0.57	0.51	12.85	12.65	13.82		
91	0.53	0.57	0.50	13.43	12.65	14.02		
92	0.68	0.56	0.62	10.51	12.85	11.68		
93	0.58	0.46	0.51	12.46	14.80	13.82		
94	0.45	0.42	0.42	14.99	15.58	15.58		
95	0.43	0.47	0.47	15.38	14.60	14.60		
96	0.30	0.26	0.33	17.92	18.70	17.34		
97	0.20	0.25	0.24	19.87	18.90	19.09		
98	0.25	0.26	0.24	18.90	18.70	19.09		

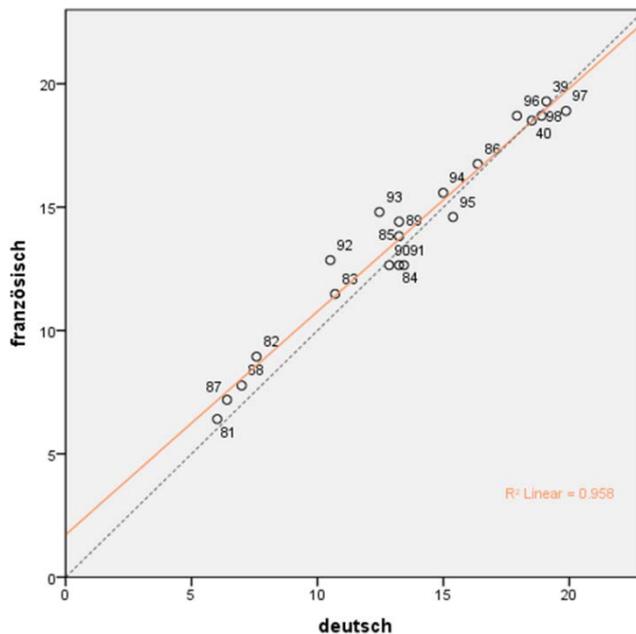


Abbildung 22: Delta-Plot für „Textverständnis“ (deutsch-französisch).

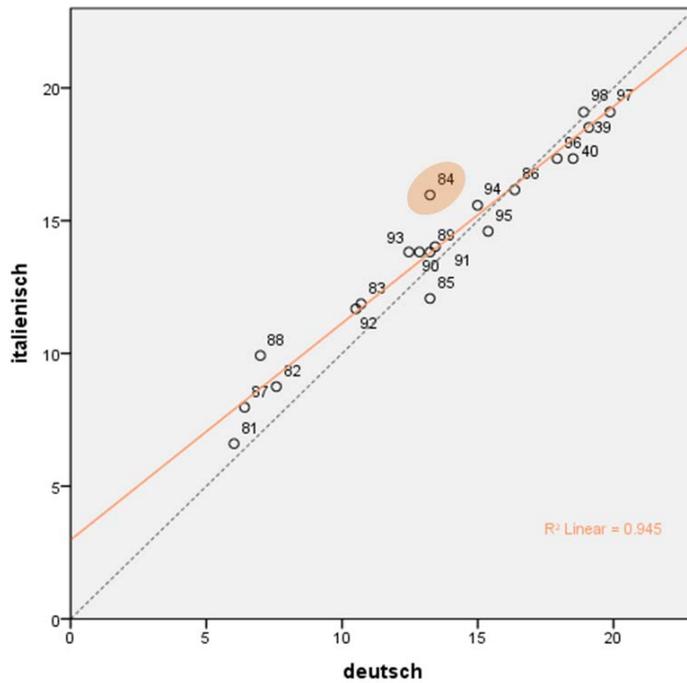


Abbildung 23: Delta-Plot für „Textverständnis“ (deutsch-italienisch).

6.2.3.3 Fakten lernen

In der Aufgabengruppe „Fakten lernen“ treten für beide Sprachgruppen Differenzen auf. Zwei Items in der französischsprachigen (kombiniert maximal 0.44 Punkte), nur ein Item in der italienischsprachigen Version (0.2 Punkte).

Tabelle 18: DIF-Analyse „Fakten lernen“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
119	0.43	0.37	0.45	15.38	16.56	14.99		
120	0.57	0.53	0.54	12.65	13.43	13.24		
121	0.77	0.76	0.71	8.75	8.94	9.92		
122	0.28	0.20	0.22	18.31	19.87	19.48		
123	0.62	0.54	0.57	11.68	13.24	12.65		
124	0.55	0.52	0.42	13.04	13.63	15.58		
125	0.65	0.58	0.58	11.09	12.46	12.46		
126	0.36	0.32	0.38	16.75	17.53	16.36		
127	0.25	0.18	0.18	18.90	20.26	20.26		
128	0.61	0.49	0.56	11.87	14.21	12.85		
129	0.33	0.23	0.32	17.34	19.29	17.53		
130	0.65	0.44	0.54	11.09	15.19	13.24	0.21	
131	0.79	0.72	0.69	8.36	9.73	10.31		
132	0.37	0.14	0.29	16.56	21.04	18.12	0.23	
133	0.51	0.44	0.49	13.82	15.19	14.21		
134	0.37	0.47	0.37	16.56	14.60	16.56		
135	0.36	0.32	0.30	16.75	17.53	17.92		
136	0.37	0.39	0.44	16.56	16.16	15.19		
137	0.63	0.52	0.43	11.48	13.63	15.38		0.20
138	0.47	0.39	0.35	14.60	16.16	16.95		

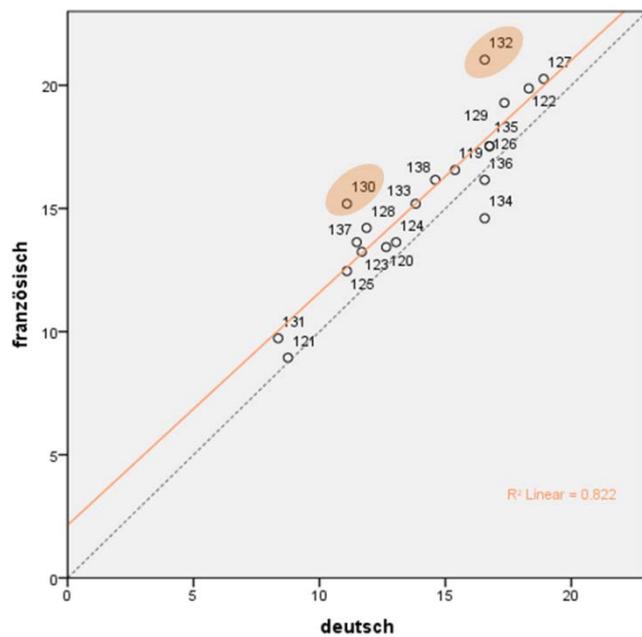


Abbildung 24: Delta-Plot für „Fakten Lernen“ (deutsch-französisch).

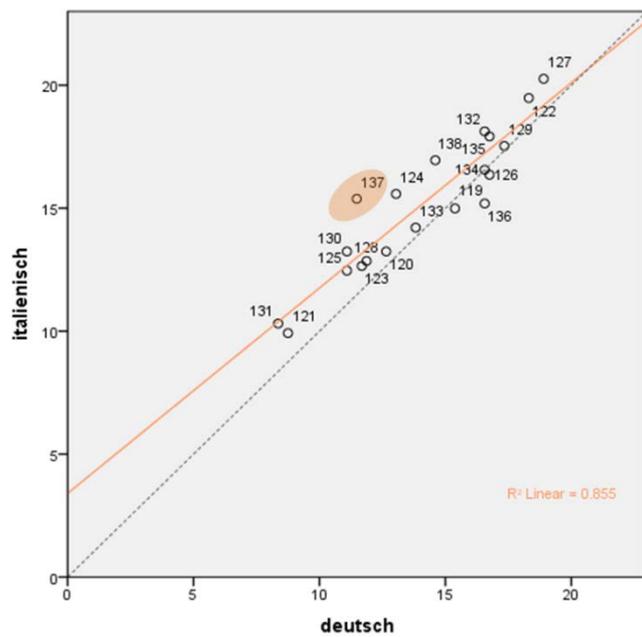


Abbildung 25: Delta-Plot für „Fakten lernen“ (deutsch-italienisch).

6.2.3.4 Diagramme und Tabellen

Auch in der Aufgabengruppe „Diagramme und Tabellen“ ist nur ein geringer Ausgleich zu vergeben. Lediglich für ein Item der italienischsprachigen Version können 0.14 Punkte gutgeschrieben werden.

