

Eignungstest für das Medizinstudium in der Schweiz

Bericht über Durchführung und Ergebnisse

B. Spicher

Dr. Spicher, Benjamin (2020)
EMS Eignungstest für das Medizinstudium 2020
Berichte des Zentrums für Testentwicklung, Nr. 27,
im Auftrag der Schweizerischen Hochschulkonferenz (SHK)

Unter Mitarbeit von M. Bernasconi,
T. Cruchaud, 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:

M. Klaus

Redaktion: T. Cruchaud



1 Inhaltsverzeichnis

2	Zusammenfassung	6
3	Résumé	7
4	Numerus clausus (NC) und Medizinstudium	8
5	Anmeldung zum Medizinstudium und Test 2020	10
5.1	Anmeldestatistik Humanmedizin	10
5.2	Anmeldestatistik Veterinärmedizin	12
5.3	Anmeldestatistik Zahnmedizin	14
5.4	Testorte und Wunschhochschulen	17
5.5	Wunschhochschule und Testort nach Wohnkanton	18
5.6	Teilnahmen am EMS (NC) nach Kantonen seit 1998	22
5.7	Testabsolvierung nach Geschlecht	23
5.8	Übernahme des Testergebnisses aus dem Vorjahr	25
5.9	Sprachgruppen	26
6	Der EMS im Corona-Jahr	27
6.1	Distanzen	27
6.2	Schutzmasken und Desinfektionsmittel	27
6.3	Einlass	27
6.4	Versetzter Testbeginn in Zürich	27
6.5	Mittagspause	27
6.6	Austeilen von Materialien	28
6.7	Entlassung der Teilnehmenden	28
7	Beschreibung des verwendeten Eignungstests	29
7.1	Berechnung der Werte 2020	31
7.2	Mittlerer Rangplatz der Aufgabengruppen	32
8	Testanwendung 2020	34
8.1	Verteilungsprüfung	34
8.2	Äquivalenz der Sprachversionen	35
8.2.1	<i>Sprachvergleich für die Aufgabengruppen</i>	<i>37</i>

8.2.2	<i>Darstellung des Korrekturverfahrens</i>	39
8.2.3	<i>Identifikation von DIF-Aufgaben und Bestimmung des Korrekturwertes</i>	40
8.2.4	<i>Effekte der Korrektur</i>	45
8.3	Vergleichbarkeit der Testlokale	48
8.3.1	<i>Versetzter Testbeginn in Zürich</i>	50
8.4	Vergleich für die Geschlechter.....	51
8.5	Vergleiche nach Wunschhochschulen	53
9	Ergebnisse zur Testgüte	54
9.1	Zuverlässigkeit	54
9.2	Binnenstruktur	55
10	Beispielaufgaben für die Aufgabengruppen	57
10.1	Muster zuordnen	57
10.2	Medizinisch-naturwissenschaftliches Grundverständnis	57
10.3	Schlauchfiguren	59
10.4	Quantitative und formale Probleme.....	59
10.5	Konzentriertes und sorgfältiges Arbeiten.....	60
10.6	Textverständnis	60
10.7	Figuren lernen.....	62
10.8	Fakten lernen.....	62
10.9	Diagramme und Tabellen.....	63
11	Literatur	64
11.1	Originaltest zur Information und Vorbereitung.....	65
11.2	Frühere Berichte des ZTD	65

2 Zusammenfassung

Die in diesem Bericht vorgestellten Ergebnisse betreffen den Eignungstest 2020 für das Medizinstudium in der Schweiz (EMS) und die Zulassung zum Studium der Medizin ab Herbstsemester 2020. Für bestimmte Fragestellungen werden Vergleichsdaten früherer Testanwendungen herangezogen.

1. Ein Numerus clausus (NC) war 2020 für Humanmedizin, Veterinärmedizin und Zahnmedizin notwendig für Personen, die sich in Basel, Bern, Freiburg oder Zürich angemeldet haben - auch für den Studiengang 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 weiterhin auch für Chiropraktik (seit 2008).
2. Die Nachfrage nach Studienplätzen für **Humanmedizin** nahm, gemessen an den Anmeldungen im Februar, geringfügig ab. Die entsprechenden (stabilen) Kapazitäten werden in dieser Disziplin zu **375%** (Vorjahr 377%) ausgelastet, in **Veterinärmedizin** zu **330%** (Vorjahr 333%), in **Zahnmedizin** zu **245%** (Vorjahr 257%).
3. **4209 Personen** (Vorjahr 4095) haben sich für die betreffenden Hochschulen bis zur gesetzten Frist im **Mai 2020** bei swissuniversities zum EMS 2020 angemeldet. Davon bewarben sich **15 Personen** um einen Disziplinenwechsel oder Einstieg in einem höheren Semester. Weitere **99 Personen** bewarben sich mit ihrem Testergebnis aus dem Vorjahr und nahmen nicht am diesjährigen Test teil.
4. **4058 Personen** haben den EMS mit gültigem Ergebnis beendet (Vorjahr 3887). 5 Personen haben den Test vorzeitig abgebrochen (Vorjahr 4). 146 Personen sind nicht zum Test erschienen oder haben ihre Testanmeldung zurückgezogen (Vorjahr 204). Die Rückzugsquoten zwischen Februar und Test sanken für Humanmedizin auf **17.4%**, für Veterinärmedizin auf **14.4%** und für Zahnmedizin auf **15%**.
5. Der Test fand am 3. Juli 2020 gleichzeitig an **8 Testorten** und in 34 Testlokalen in drei Sprachen (Deutsch: Basel, Bern, Chur, Luzern, St. Gallen, Zürich; Französisch: Freiburg; Italienisch: Lugano) statt. Die Wahl des Testortes war unabhängig vom zukünftigen Studienort möglich.
6. Aufgrund der **Corona-Pandemie** wurde der EMS 2020 um drei Aufgabengruppen gekürzt und ohne Mittagspause durchgeführt. Diverse weitere Massnahmen stellten sicher, dass die Vorgaben des Bundes bei der Durchführung eingehalten werden konnten.

3 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 2020. Pour certaines problématiques, des données comparatives des éditions antérieures du test (1998-2019) y figurent également.