Tabelle 19: DIF-Analyse „Diagramme und Tabellen“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
139	0.70	0.70	0.55	10.12	10.12	13.04		
140	0.57	0.64	0.50	12.65	11.29	14.02		
141	0.58	0.60	0.54	12.46	12.07	13.24		
142	0.55	0.48	0.52	13.04	14.41	13.63		
143	0.58	0.59	0.47	12.46	12.26	14.60		
144	0.50	0.50	0.41	14.02	14.02	15.77		
145	0.42	0.46	0.45	15.58	14.80	14.99		
146	0.46	0.34	0.38	14.80	17.14	16.36		
147	0.34	0.32	0.34	17.14	17.53	17.14		
148	0.46	0.41	0.49	14.80	15.77	14.21		
149	0.37	0.33	0.29	16.56	17.34	18.12		
150	0.42	0.34	0.39	15.58	17.14	16.16		
151	0.35	0.37	0.36	16.95	16.56	16.75		
152	0.31	0.18	0.24	17.73	20.26	19.09		
153	0.28	0.19	0.20	18.31	20.07	19.87		
154	0.30	0.27	0.16	17.92	18.51	20.65		0.14
155	0.24	0.22	0.23	19.09	19.48	19.29		
156	0.13	0.14	0.14	21.24	21.04	21.04		
157	0.25	0.25	0.28	18.90	18.90	18.31		
158	0.15	0.11	0.09	20.85	21.63	22.02		

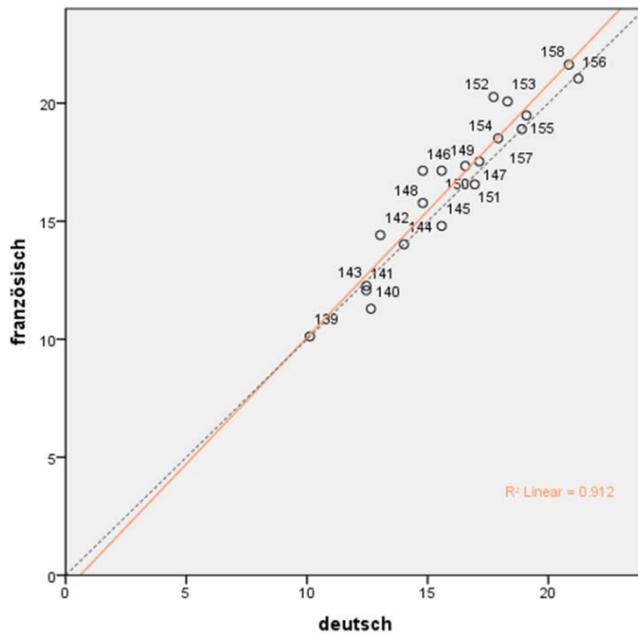


Abbildung 26: Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch).

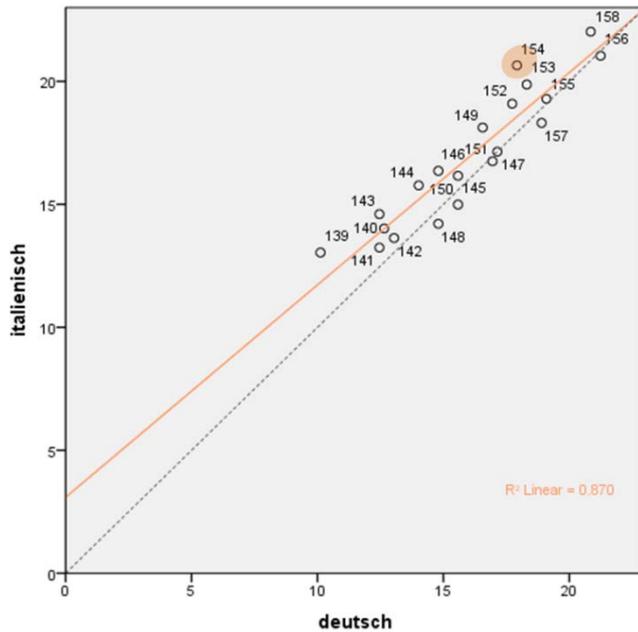


Abbildung 27: Delta-Plot für „Diagramme und Tabellen“ (deutsch-italienisch).

6.2.4 Effekte der Korrektur

Die durch den Ausgleich resultierenden Bonuspunkte können Tabelle 20 entnommen werden. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der maximale Bonus beträgt bei der französischen Sprachgruppe 1 Punkt, die italienische Sprachgruppe erhält auf Testebene keine Bonuspunkte, die zusätzlichen Punkteanteile auf Ebene der Aufgabengruppen haben aber Auswirkungen auf die Bestimmung des mittleren Rangplatzes.

Tabelle 20: Mittelwerte und Standardabweichungen der Korrekturwerte für die Personen der französisch- und italienischsprachigen Gruppen.

		n	Minimum	Maximum	m	s
franz.	Med.-naturw. Grundverständnis	300	0	0.20	0.13	0.09
	Fakten lernen	300	0	0.44	0.32	0.13
	GESAMT (Gerundet)	300	0	1.00	0.34	0.47
ital.	Textverständnis	181	0	0.14	0.08	0.07
	Fakten lernen	181	0	0.20	0.11	0.10
	Diagramme und Tabellen	181	0	0.14	0.12	0.05
	GESAMT (Gerundet)	181	0	0.00	0.11	0.00

Die Mittelwertsdifferenzen für jede Aufgabengruppe nach dem Sprachausgleich sind in Tabelle 21 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 (auf dem 1%-Niveau) höhere Werte erzielt haben als die beiden anderen Sprachgruppen, die sich untereinander nicht signifikant unterscheiden.

Auch nach dem Sprachausgleich bleiben Differenzen zwischen den Sprachgruppen in sprachabhängigen und sprachunabhängigen Aufgabengruppen erhalten – diese können jedoch nicht mehr in der Übersetzung des Tests begründet sein. In Abbildung 28 sind die Ausmasse der (ungerundeten) Bonuspunkte für die bisherigen Testdurchführungen dargestellt. Die senkrechten Linien bezeichnen Mittelwert und Standardabweichung der Jahrgänge 1998-2016.

Tabelle 21: Test-Prozentrang und Punktwerte inklusive Sprachausgleich.

		n	Mittelwert	Stand.abw.	Minimum	Maximum
Test-Prozentrang $d > (f,i)^{**}$	deutsch	3286	52.3	28.2	10	100
	französisch	300	42.0	26.6	10	98
	italienisch	181	45.6	27.3	10	100
Quantitative und formale Probl.	deutsch	3286	8.7	3.2	1	19
	französisch	300	8.6	3.0	2	17
	italienisch	181	8.6	3.1	2	18
Schlauchfiguren $(d,i) > f^{**}$	deutsch	3286	15.8	3.3	3	20
	französisch	300	14.4	3.7	4	20
	italienisch	181	15.5	3.8	4	20
Textverständnis $d > (f,i)^*$	deutsch	3286	10.1	2.8	1	18
	französisch	300	9.7	2.9	2	17
	italienisch	181	9.6	2.9	3	18
Med.-naturwiss. Grundverständnis	deutsch	3286	9.8	3.5	0	20
	französisch	300	9.3	3.4	1	19
	italienisch	181	9.2	3.2	2	19
Figuren lernen $d > (f,i)^{**}$	deutsch	3286	14.3	3.5	2	20
	französisch	300	13.1	3.4	2	20
	italienisch	181	13.8	3.6	5	20
Fakten lernen $d > (f,i)^{**}$	deutsch	3286	10.0	3.6	0	20
	französisch	300	8.9	2.8	1	18
	italienisch	181	8.9	3.6	1	19
Muster zuordnen $(d,i) > f^{**}$	deutsch	3286	11.3	3.4	0	19
	französisch	300	10.6	3.2	1	18
	italienisch	181	11.9	3.5	2	19
Diagramme und Tabellen $d > (f,i)^{**}$	deutsch	3286	8.0	2.9	0	18
	französisch	300	7.4	2.5	2	15
	italienisch	181	7.2	2.7	1	14
Konzentriertes und sorgf. Arbeiten $d > (f,i)^{**}$	deutsch	3286	10.8	4.6	0	20
	französisch	300	9.6	4.6	0	20
	italienisch	181	9.7	4.4	0	19

**/*: signifikanter Einfluss des Faktors „Sprache“ bei varianzanalytischer Prüfung (1%- bzw. 5%-Niveau); darunter signifikante Unterschiede bei multiplem Mittelwertvergleich.