1. En 2020, un numerus clausus (NC) a été nécessaire à nouveau en médecine humaine, en médecine vétérinaire et en médecine dentaire pour les personnes inscrites à Bâle, Berne, Fribourg ou Zurich. En médecine humaine, cela concerne également les nouvelles filières avec NC à l'EPF Zurich, aux Universités de Lucerne et de St-Gall en coopération avec l'Université de Zurich ainsi qu'à l'Università della Svizzera italiana dont le cursus commence à Bâle. À l'Université de Zurich, le NC se poursuit pour la chiropractie (depuis 2008).
2. La demande de places d'études en **médecine humaine** a légèrement diminué, sur la base du nombre d'inscriptions reçues en février. Les capacités correspondantes (qui restent stables) sont dépassées dans cette discipline à hauteur de **375 %** (377 % l'année précédente), celles en **médecine vétérinaire** à hauteur de **330 %** (333 % en 2019) et en **médecine dentaire** à hauteur de **245 %** (257 % en 2019).
3. **4209 personnes** (4095 en 2019) se sont annoncées pour ces hautes écoles auprès de swissuniversities jusqu'au délai imparti de **mai 2020**, en vue de la passation du test AMS 2020. Parmi elles, **15 personnes** ont sollicité un changement de discipline ou un passage à un semestre supérieur et **99 personnes** ont fait valoir leur résultat de l'année précédente et n'ont ainsi pas pris part au test 2020.
4. **4058 personnes** ont terminé l'AMS avec des résultats valables (3887 en 2019). 5 personnes ont interrompu prématurément le test (4 en 2019). 146 personnes ont retiré leur inscription au test ou ne se sont pas présentées à celui-ci (204 en 2019). Les taux de retrait entre février et la date de passation du test sont tombés à **17,4 %** pour la médecine humaine, **14,4 %** pour la médecine vétérinaire et **15 %** pour la médecine dentaire.
5. Le test s'est déroulé le 3 juillet 2020, simultanément **sur 8 sites** différents comprenant au total 34 locaux de test et ce, dans trois langues (allemand : 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.
6. En raison de la **pandémie du coronavirus**, l'AMS 2020 a été raccourci de trois groupes d'exercices et il a été réalisé sans pause de midi. Diverses autres mesures ont permis de garantir que les directives du Conseil fédéral soient respectées lors de sa mise en œuvre.

4 Numerus clausus (NC) und Medizinstudium

Die Nachfrage nach Studienplätzen an Universitäten mit NC nahm 2020 gegenüber dem Vorjahr geringfügig ab und scheint sich auf hohem Niveau stabilisiert zu haben. Demgegenüber stieg die Nachfrage an Universitäten ohne NC (2020 insbesondere Genf) zum zweiten Mal in Folge deutlich an.

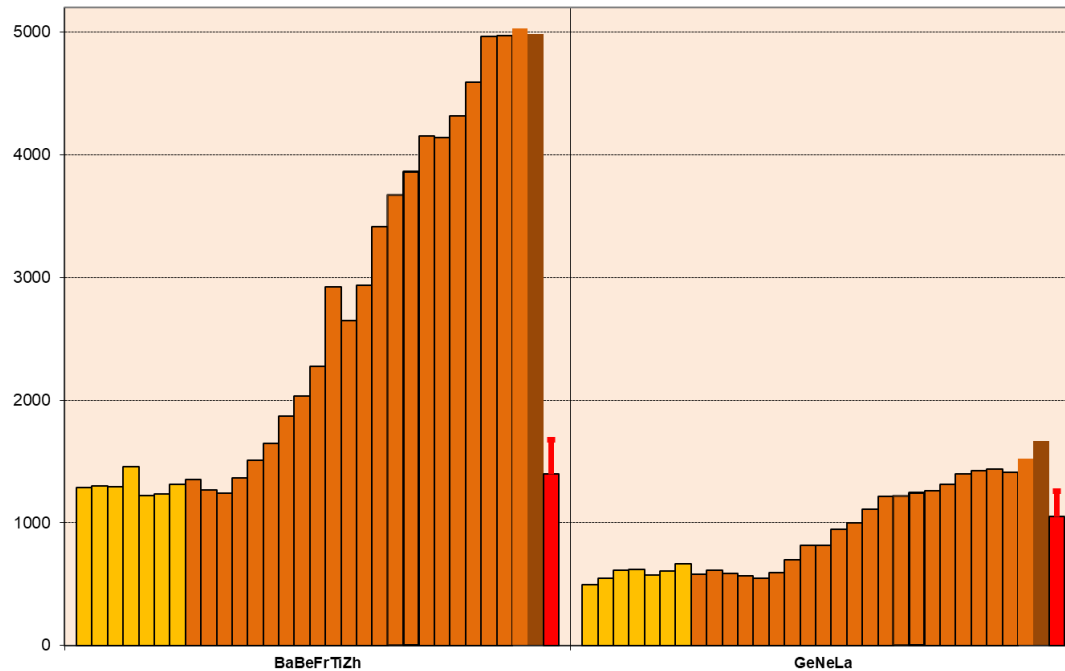


Abbildung 1: Anmeldungen zum Medizinstudium 1991 bis 2020 nach Gruppen (mit NC vs. ohne NC). Hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule (rot): Kapazität 2020 mit 20%-Überschreitungskriterium als Grenze für Auslösung des NC.

Auf jeden Studienplatz in Humanmedizin bewerben sich an den Universitäten mit NC 3.75 Interessierte, in der Veterinärmedizin sind es 3.3, in der Zahnmedizin 2.45.