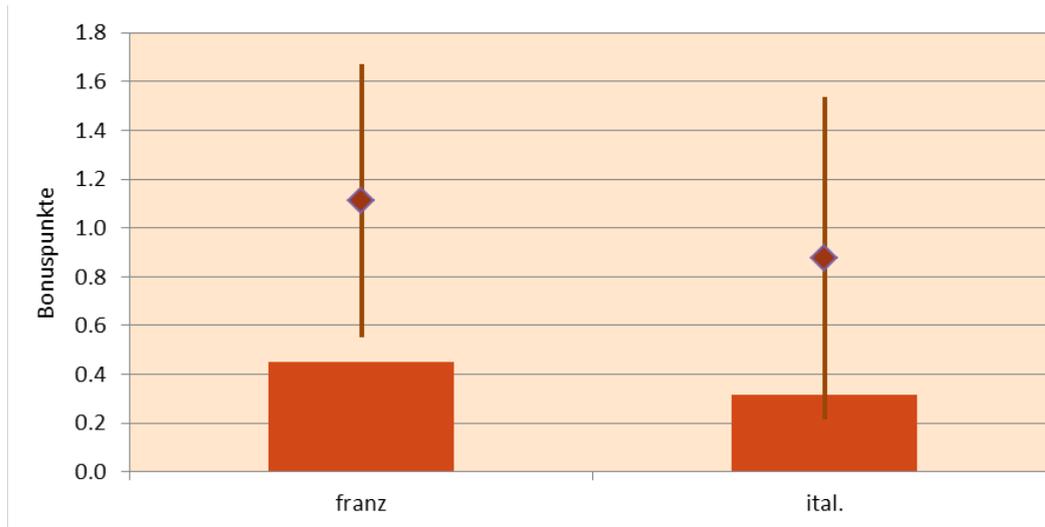


Abbildung 28: Sprachausgleich 2017 (ungerundet) und Mittelwert mit Standardabweichung aller Testdurchführungen 1998 – 2016

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.

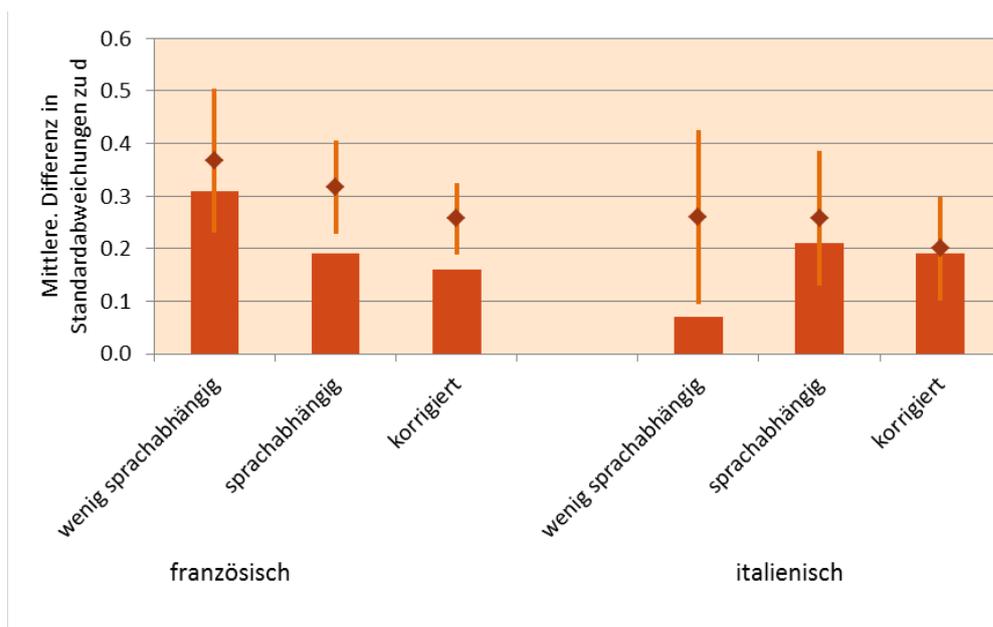


Abbildung 29: Mittlere Abweichung (Delta z) für wenig sprachabhängige und sprachabhängige Aufgabengruppen (korrigiert und unkorrigiert), 2017 und Vergleich zu 1998-2016 (Mittelwert und Streuung).

2017 weist die französische Sprachgruppe bei den sprachabhängigen Aufgabengruppen eine geringere Differenz aus als bei den wenig sprachabhängigen. Nach der Korrektur wird diese weiter vermindert. In der italienischen Version ist das Umgekehrte der Fall. Alle Differenzen liegen in den unteren Bereichen aller bisherigen Testdurchführungen.

6.3 Vergleichbarkeit der Testlokale

Die Durchführungsbedingungen an den einzelnen Testorten sind standardisiert. Hinsichtlich des Test-Prozentrangs erreicht ein kleines (deutschsprachigen) Testlokal in Bern die geringsten Werte. Weitere Differenzen treten vereinzelt (in vergleichsweise kleineren Lokalen) auf – aus den Testdurchführungsprotokollen ergeben sich keine Hinweise auf mögliche Ursachen.

Tabelle 22: Varianzanalytische Prüfung der Homogenität für Test- und Punktwerte zwischen den (deutschsprachigen) Testlokalen.

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testprozentrang	<i>Zwischen Gruppen</i>	38270.57	30	1275.69	1.612	0.019
	<i>Innerhalb Gruppen</i>	2576532.22	3255	791.56		
	<i>Total</i>	2614802.79	3285			
Muster zuordnen	<i>Zwischen Gruppen</i>	564.07	30	18.80	1.609	0.019
	<i>Innerhalb Gruppen</i>	38037.95	3255	11.69		
	<i>Total</i>	38602.03	3285			
Med.-naturwiss. Grundverständnis	<i>Zwischen Gruppen</i>	286.51	30	9.55	0.782	0.796
	<i>Innerhalb Gruppen</i>	39769.39	3255	12.22		
	<i>Total</i>	40055.91	3285			
Schlauchfiguren	<i>Zwischen Gruppen</i>	396.32	30	13.21	1.253	0.162
	<i>Innerhalb Gruppen</i>	34310.28	3255	10.54		
	<i>Total</i>	34706.60	3285			
Quant. und formale Probleme	<i>Zwischen Gruppen</i>	832.01	30	27.73	2.729	0.000
	<i>Innerhalb Gruppen</i>	33082.47	3255	10.16		
	<i>Total</i>	33914.48	3285			
Textverständnis	<i>Zwischen Gruppen</i>	300.79	30	10.03	1.318	0.116
	<i>Innerhalb Gruppen</i>	24768.84	3255	7.61		
	<i>Total</i>	25069.63	3285			
Figuren lernen	<i>Zwischen Gruppen</i>	439.80	30	14.66	1.218	0.193
	<i>Innerhalb Gruppen</i>	39190.07	3255	12.04		
	<i>Total</i>	39629.87	3285			
Fakten lernen	<i>Zwischen Gruppen</i>	558.04	30	18.60	1.457	0.052
	<i>Innerhalb Gruppen</i>	41556.93	3255	12.77		
	<i>Total</i>	42114.97	3285			
Diagramme und Tabellen	<i>Zwischen Gruppen</i>	334.18	30	11.14	1.358	0.092
	<i>Innerhalb Gruppen</i>	26691.97	3255	8.20		
	<i>Total</i>	27026.15	3285			
Konzentr. u. sorgf. Arbeiten	<i>Zwischen Gruppen</i>	909.52	30	30.32	1.436	0.059
	<i>Innerhalb Gruppen</i>	68699.42	3255	21.11		
	<i>Total</i>	69608.94	3285			

Aus den Protokollen der Testabnahme und der Analyse der erzielten Ergebnisse nach Aufgabengruppen ergeben sich keine Hinweise, die als Einschränkungen der Chancengleichheit zu

bewerten wären. Insgesamt gibt es keinen Anlass, Korrekturen aufgrund der Testlokalzuordnung vorzunehmen.

Tabelle 23: Darstellung der detaillierten Ergebnisse nach Testlokalen (deutschsprachig).

Lokal	Testprozentrang	Muster zuordnen	Med.-naturwiss. Grundverständnis	Schlauchfiguren	Quant. und formale Probleme	Textverständnis	Figuren lernen	Fakten lernen	Diagramme und Tabellen	Konzentr. u. sorgf. Arbeiten	Lokalgrösse
1	47.6	11.1	9.1	15.7	7.7	9.6	14.4	9.8	7.4	10.6	83
2	52.7	11.8	9.6	16.1	8.2	9.9	14.1	10.4	8.0	10.7	145
3	57.3	11.9	9.9	16.5	8.1	10.8	15.1	9.8	8.2	11.4	77
4	55.5	12.0	10.0	16.1	8.0	10.0	15.0	10.7	7.4	11.1	84
5	48.8	11.4	9.3	16.0	8.1	9.8	14.0	10.2	7.2	10.1	72
6	48.2	10.9	9.8	15.5	8.4	9.7	14.1	9.2	7.8	10.1	135
7	50.3	11.1	9.9	15.5	8.1	10.4	14.3	9.5	7.9	10.5	141
8	51.0	11.2	10.1	15.5	8.8	10.3	13.5	9.8	8.1	10.8	95
9	50.7	10.8	9.9	16.1	7.8	10.6	14.7	9.8	7.3	10.8	60
10	50.1	11.3	9.8	15.6	8.5	10.2	13.8	9.7	8.0	10.8	142
11	44.3	10.4	9.6	15.3	8.2	9.5	13.7	9.2	7.6	9.2	72
12	49.7	11.3	9.6	15.7	8.3	9.9	14.5	9.3	7.9	10.4	73
13	60.3	12.3	10.3	16.4	9.2	10.5	14.7	10.7	8.3	11.9	131
14	52.5	11.2	9.9	16.0	8.9	9.9	13.8	9.8	8.1	11.0	155
15	55.2	11.8	9.9	16.0	9.5	10.4	14.4	9.7	8.5	10.5	94
16	57.5	11.4	10.3	16.1	9.4	10.3	14.0	10.5	8.8	11.5	92
17	54.4	11.4	10.3	16.4	9.6	10.1	14.2	10.1	8.1	10.2	90
18	54.8	11.4	9.8	16.1	9.1	10.2	14.1	10.4	8.4	11.5	88
19	54.7	12.2	9.4	15.9	8.9	10.6	14.8	9.7	7.8	10.8	73
20	48.9	10.9	9.7	15.5	8.4	10.1	13.7	9.8	7.6	10.6	105
21	57.8	11.7	10.1	16.0	9.9	10.3	14.7	9.9	8.1	11.3	84
22	54.9	11.8	10.0	16.1	9.1	10.1	14.1	10.8	8.3	10.4	89
23	54.3	11.7	9.6	16.0	8.8	10.1	14.3	10.6	7.7	10.5	94
24	54.3	11.4	9.8	16.7	8.8	10.3	14.7	10.2	7.4	10.9	47
25	51.9	11.2	9.5	15.4	9.6	10.3	14.2	10.7	8.2	11.1	41
26	54.6	11.1	9.9	15.8	9.0	10.7	14.6	9.9	8.1	11.2	155
27	51.2	11.3	9.3	16.0	8.5	9.9	14.3	9.9	7.8	10.9	140
28	52.9	11.4	9.8	15.6	8.9	10.3	14.4	9.9	7.9	11.5	136
29	48.7	11.2	9.5	15.4	8.3	9.8	14.4	9.6	8.0	10.5	132
30	46.7	10.6	9.9	15.0	8.4	10.4	13.6	9.9	7.9	9.8	98
31	51.7	11.1	9.5	15.7	8.8	10.0	14.3	9.9	8.2	10.8	263
Total	52.3	11.3	9.8	15.8	8.7	10.1	14.3	10.0	8.0	10.8	3286

Markiert: Mittelwerte für Test-Prozentrang und die Punktwerte der Aufgabengruppen im multiplen Mittelwertsvergleich nach Tukey, die signifikant tiefer als in mindestens einem anderen Testlokal liegen.

6.4 Vergleich für die Geschlechter

Die Forderung nach Gleichbehandlung der Geschlechter konnte in allen bisherigen Testdurchführungen 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. Nach 2006 hat sich der Männeranteil vorerst bei 40% stabilisiert, mittlerweile ist er allerdings weiter auf rund 35% gesunken (2017 gegenüber dem Vorjahr erstmals wieder leicht erhöht). Das ähnliche Verhältnis findet sich seit 2010 in der Zahnmedizin. In der Veterinärmedizin liegt der Männeranteil unter 15%. Die mittleren Test-Prozentränge nach Geschlecht, Sprache und Disziplin können Tabelle 24 entnommen werden.

Tabelle 24: Statistiken für den Test-Prozentrang nach Geschlechtern für Sprache und Disziplin.

		Testsprache			Wunschdisziplin			Gesamt
		deutsch	französisch	italienisch	HM	VM	ZM	
männlich	m	53.4	42.1	45.4	53.2	41.4	47.4	52.2
	s	28.5	27.3	26.8	28.8	23.9	26.0	28.5
	n	n=1099	n=81	n=63	n=1107	n=59	n=77	n=1243
weiblich	m	51.8	41.9	45.7	52.5	45.1	37.6	50.6
	s	28.0	26.3	27.8	28.3	25.6	24.4	28.0
	n	n=2187	n=219	n=118	n=2015	n=381	n=128	n=2524

m: Mittelwert, s: Standardabweichung, n: Personenzahl

Für die zahlenmässig bedeutendste deutschsprachige Teilnehmergruppe wird 2017 ein Unterschied von lediglich 1.6 Test-Prozentrangpunkten registriert, dies entspricht 1.64 (Rohwert-)Punkten und bewegt sich damit im untersten Bereich der bis anhin bekannten Differenzen. Es gibt weiterhin keinen Anlass, von testbedingten Genderdifferenzen auszugehen, die über die unterschiedliche Repräsentativität der Gruppen hinausgehen würden.

Tabelle 25: Kombinierte Statistiken für den Test-Prozentrang nach Geschlechtern, Sprache und Disziplin.

		Humanmedizin			Veterinärmedizin			Zahnmedizin		
		d	f	i	d	f	i	d	f	i
männlich	m	54.4	43.4	44.2	39.9	35.0	63.0	48.1	33.0	10.0
	s	28.8	27.4	26.8	22.2	28.1	18.4	25.9	0.0	0.0
	n	982	69	56	42	11	6	75	1	1
weiblich	m	53.3	45.4	44.5	47.6	38.6	49.5	37.3		80.0
	s	28.3	27.6	27.7	25.3	24.7	27.8	24.2		0.0
	n	1812	107	96	248	112	21	127		1

Betrachtet man die Unterschiede disziplin- und sprachspezifisch, zeigen sich die aus den Vorjahren bekannten Differenzierungen. Lediglich für die deutschsprachige Gruppe mit der Wunschdisziplin Zahnmedizin wird eine signifikante Differenz zwischen den Geschlechtern festgestellt (Tabelle 26).

Tabelle 26: Prüfung zur Varianzhomogenität und der Mittelwertunterschiede (t-Test) für Test-Prozentrang zwischen Geschlechtern (nur für ausreichende Zellengrößen).

Sprache	Disziplin	Levene-Test Gleichheit der Varianzen			t-Test Gleichheit des Mittelwerts		
			F	Sig.	t	df	Sig.
Deutsch	HM	homogen	1.19	0.28	0.89	2792	0.371
	VM	homogen	3.64	0.06	-1.85	288	0.065
	ZM	homogen	0.82	0.37	2.98	200	0.003
Französisch	HM	homogen	0.12	0.73	-0.48	174	0.635
Italienisch	HM	homogen	0.02	0.90	-0.07	150	0.941

Abbildung 30 verdeutlicht die Vergleichbarkeit der Geschlechterdifferenzen mit den Vorjahren. Die diesjährigen Punktwert-Unterschiede auf Ebene der Aufgabengruppen fallen fast durchwegs geringer aus als in den Vorjahren (Mittelwert und Standardabweichung), weisen aber weiterhin die bekannten Ausprägungsrichtungen auf.

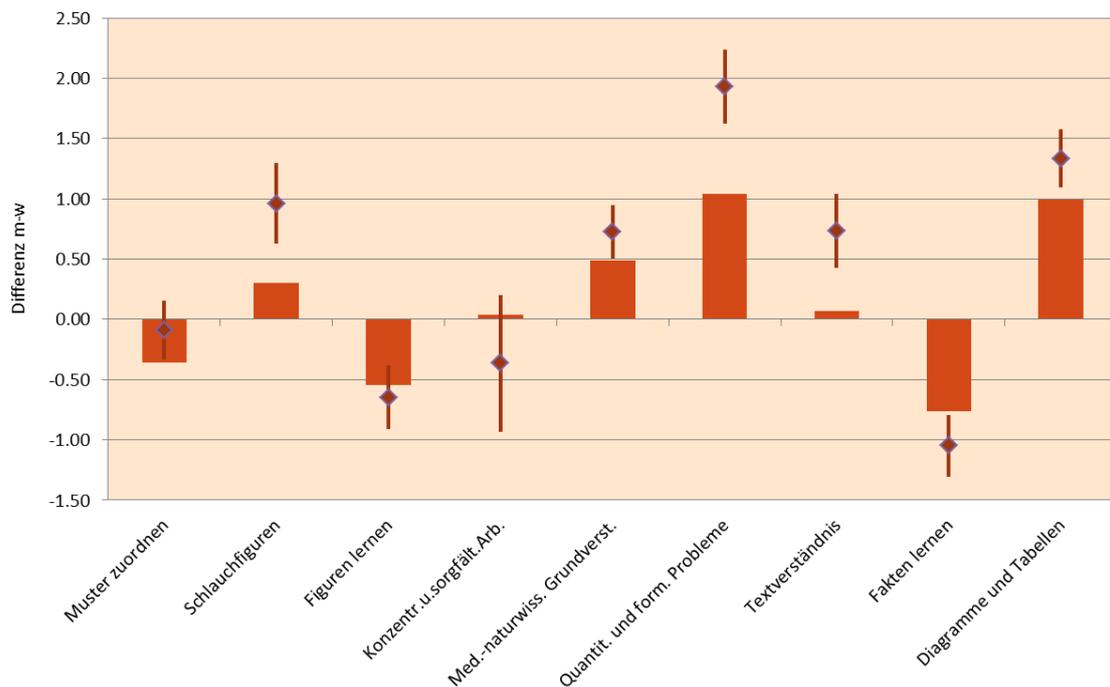


Abbildung 30: Geschlechtsspezifische Differenzen nach Aufgabengruppen (deutschsprachig, HM) mit Vergleich für die Vorjahre 1998-2016.