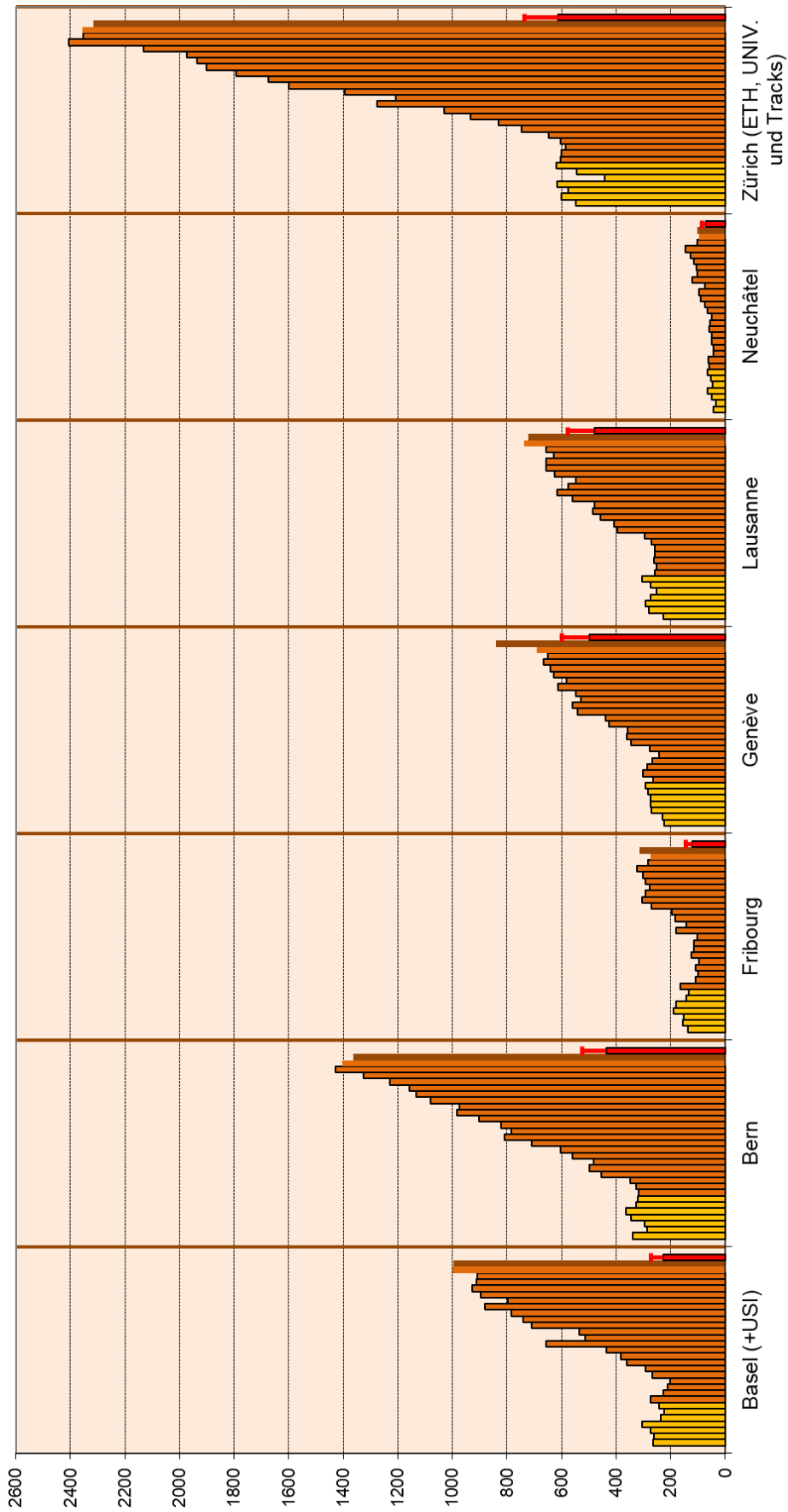


Abbildung 2: Anmeldezahlen pro Hochschule 1991 bis 2020. Hellere Säulen: Jahre ohne NC; dunklere Säulen: Jahre mit NC; letzte Säule (rot): Kapazität 2020 mit 20%-Überschreitungskriterium.

5 Anmeldung zum Medizinstudium und Test 2020

5.1 Anmeldestatistik Humanmedizin

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

	Humanmedizin (seit 2008 inkl. Chiropraktik)												
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Kapazitäten (HS mit NC)	603	603	653	653	653	713	793	793	793	997	1107	1107	1107
Anmeldungen im Februar	2029	2324	2651	2936	3120	3270	3310	3491	3754	4109	4111	4178	4152
... in % zu Kapazität	336	385	406	450	478	459	417	440	473	410	371	377	375
Anmeldungen zum EMS Mai	1612	1765	2109	2295	2461	2547	2653	2771	3000	3250	3273	3405	3499
Absolvierung EMS	1535	1664	2016	2172	2337	2403	2525	2659	2864	3122	3145	3234	3370
Übernahme Ergebnis Vorjahr	52	66	54	65	54	60	42	59	70	76	72	51	73
Bewerbungen Juli	1576	1707	2040	2212	2363	2426	2533	2684	2916	3186	3202	3281	3428
Rückzugsquote zwischen Februar und Juli (%)	22.3	26.6	23.0	24.6	24.3	25.8	23.5	23.1	22.3	22.5	22.1	21.5	17.4
Zugeteilte Studienplätze mit „Überbuchung“	726	707	763	745	745	805	886	877	881	1091	1211	1198	1201^m
Abgewiesene Bewerbungen	850	1000	1277	1467	1618	1621	1647	1807	2035	2095	1991	2083	2228^m
% Bewerbungen, die Studienplatz erhalten	45	41	37	34	32	33	35	33	30	34	38	37	35^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 Nachrückende möglichst vermieden werden können. Der Anteil der Rückzüge zwischen Februar (Anmeldung Studium) und der Testteilnahme ging 2020 stark zurück (auf 17.4%), so dass trotz einer geringeren Anzahl Anmeldungen zum Medizinstudium ein neuer Höchststand hinsichtlich der Bewerbungen im Juli (also mit gültigem Testergebnis) erreicht wurde.

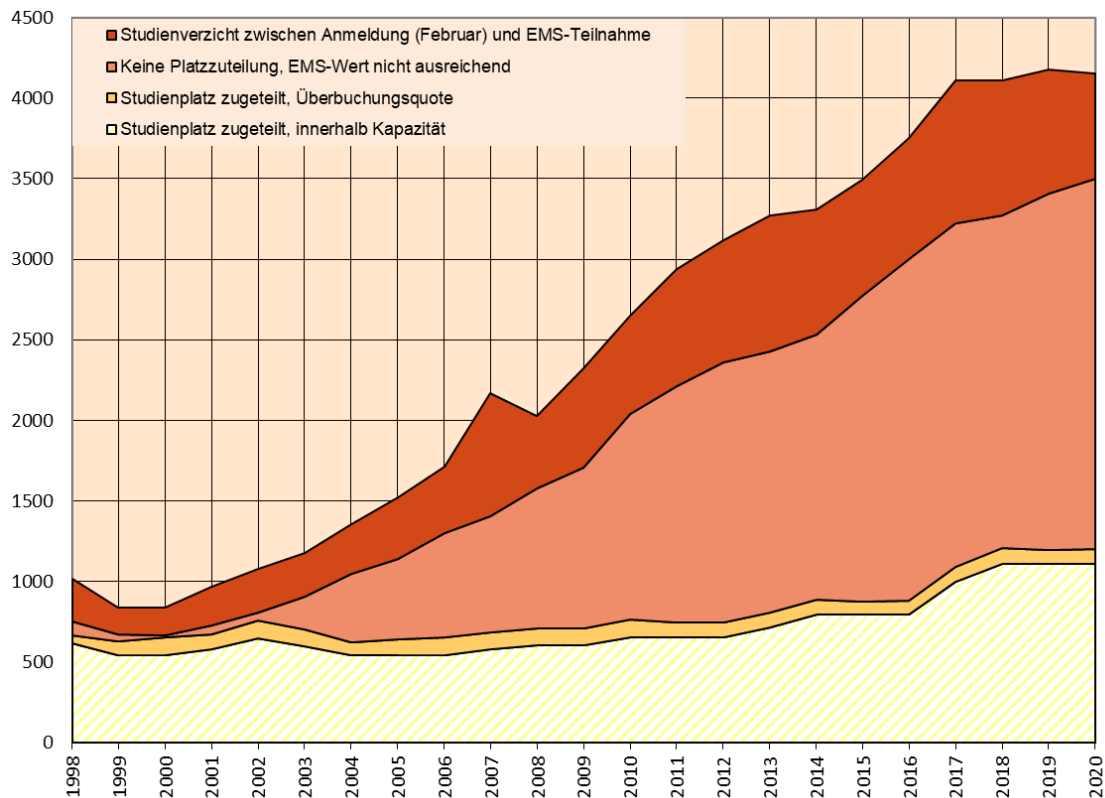


Abbildung 3: Anmeldeverlauf Humanmedizin.

Die bis 2017 fast stetig zunehmenden Anmeldezahlen zum Medizinstudium haben sich seit drei Jahren auf hohem Niveau stabilisiert. Durch den deutlich verringerten Anteil an Personen, die ihre Anmeldung vor dem Testtag zurückzogen (oder nicht zum Test erschienen sind) ist ein Anstieg der aufgrund des Testergebnisses nicht zugelassenen Personen erkennbar. Die Zulassungsquote liegt dementsprechend leicht tiefer als im Vorjahr, allerdings nach den massiven Erhöhungen der Studienplätze in den Jahren 2017 und 2018 weiterhin im Bereich der Jahre 2010/2011. Entsprechend bleibt der Numerus clausus im Verhältnis zu anderen Ländern 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/innen ist es im Vergleich zu vielen anderen Ländern nach wie vor wahrscheinlicher, zum Medizinstudium zugelassen zu werden.

5.2 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												
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Kapazitäten (HS mit NC)	150	150	150	150	150	150	150	150	150	150	150	166	166
Anmeldungen im Februar	365	371	432	408	406	478	487	476	496	566	552	553	548
... in % zu Kapazität	243	247	288	272	271	319	325	317	331	377	368	333	330
Anmeldungen zum EMS Mai	290	294	360	308	327	365	403	404	427	459	455	455	475
Absolvierung EMS	272	281	342	295	317	350	384	375	408	440	439	425	461
Übernahme Ergebnis Vorjahr	15	16	11	16	8	24	17	8	8	16	5	10	8
Bewerbungen Juli	287	297	353	311	325	374	401	383	416	454	444	435	469
Rückzugsquote zwischen Februar und Juli (%)	21.3	19.9	18.0	23.8	20.0	21.8	17.7	19.5	16.1	19.8	19.6	21.3	14.4
Zugeteilte Studienplätze mit „Überbuchung“	173	173	173	173	173	173	173	173	167	167	163	180	180^m
Abgewiesene Bewerbungen	114	124	180	138	152	201	228	210	249	287	281	255	289^m
% Bewerbungen, die Studienplatz erhalten	60	58	49	56	53	46	43	45	40	37	37	41	38^m

^m Stand Oktober gemäss Modell.

Im Fach Veterinärmedizin blieb die Zahl der Anmeldungen bis zum Februar stabil, auch hier sank jedoch die Rückzugsquote zwischen Februar und Test deutlich ab. Auch in der Veterinärmedizin hat die Anzahl gültiger Testergebnisse 2020 einen Höchststand erreicht – die Zulassungsquote sank auf 38%.