6.5 Vergleiche nach Wunschhochschulen

Die Tabelle 27 führt die Test-Prozentränge getrennt nach Disziplin und gewünschter Hochschule (erste Wahl) aller deutschsprachigen Kandidaten auf. Für die Bewerber „Humanmedizin“ sind signifikante Differenzen feststellbar.

Nach der Zulassung treten in keinen Disziplinen mehr signifikante Unterschiede auf.

Tabelle 27: Test-Prozentränge und Ergebnisse der Varianzanalyse pro Disziplin nach Wunschhochschulen (d).

	Wunschhochschule	n	Mittelwert	Standardabw.	Quadratsumme	Df	MQ	F	Sig.
HM	Bern	721	50.0	27.9	25809.7	6	4301.6	5.4	.000
	Basel	604	54.8	28.9					
	ETH	172	56.8	29.8					
	Freiburg	64	63.1	26.7					
	Zürich	1139	55.3	28.1					
	Zürich (LU)	39	48.3	28.6					
	Zürich (SG)	55	45.4	26.2					
VM	Bern	138	45.8	24.0	196.5	1	196.5	0.3	0.578
	Zürich	152	47.5	26.2					
ZM	Bern	50	38.1	22.2	1802.3	2	901.2	1.4	0.245
	Basel	53	38.9	25.8					
	Zürich	99	44.5	26.3					

Tabelle 28: Test-Prozentränge und Ergebnisse der Varianzanalyse pro Disziplin nach Wunschhochschulen, nur zugelassene, deutschsprachige Personen.

	Wunschhochschule	n	Mittelwert	Standardabw.	Quadratsumme	Df	MQ	F	Sig.
HM	Bern	208	85.3	8.8	198.2	6	33.0	0.4	.846
	Basel	219	86.2	8.6					
	ETH	69	86.9	8.0					
	Freiburg	31	86.5	7.2					
	Zürich	407	85.8	8.6					
	Zürich (LU)	11	85.3	8.8					
	Zürich (SG)	10	84.6	10.4					
VM	Bern	45	73.4	11.6	72.3	1	72.3	0.5	0.466
	Zürich	61	75.0	11.6					
ZM	Bern	26	56.0	13.9	1225.3	2	612.6	2.5	0.086
	Basel	26	61.7	15.9					
	Zürich	56	64.3	16.3					

7 Ergebnisse zur Testgüte

7.1 Zuverlässigkeit

Die zulassungsrelevante Reliabilität (Zuverlässigkeit) über den Gesamttest kann anhand zweier Koeffizienten berechnet werden. Die Reliabilitätsschätzung nach der Testhalbierungsmethode (Teilung nach gerad- und ungeradzahligem 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 der entsprechenden Aufgabengruppen.

Entscheidend sind die Kennwerte des Punktwertes, welcher nach der Standardisierung als Test-Prozentrang für die Zulassung verwendet wird.

Der Zuverlässigkeitswert des Punktwertes liegt 2017 weiterhin im gewohnt hohen Bereich der Vorjahre. Auch das Testprofil erreicht mit über 0.8 eine vergleichbare Konsistenz. Hierbei ist zu beachten, dass zu hohe Werte für sehr gleichartige Aufgabengruppen (mit der Frage, ob man einzelne weglassen kann) sprechen würden, sehr niedrige Werte für eine heterogene Testatterie, die nicht ohne weiteres zu einem Test-Prozentrang 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. Test-Prozentrang für die Eignungsmessung zu verwenden.

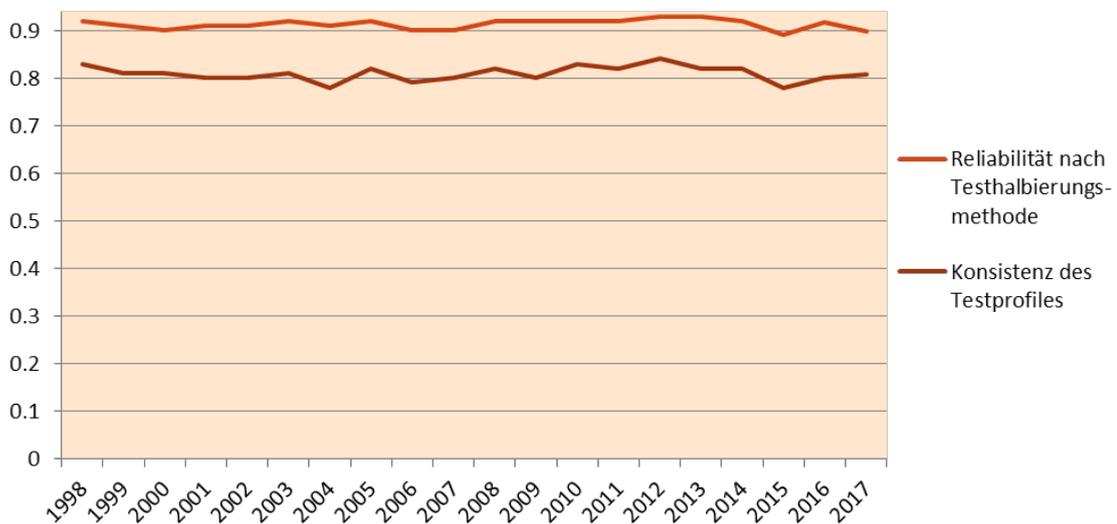


Abbildung 31: Reliabilität des Punktwertes (Split Half) und Konsistenz des Testprofils 1998 bis 2016.

Die Zuverlässigkeitswerte der einzelnen Aufgabengruppen sind für die Testverwendung nicht bedeutsam, da keine Entscheidungen auf der Basis einzelner Aufgabengruppen getroffen und diese nicht einzeln interpretiert werden. Hier müssten lediglich Eingriffe erfolgen, wenn der Test-Prozentrang selbst eine zu geringe Zuverlässigkeit aufweisen würde (was aber ausgeschlossen werden kann).

7.2 Binnenstruktur

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

Tabelle 29: Korrelationen zwischen Punktwerten der Aufgabengruppen und Gesamt-Punktwert.

	Med.-naturwis. Grundv.	Schlauchfiguren	Quant. u. formale Probl.	Textverständnis	Figuren lernen	Fakten lernen	Diagr. und Tabellen	Konzent. u. sorgf. Arbeiten	Punkt-wert
Muster zuordnen	.279	.548	.295	.203	.444	.395	.207	.334	.657
Med.-naturw. Grundverst.		.343	.459	.519	.262	.295	.452	.312	.677
Schlauchfiguren			.324	.242	.455	.407	.286	.334	.691
Quant. u. form. Probleme				.333	.226	.244	.458	.324	.631
Textverständnis					.219	.231	.372	.271	.567
Figuren lernen						.423	.185	.334	.633
Fakten lernen							.193	.241	.609
Diagramme und Tabellen								.279	.576
Konzent. u. sorgf. Arbeiten									.645

Eigenwerte und Varianzanteile der unrotierten Lösung für die neun Aufgabengruppen entsprechen weiterhin den bekannten Größenordnungen (Tabelle 30).

Tabelle 30: Faktorenanalyse: Verlauf der unrotierten Lösung (2013-2017).