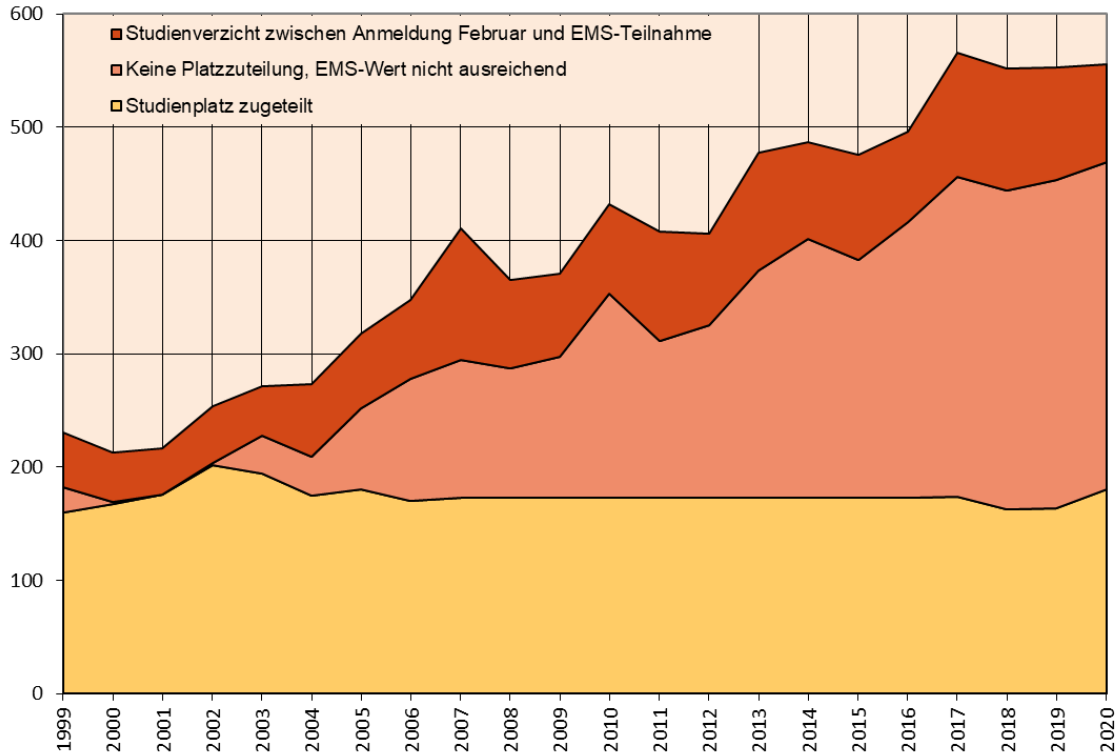


Abbildung 4: Anmeldeverlauf für Veterinärmedizin 1999 bis 2020 im Vergleich (Univ. mit NC).

5.3 Anmeldestatistik Zahnmedizin

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

	Zahnmedizin												
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Kapazitäten (Univ. mit NC)	142	142	142	142	139	139	131	131	131	117	117	117	117
Anmeldungen im Februar	197	239	315	324	332	396	343	352	339	292	311	301	287
... in % zu Kapazität	138	168	222	228	239	285	262	269	259	250	266	257	245
Anmeldungen zum EMS Mai	164	181	233	250	257	284	273	275	267	221	236	257	235
Absolvierung EMS	154	174	220	239	248	275	264	258	251	205	224	228	227
Übernahme Ergebnis Vorjahr	13	31	37	37	25	39	26	19	18	18	16	13	18
Bewerbungen Juli	167	205	257	276	273	313	290	277	269	223	240	241	244
Rückzugsquote zwischen Februar und Juli (%)	15.2	14.3	18.4	14.8	17.8	21.0	15.5	21.3	20.6	23.6	22.8	19.9	15.0
Zugeteilte Studienplätze mit „Überbuchung“	167	170	165	159	159	161	152	150	150	127	131	131	131^m
Abgewiesene Bewerbungen	0	35	92	117	114	152	138	127	119	96	109	110	113^m
% Bewerbungen, die Studienplatz erhalten	99	83	64	58	58	51	52	54	56	57	55	54	54^m

^m Stand Oktober gemäss Modell.

In der Zahnmedizin verringerte sich die Rückzugsquote zwischen der Anmeldung im Februar und dem Testtag weiter von rund 20% aus dem Vorjahr auf 15%. Trotz etwas geringerer Anmeldezahlen im Februar resultierten dadurch eine praktisch identische Bewerbungszahl und Zulassungsquote wie im Jahr 2019.

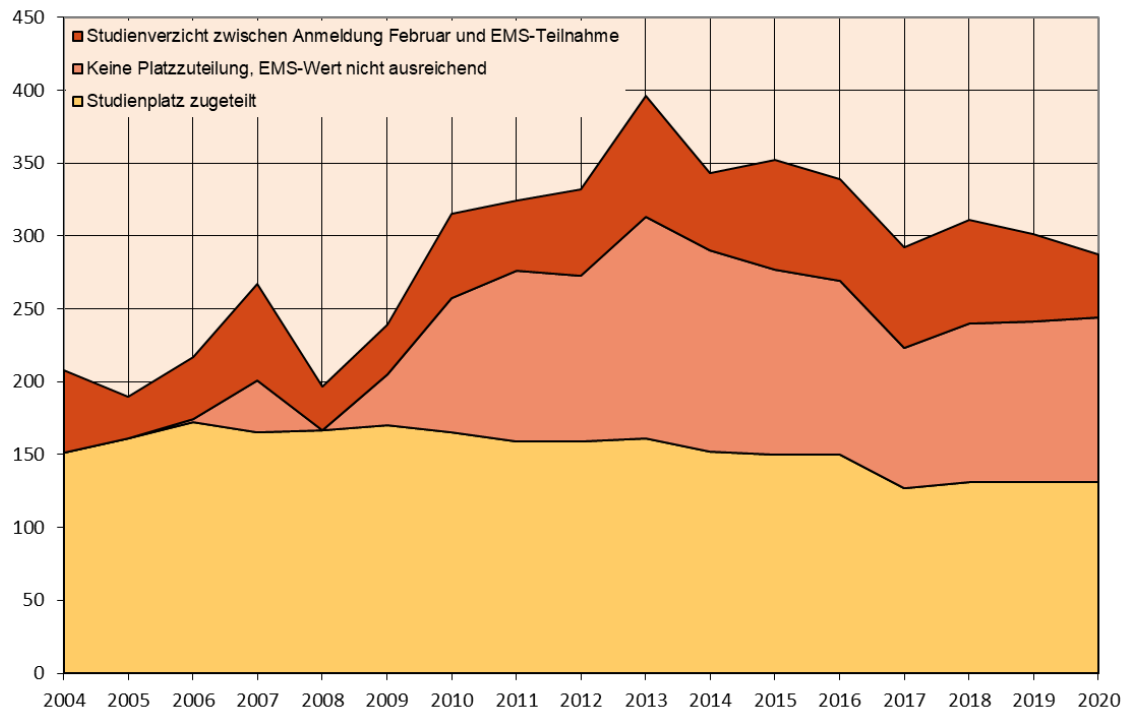


Abbildung 5: Anmeldeverlauf für Zahnmedizin 2004 bis 2020 (Univ. mit NC).

Seit Jahren ist die Durchlässigkeit zwischen Zahn- und Humanmedizin stark eingeschränkt. Dadurch wird verhindert, dass die Zulassung zum Studium der Zahnmedizin als „Hintereingang“ zur Humanmedizin missbraucht wird, und sich nur Personen zum Zahnmedizinstudium bewerben, die auch tatsächlich Interesse für diese Disziplin aufweisen. Für einen Wechsel zur Humanmedizin bedarf es weiterhin folgender 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 ein Drittel der Kandidaten und Kandidatinnen 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
2020	14* (5)	4 (1)	2 (0)	4 (2)	4 (2)
2019	4 (3)				4 (3)
2018	15 (5)	4 (1)	2 (1)	3 (1)	6 (2)
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 1 (0) Person mit Wechselwunsch nach Zahnmedizin (BS) .

** 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.

2020 traten 14 Personen mit Wechselwunsch von Zahn- zu Humanmedizin zum EMS an. Fünf davon hätten das Zulassungskriterium erreicht – ein tatsächlicher Wechsel hängt von den freien Kapazitäten im betroffenen Jahrgang „Humanmedizin“ ab.

Der Verlauf der Zulassungsquoten für alle Disziplinen über die Jahre ist in Abbildung 6 dargestellt.

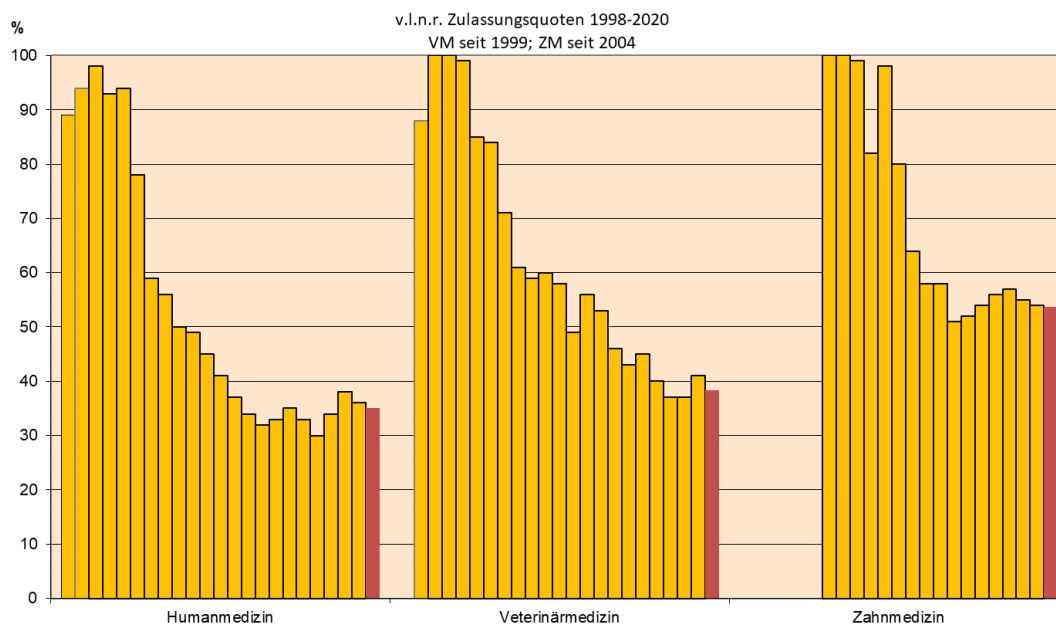


Abbildung 6: Anteil Personen mit gültigem Test-Prozentrang, denen ein Studienplatz zugewiesen werden konnte.

5.4 Testorte und Wunschhochschulen

Tabelle 5: Testorte und Wunschhochschulen (erste Wahl) nach Disziplinen

Disziplin	Testort	Wunschhochschulen								Gesamt
		BS	BE	ETH	FR	USI	ZH	ZH (LU)	ZH (SG)	
Human- medizin	Basel	501	20	14	2	1	21	0	1	560
	Bern	22	521	7	35	0	22	0	0	607
	Chur	21	39	7	2	0	51	1	5	126
	Freiburg	4	8	1	172	0	4	0	1	190
	Lugano	37	10	7	30	7	37	0	0	128
	Luzern	56	121	15	8	0	70	39	3	312
	St. Gallen	44	41	17	4	0	79	0	64	249
	Zürich	78	65	127	6	2	896	11	13	1198
	Gesamt	763	825	195	259	10	1180	51	87	3370
Veterinär- medizin	Basel		19				21			40
	Bern		73				3			76
	Chur		8				14			22
	Freiburg		120				6			126
	Lugano		7				8			15
	Luzern		19				19			38
	St. Gallen		2				28			30
	Zürich		2				112			114
	Gesamt		250				211			461
Zahnmedizin	Basel	32	0				0			32
	Bern	5	49				2			56
	Chur	0	0				1			1
	Freiburg	0	2				1			3
	Lugano	2	0				5			7
	Luzern	6	9				13			28
	St. Gallen	2	2				15			19
	Zürich	9	5				67			81
	Gesamt	56	67				104			227
Gesamt	Basel	533	39	14	2	1	42	0	1	632
	Bern	27	643	7	35	0	27	0	0	739
	Chur	21	47	7	2	0	66	1	5	149
	Freiburg	4	130	1	172	0	11	0	1	319
	Lugano	39	17	7	30	7	50	0	0	150
	Luzern	62	149	15	8	0	102	39	3	378
	St. Gallen	46	45	17	4	0	122	0	64	298
	Zürich	87	72	127	6	2	1075	11	13	1393
Gesamt	819	1142	195	259	10	1495	51	87	4058	

Die Testorte entsprechen in der Mehrheit dem Standort der gewünschten Hochschule. Für die Zuteilung der Studienplätze ist der Ort der Testabsolvierung nicht relevant. In Freiburg wird der EMS in der französischen Version durchgeführt, in Lugano in der italienischen.

5.5 Wunschhochschule und Testort nach Wohnkanton

Die folgenden Tabellen zeigen die Wunschhochschulen nach Wohnkanton für die einzelnen Disziplinen. Angegeben sind auch Vergleichswerte seit 2014.

Tabelle 6: Gewünschte Hochschulen Humanmedizin nach Wohnkantonen 2020 und Vergleich mit 2014 bis 2019.