Unrotierte Lösung															
Faktor	Eigenwert					% Varianz					Kumuliert %				
	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017
1	4.02	4.02	3.85	3.64	3.61	40.16	40.18	38.45	40.40	40.14	40.16	40.18	38.45	40.40	40.14
2	1.58	1.60	1.41	1.44	1.31	15.80	15.97	14.06	15.96	14.59	55.95	56.16	52.51	56.37	54.72
3	0.89	0.78	0.87	0.92	0.77	8.89	7.79	8.72	10.23	8.50	64.84	63.94	61.22	66.60	63.22
4	0.68	0.70	0.83	0.71	0.71	6.75	6.98	8.26	7.89	7.91	71.60	70.92	69.48	74.49	71.13
5	0.58	0.63	0.66	0.58	0.63	5.81	6.27	6.61	6.41	6.96	77.40	77.19	76.09	80.90	78.09
6	0.51	0.53	0.56	0.50	0.56	5.06	5.34	5.59	5.60	6.18	82.46	82.53	81.68	86.51	84.27
7	0.50	0.52	0.54	0.48	0.53	5.01	5.20	5.44	5.32	5.83	87.47	87.74	87.12	91.82	90.10
8	0.49	0.48	0.47	0.37	0.46	4.85	4.84	4.66	4.16	5.10	92.32	92.58	91.77	95.99	95.20
9	0.40	0.38	0.43	0.36	0.43	4.04	3.83	4.25	4.01	4.80	96.35	96.41	96.02	100	100.00
10	0.37	0.36	0.40	.*	.*	3.65	3.59	3.98	.*	.*	100	100	100	.*	.*

* Nur noch 9 Aufgabengruppen

Tabelle 31: Varianzanteile der 2-Faktorenlösungen 2013-2017 (rotiert).

Varimax-rotierte Lösungen														
Eigenwert					% Varianz					Kumuliert %				
2013	2014	2015	2016	2017	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017
3.2	3.1	3.2	2.8	2.5	32.2	30.9	31.7	30.9	28.0	32.2	30.9	31.7	30.9	28.0
2.4	2.5	2.1	2.3	2.4	23.7	25.3	20.8	25.4	26.7	55.9	56.2	52.5	56.4	54.7

Ein allgemeiner Fähigkeitsfaktor klärt weiterhin 40% der Varianz auf – es gibt quasi einen grossen „gemeinsamen Nenner“ in allen Aufgabengruppen. Mit der Zwei-Faktorenlösung können nach wie vor deutlich über 50% der Varianz aufgeklärt werden. 2017 tendiert Aufgabe „Konzentriertes und sorgfältiges Arbeiten“ wie in den Anfangsjahren zum Faktor „Werkzeugfunktionen“ – aufgrund der wechselnden Instruktionen kann diese Aufgabengruppe mehr Anteile des einen oder anderen Faktors enthalten.

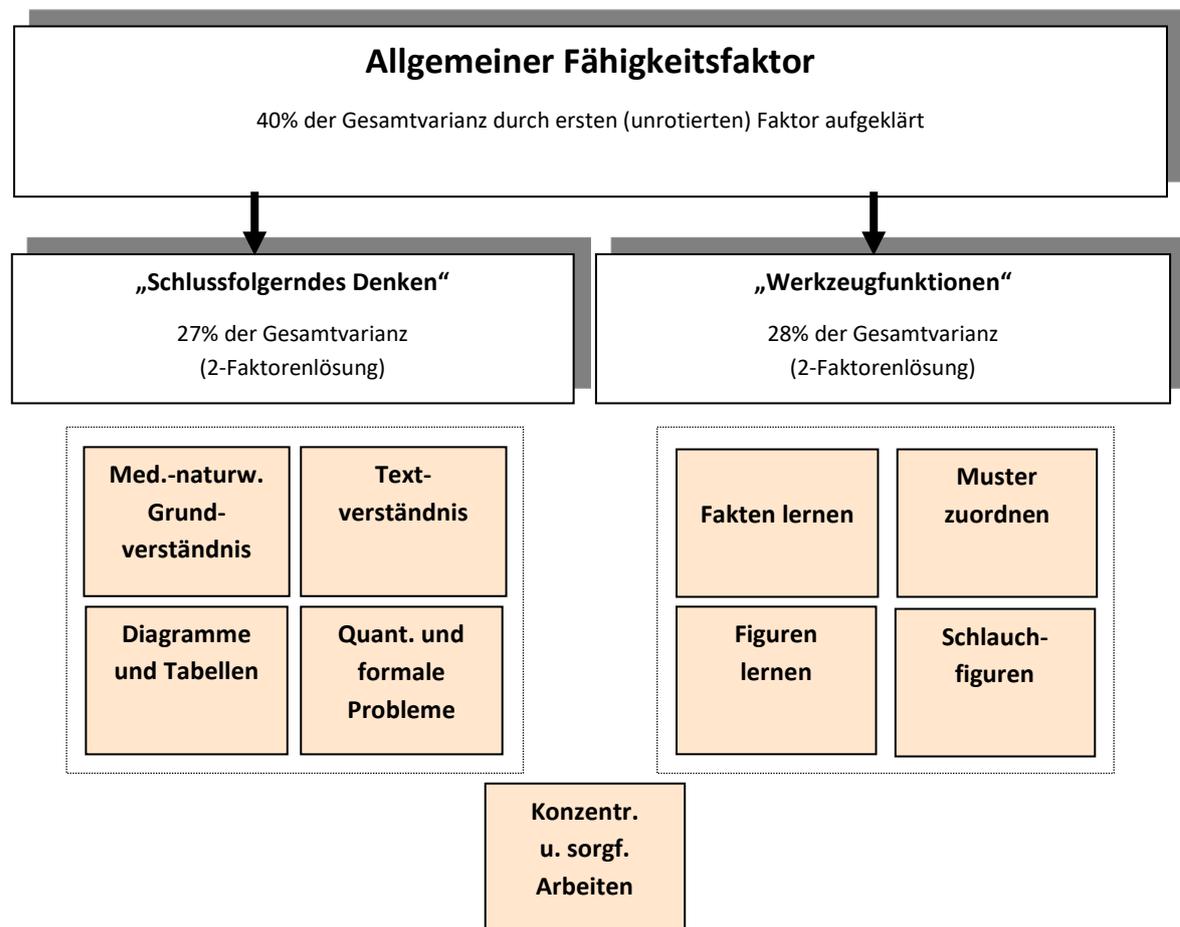


Abbildung 32: Struktur der Aufgabengruppen des EMS, hierarchische Faktorenstruktur 2017.

7.3 Konzentriertes und sorgfältiges Arbeiten

Diese Aufgabengruppe wird jährlich mit einer variierenden Instruktion durchgeführt, die vorher nicht bekannt ist.

Folgende Parameter sind (unabhängig von der konkreten Anforderung) relevant:

Richtige: Target richtig markiert (Wertebereich 0...400)

Fehler I: Target nicht markiert, „übersehen“ (Wertebereich 0...400)

Fehler II: Distraktor markiert, „falscher Alarm“ (Wertebereich 0...1200)

Der Punktwert (PW) berechnet sich wie folgt:

$$\text{PW} = \text{Richtige} - \text{Fehler I} - \text{Fehler II}$$

Werden mehr Fehler als Richtige Punkte erzielt, 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. Dieses Verfahren gewährleistet, dass auch unterschiedlich schwierige Konzentrationstests zu einer vergleichbaren Bepunktung über die einzelnen Jahre führen und dass der differenzierte Wertebereich nicht durch Ausreisser eingeschränkt wird.

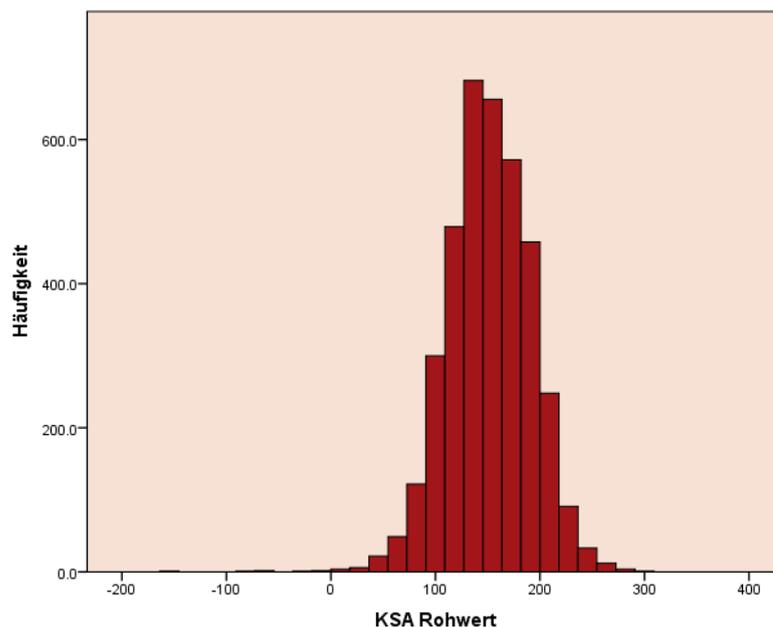


Abbildung 33: Verteilung des Rohwertes für „Konzentriertes und sorgfältiges Arbeiten“ 2017, negative Werte nicht Null gesetzt, Darstellung bis -200. 21 Personen erreichten Werte unter -200.