Wohnkanton/ Wohnort	Humanmedizin 2020 gewünschte Hochschule									Total Vorjahre					
	BS	BE	ETH	FR	USI	ZH	ZH (LU)	ZH (SG)	Total	2019	2018	2017	2016	2015	2014
AG	124	54	26	5	0	101	10	0	320	282	286	257	252	226	226
AI	0	5	0	0	0	1	0	1	7	6	3	7	9	10	8
AR	8	6	2	0	0	3	0	5	24	27	29	29	25	24	23
BE	9	405	6	24	0	15	0	0	459	480	486	475	389	372	370
BL	228	6	6	1	0	7	0	1	249	204	172	166	169	173	129
BS	144	4	7	0	1	5	0	0	161	184	139	139	150	148	126
FR	2	20	1	96	0	2	0	0	121	102	100	111	101	94	106
GE	0	1	0	8	0	1	0	0	10	7	18	12	14	12	3
GL	1	0	0	0	0	9	0	0	10	14	8	15	11	7	12
GR	13	34	2	2	2	34	1	3	91	81	78	81	70	60	71
JU	3	0	0	6	0	0	0	0	9	16	9	18	16	8	5
LU	32	70	6	2	0	28	19	0	157	180	189	196	193	192	162
NE	0	0	0	10	0	1	0	0	11	3	5	8	7	7	7
NW	7	6	2	2	0	5	4	0	26	17	15	20	22	8	12
OW	1	8	0	0	0	4	2	0	15	16	9	14	12	13	14
SG	34	27	11	5	0	59	0	56	192	190	208	199	186	155	167
SH	5	3	0	0	0	20	0	0	28	25	29	24	13	26	26
SO	43	66	2	0	0	16	0	0	127	110	125	105	102	85	98
SZ	6	9	2	1	0	43	4	1	66	62	66	69	60	62	53
TG	12	10	12	1	0	45	0	13	93	101	98	78	81	68	57
TI	37	9	7	29	5	39	0	0	126	119	122	141	116	94	93
UR	0	4	0	0	0	3	1	0	8	10	11	15	10	5	9
VD	0	2	0	20	0	2	0	1	25	20	30	33	21	26	20
VS	9	15	2	37	0	3	0	0	66	56	62	57	61	43	30
ZG	12	23	9	2	0	24	6	1	77	69	63	68	54	55	51
ZH	12	27	86	3	2	685	4	4	823	787	723	724	657	630	586
FL*	0	4	3	0	0	9	0	0	16	20	21	9	10	11	14
Ausland	21	7	3	5	0	16	0	1	53	46	41	40	53	45	47
Total	763	825	195	259	10	1180	51	87	3370	3234	3145	3110	2864	2659	2525

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

Tabelle 7: Gewünschte Hochschulen Veterinärmedizin nach Wohnkantonen 2020 und Vergleich mit 2010 bis 2019.

Wohnkanton/ Wohnort	Veterinärmedizin 2020 gewünschte Hochschule			Total Vorjahre									
	BE	ZH	Total	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010
AG	5	27	32	29	27	26	27	23	24	24	21	21	24
AI	0	0	0	2	1	1	2	1	1	0	1	0	0
AR	0	2	2	2	1	3	3	4	4	5	2	2	1
BE	56	1	57	62	83	61	59	59	55	51	45	41	62
BL	9	12	21	15	14	15	12	20	17	15	17	10	14
BS	8	2	10	3	6	12	9	10	7	5	5	4	7
FR	24	0	24	28	23	19	17	12	17	26	10	17	17
GE	16	4	20	19	27	16	17	21	18	12	12	12	16
GL	1	6	7	2	3	4	7	4	5	4	2	3	2
GR	5	7	12	12	3	10	11	15	11	12	12	14	10
JU	5	0	5	3	5	5	2	3	4	4	2	1	1
LU	15	9	24	21	20	26	24	16	16	17	19	20	28
NE	7	0	7	11	16	15	17	6	11	8	7	3	3
NW	1	1	2	1	1	2	2	2	2	2	1	0	2
OW	2	0	2	1	2	1	1	2	4	2	1	2	1
SG	2	28	30	22	22	23	24	23	15	16	16	14	13
SH	0	4	4	2	4	8	2	6	5	2	2	4	2
SO	10	3	13	12	10	14	10	7	13	12	7	9	8
SZ	0	8	8	11	6	2	2	4	11	12	12	10	3
TG	1	13	14	11	14	11	12	14	10	10	7	4	7
TI	8	9	17	17	18	27	17	9	24	12	17	17	17
UR	0	1	1	2	0	1	1	3	0	0	4	0	0
VD	44	1	45	45	40	55	39	29	32	27	26	27	29
VS	27	0	27	15	16	12	25	20	16	10	12	7	17
ZG	1	3	4	7	4	8	4	4	4	5	5	2	3
ZH	0	65	65	57	68	54	57	53	54	48	45	44	50
FL*	0	2	2	4	2	3	3	0	1	2	1	2	1
Ausland	3	3	6	9	3	4	2	5	3	7	6	5	4
Total	250	211	461	425	439	438	408	375	384	350	317	295	342

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

Tabelle 8: Gewünschte Hochschulen Zahnmedizin nach Wohnkantonen 2020 und Vergleich mit 2010 bis 2019.