Die Verteilung der Rohwerte ist in Abbildung 33 dargestellt. Personen mit negativen Werten (mehr Fehler als Richtige) haben die Aufgabe nicht instruktionsgemäss bearbeitet. Insgesamt differenziert die Version 2017 im gesamten Wertebereich, niemand hat das Punktemaximum (400 Rohwertpunkte) erreicht. Auslassungsfehler sind deutlich häufiger als Fehlmarkierungen, im Durchschnitt wurde bis Zeile 22 (von 40) gearbeitet. Ein detaillierter Vergleich für die Sprachgruppen mit Differenzierung nach Fehlertyp ist in Tabelle 32 dargestellt.

Tabelle 32: Sprachspezifische Statistiken für Parameter der Aufgabengruppe „Konzentriertes und sorgfältiges Arbeiten“ und gesamt.

		Mittelwert	Standard- abweichung	Minimum	Maximum
Deutsch	Richtige	185.4	36.9	8	320
	Fehler	35.2	32.3	1	530
	... Auslassungen (I)	32.7	24.6	0	268
	... Fehlmarkierungen (II)	2.5	11.8	0	262
	letztes bearbeitetes Zeichen	868.4	175.7	353	1598
Französisch	Richtige	171.5	38.2	13	311
	Fehler	32.8	32.4	0	291
	... Auslassungen (I)	29.4	23.8	0	172
	... Fehlmarkierungen (II)	3.4	17.0	0	191
	letztes bearbeitetes Zeichen	799.5	183.0	346	1519
Italienisch	Richtige	175.4	40.6	21	258
	Fehler	39.1	41.5	2	371
	... Auslassungen (I)	34.4	24.9	2	204
	... Fehlmarkierungen (II)	4.7	19.7	0	167
	letztes bearbeitetes Zeichen	835.2	171.3	377	1235
Gesamt	Richtige	183.8	37.4	8	320
	Fehler	35.2	32.8	0	530
	... Auslassungen (I)	32.5	24.6	0	268
	... Fehlmarkierungen (II)	2.7	12.8	0	262
	letztes bearbeitetes Zeichen	861.3	177.1	346	1598

Abbildung 34 zeigt den Scatterplot für Mengenleistung und die Genauigkeit der Bearbeitung. Es besteht ein signifikanter, aber in der Ausprägung eher schwacher negativer Zusammenhang ($r=-.270$) zwischen Mengenleistung und Fehlerzahl (Tabelle 33).

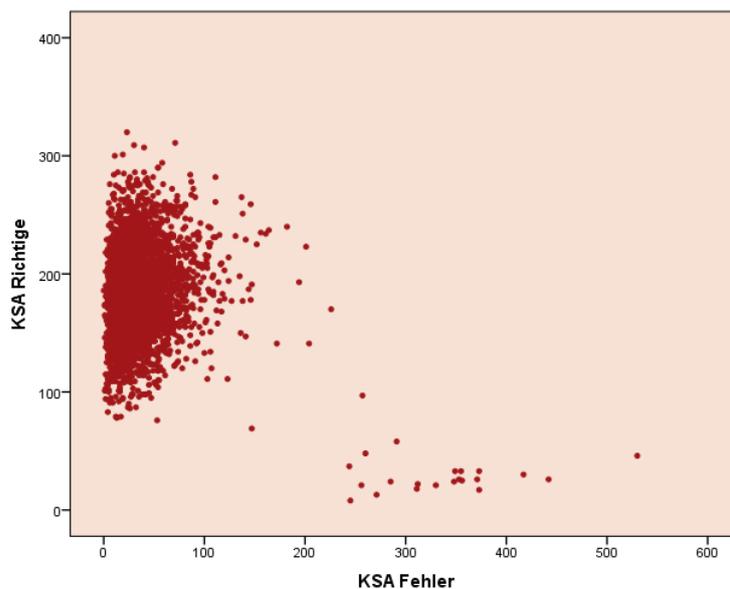


Abbildung 34: Scatterplot für alle Personen – Mengenleistung (Richtige) und Genauigkeit (Fehlerzahl).

Tabelle 33: Korrelationen der Parameter im Konzentrationstest

	Fehler gesamt	Auslassungen (I)	Fehlmarkierungen (II)
Richtige	-.124**	-.024	-.270**
Fehler gesamt		.940**	.757**
Auslassungen (I)	.940**		.088**

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

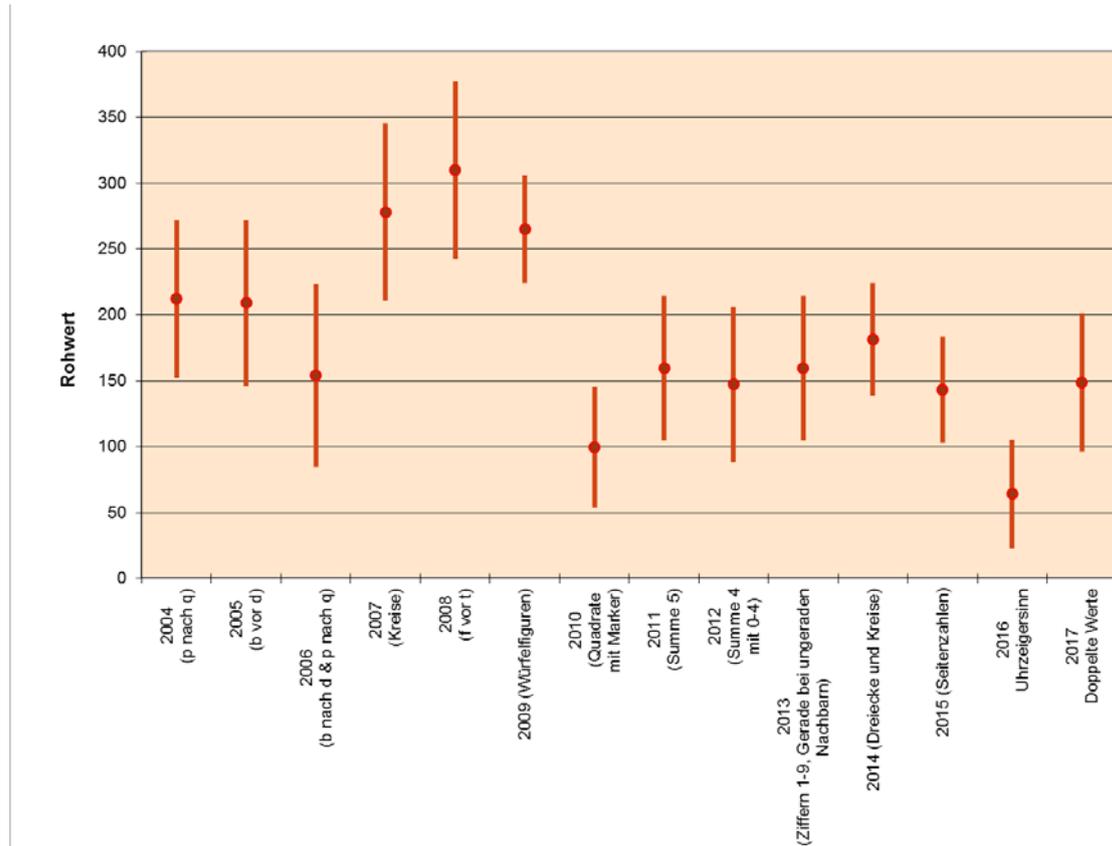


Abbildung 35: Leistungen „Konzentriertes und sorgfältiges Arbeiten“ (m±s) seit Einführung variierender Instruktionen und Zeichen (2004).

Trotz unterschiedlicher Schwierigkeiten der einzelnen Versionen dieser Aufgabe bleiben die Varianzen stabil, was eine Voraussetzung vergleichbarer Differenzierungsfähigkeit zwischen den Jahren ist. Die Version 2017 wies im Vergleich der bisherigen Versionen einen mittleren Schwierigkeitsgrad auf.

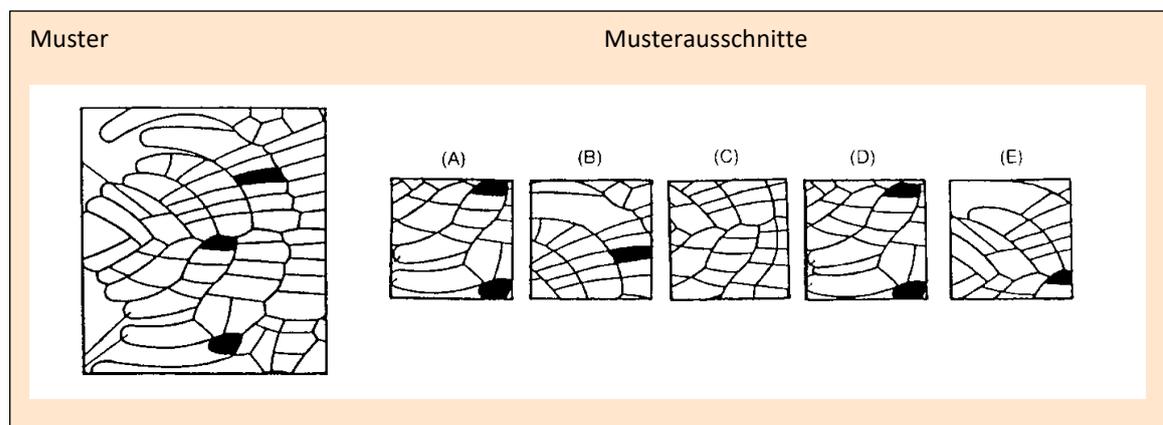
8 Beispielaufgaben für die Aufgabengruppen

Nachfolgend wird pro Aufgabengruppe eine Beispielaufgabe dargestellt. So können lediglich die Prinzipien der Aufgabenstruktur verdeutlicht werden. Die Aufgaben unterscheiden sich innerhalb jeder Aufgabengruppe bezüglich des Schwierigkeitsgrades und der Anforderung.

8.1 Muster zuordnen

In dieser Aufgabengruppe 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:



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 diese Aufgabengruppe 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.

8.2 Medizinisch-naturwissenschaftliches Grundverständnis

Hier wird das Verständnis für Fragen der Medizin und der Naturwissenschaften geprüft. 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 Aufgabengruppen 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.

8.3 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.

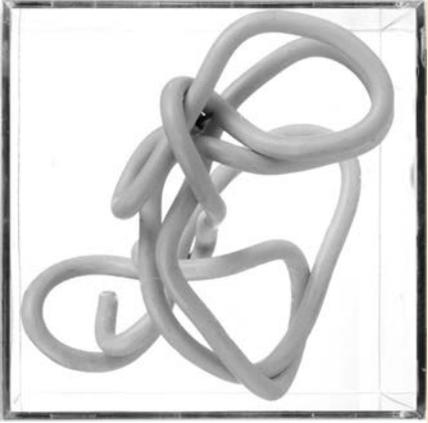
Jede Aufgabe besteht aus zwei Abbildungen eines durchsichtigen Würfels, in dem sich ein, zwei oder drei Kabel befinden. Die erste Abbildung (links) zeigt stets die Vorderansicht des Würfels; auf dem rechten Bild daneben, in welchem derselbe Würfel noch einmal abgebildet ist, soll die Testteilnehmerin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

1)



Hier sehen Sie den Würfel von vorne!

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



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

8.4 Quantitative und formale Probleme

Mit Hilfe dieses Aufgabengruppen 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.

8.5 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.

8.6 Textverständnis

Mit Hilfe dieser Aufgabengruppe 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 (T3) und Thyroxin (T4). 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 T3 und T4 als inaktive Speicherformen enthalten sind. Beim Menschen ist in den Follikeln so viel T3 und T4 gespeichert, dass der Organismus damit für etwa 10 Monate versorgt werden kann.

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

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

Die letzten Schritte der Hormonbildung finden in den Follikeln, also ausserhalb der einzelnen Epithelzellen, statt. In dort vorhandene sogenannte Tyrosin-Reste (des Thyreoglobulins) wird zunächst ein Jodatomb eingebaut. So entstehen Monojodtyrosin-Reste (MIT), von denen ein Teil durch die Bindung je eines weiteren Jodatoms in Dijodtyrosin-Reste (DIT) umgewandelt wird. Durch die Verknüpfung von je zwei DIT-Resten entsteht schliesslich T₄, während aus der Verbindung je eines MIT-Restes mit einem DIT-Rest T₃ hervorgeht. T₃ und T₄ werden dann in den Follikeln gespeichert und bei Bedarf über die Epithelzellen ins Blut freigesetzt.

Diese Freisetzung von T₃ und T₄ 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 T₃ und T₄; es ist hinsichtlich seiner eigenen Sekretionsrate jedoch abhängig von der Stimulation durch das hypothalamische Hormon TRH. Die TRH-Sekretion wiederum wird z.B. durch Kälte stimuliert, während Wärme hemmend wirken kann. Neben diesen übergeordneten Steuerungsmechanismen existiert noch ein sogenannter Rückkoppelungsmechanismus: Eine hohe Konzentration von T₃ und T₄ 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 T₃ und T₄ 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 T₃ und T₄ können ebenso wie TSH und TRH für diagnostische und therapeutische Zwecke synthetisch hergestellt werden.

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

Welcher der folgenden Vorgänge gehört nicht zu den im Text beschriebenen Schritten, die zur Bildung von T₃ führen?

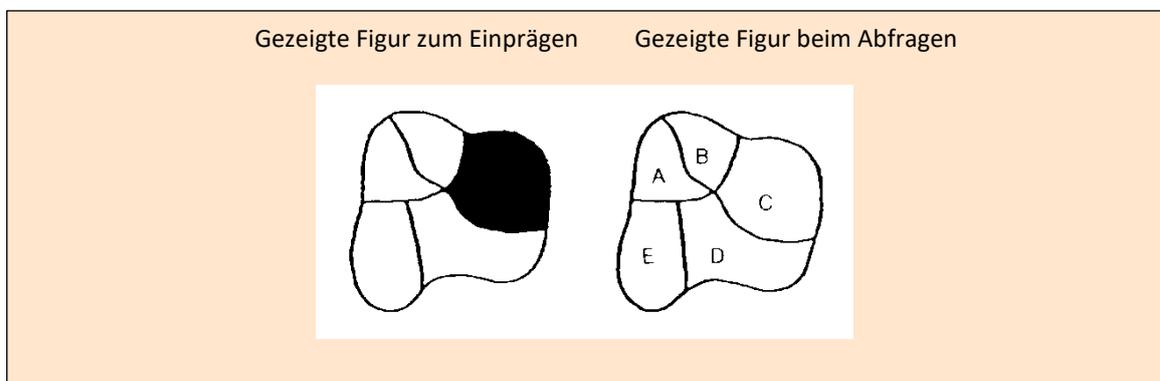
- (A) Transport von Jod aus den Epithelzellen in die Follikel
- (B) Umwandlung von Jod in Jodid in den Follikeln
- (C) Transport von Jodid aus dem Blut in die Epithelzellen
- (D) Verknüpfung von MIT- und DIT-Resten in den Follikeln
- (E) Verknüpfung von Jod und Tyrosin-Resten in den Follikeln

Für die Beantwortung dieser Frage ist das Verständnis der im obigen Text unterstrichenen Stellen wichtig (im Original sind selbstverständlich keine Hervorhebungen). Der Text sagt nichts über eine Umwandlung von Jod in Jodid in den Follikeln aus, und auch der umgekehrte Prozess, die Umwandlung von Jodid in Jod, findet nicht in den Follikeln statt, sondern in den Epithelzellen. Somit gehört der Vorgang (B) nicht zu den vom Text beschriebenen Schritten, die zur Bildung von T3 führen.

8.7 Figuren lernen

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

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



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.

8.8 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

8.9 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 Aufgabengruppen „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 Aufgabengruppen 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.

9 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: studienfeld-bezogene 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, (2014): Wollen wir nicht lieber diplomierte Ärzte? Sind fachliche und soziale Kompetenzen Gegensätze? *Schweizerische Ärztezeitung* 2014; 95: 6 S. 194 f.

Hänsgen, K.-D. (2013): Ohne Numerus clausus sinkt die Qualität des Studiums. Der externe Standpunkt, *NZZ* am Sonntag, 28.4. 2013, S. 17.

Hänsgen, K.-D (2007): Numerus clausus in der Medizin – werden die Richtigen ausgewählt für Studium und Beruf? *Schweizerische Ärztezeitung | Bulletin des médecins suisses | Bollettino dei medici svizzeri* | 2007; 88: 46.

Hänsgen, K.-D. (2007): Wie bilden wir die richtigen Studierenden zu Ärzten aus? *Neue Zürcher Zeitung*, 18. Oktober 2007.

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.

Hängsen, 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.

Hängsen, 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.

Hängsen, 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.

9.1 Originaltest zur Information und Vorbereitung

Es existieren in allen drei Sprachen je zwei veröffentlichte Originalversionen.

Deutsch: www.unifr.ch/ztd/ems/vord.htm

Französisch: www.unifr.ch/ztd/ems/vorf.htm

Italienisch: www.unifr.ch/ztd/ems/vori.htm

9.2 Frühere Berichte des ZTD

Über EMS-Seite: www.unifr.ch/ztd/ems
Im Menü rechts inklusive aller Berichte und FAQ

EMS-WIKI-Blog www.ztd.ch
Kommentare und Fakten