Wohnkanton/ Wohnort	Zahnmedizin 2020 gewünschte Hochschule				Total Vorjahre									
	BS	BE	ZH	Total	2019	2018	2017	2016	2015	2014	2013	2012	2011	2010
AG	4	3	16	23	30	27	28	27	22	23	26	33	23	16
AI	0	0	1	1	0	0	0	0	0	1	2	0	1	0
AR	0	0	1	1	2	1	1	1	4	2	2	1	3	5
BE	3	41	1	45	44	46	38	55	43	43	48	38	36	31
BL	13	0	0	13	25	15	7	14	14	17	17	12	19	10
BS	10	0	0	10	13	12	14	11	15	14	23	22	10	10
FR	1	3	0	4	7	8	1	7	2	4	4	8	7	7
GE	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GL	0	0	0	0	1	1	3	2	2	2	2	3	3	0
GR	0	0	0	0	2	5	9	7	3	3	6	6	5	6
JU	0	0	0	0	1	1	1	0	1	1	0	0	0	2
LU	4	6	6	16	10	14	12	14	14	22	14	13	15	19
NE	0	0	0	0	1	1	0	1	0	0	0	0	1	0
NW	0	1	0	1	1	0	1	1	1	2	4	1	1	2
OW	0	0	1	1	3	0	1	3	1	2	1	0	1	1
SG	1	3	14	18	16	22	20	19	27	21	24	24	12	13
SH	0	0	1	1	1	1	2	4	3	4	1	2	0	0
SO	7	4	2	13	10	7	13	8	7	12	15	15	15	14
SZ	3	0	4	7	4	3	6	3	6	8	5	3	4	8
TG	1	1	6	8	3	4	5	9	8	5	5	5	2	9
TI	1	1	5	7	2	6	2	4	4	8	4	9	8	7
UR	0	0	1	1	0	0	0	0	2	1	1	3	2	2
VD	0	0	0	0	1	0	0	1	2	2	1	0	2	2
VS	1	1	0	2	2	0	0	1	0	2	3	0	0	2
ZG	2	2	0	4	5	5	5	7	8	7	3	2	2	1
ZH	2	0	45	47	40	42	32	50	61	53	58	43	56	44
FL*	0	0	0	0	1	1	1	1	2	2	4	0	3	1
Ausland	3	1	0	4	3	2	3	1	6	3	2	5	8	8
Total	56	67	104	227	228	224	205	251	258	264	275	248	239	220

*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 2020 (alle Teilnehmenden).

Wohnkanton	Testort								Total
	Basel	Bern	Chur	Freiburg	Lugano	Luzern	St. Gallen	Zürich	
AG	87	25	0	0	0	37	1	225	375
AI	0	0	0	0	0	0	8	0	8
AR	1	0	1	0	0	1	21	3	27
BE	5	524	1	17	0	1	1	12	561
BL	272	3	0	0	0	0	0	8	283
BS	172	2	1	0	2	1	1	2	181
FR	1	40	0	107	0	0	0	1	149
GE	0	0	0	29	0	0	0	1	30
GL	2	0	2	0	0	0	0	13	17
GR	0	0	87	0	5	1	2	8	103
JU	1	0	0	13	0	0	0	0	14
LU	2	2	0	0	0	190	0	3	197
NE	0	0	0	18	0	0	0	0	18
NW	1	0	0	0	0	26	1	1	29
OW	0	0	0	0	0	18	0	0	18
SG	1	2	36	1	0	0	174	26	240
SH	1	0	0	0	0	0	6	26	33
SO	51	83	0	0	0	3	0	16	153
SZ	2	2	3	0	0	23	2	49	81
TG	1	2	0	0	0	0	61	51	115
TI	0	1	2	0	140	1	1	5	150
UR	0	1	0	0	0	8	0	1	10
VD	0	5	0	65	0	0	0	0	70
VS	2	33	0	58	0	1	0	1	95
ZG	3	0	0	0	0	57	1	24	85
ZH	5	3	1	3	2	9	9	903	935
FL*	0	1	14	0	0	0	2	1	18
Ausland	22	10	1	8	1	1	7	13	63
Total	632	739	149	319	150	378	298	1393	4058

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

5.6 Teilnahmen am EMS (NC) nach Kantonen seit 1998

Seit 1998 wurden in der Schweiz 52 542 gültige Ergebnisse im Rahmen des NC erzielt. Seit Beginn unterliegt Humanmedizin, seit 1999 Veterinärmedizin und erst seit 2004 Zahnmedizin einem NC. Zu beachten sind die substanziellen Teilnehmerzahlen aus der Romandie – insbesondere für die Disziplin Veterinärmedizin. Sie machen deutlich, dass adäquate Vorbereitungsmöglichkeiten auch für diese Sprachgruppe, unabhängig von den nicht dem NC unterworfenen französischsprachigen Studienangeboten, wichtig sind.

Tabelle 10: Gesamtzahl der Teilnehmer am EMS 1998 bis 2020 – nur Hochschulen mit NC.

Wohnkanton/ Wohnort	Disziplin			Geschlecht		Sprache			Total
	Human- medizin	Veterinär- medizin	Zahn- medizin	männlich	weiblich	deutsch	franzö- sisch	italie- nisch	
AG	3557	442	351	1483	2867	4338	4	8	4350
AI	109	17	9	68	67	135	0	0	135
AR	365	42	35	163	279	441	1	0	442
BE	6280	986	651	2717	5200	7752	163	2	7917
BL	2451	257	222	1058	1872	2923	7	0	2930
BS	2180	131	194	956	1549	2494	2	9	2505
FR	1595	319	82	695	1301	712	1281	3	1996
GE	104	288	0	74	318	17	375	0	392
GL	172	59	25	71	185	255	1	0	256
GR	1219	211	84	536	978	1439	1	74	1514
JU	125	70	11	77	129	15	189	2	206
LU	2592	359	207	1171	1987	3151	1	6	3158
NE	80	177	6	56	207	20	243	0	263
NW	229	31	19	111	168	279	0	0	279
OW	190	23	17	65	165	230	0	0	230
SG	2725	361	297	1156	2227	3381	2	0	3383
SH	394	61	31	175	311	486	0	0	486
SO	1515	196	183	692	1202	1892	2	0	1894
SZ	800	116	73	288	701	986	2	1	989
TG	1217	201	93	453	1058	1511	0	0	1511
TI	1599	289	93	808	1173	47	1	1933	1981
UR	169	21	20	73	137	210	0	0	210
VD	283	608	12	191	712	49	854	0	903
VS	713	248	32	347	646	545	448	0	993
ZG	832	99	61	376	616	990	1	1	992
ZH	9851	994	696	3958	7583	11477	26	38	11541
FL*	197	28	27	80	172	252	0	0	252
Ausland	683	88	63	327	507	738	69	27	834
Total	42226	6722	3594	18225	34317	46765	3673	2104	52542

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

5.7 Testabsolvierung nach Geschlecht

In **Humanmedizin** beträgt das Geschlechterverhältnis der Teilnahmen weiterhin ungefähr 65 zu 35 zugunsten der Frauen (vergleiche Abbildung 8). Gegenüber 2019 hat der Anteil männlicher Kandidaten im aktuellen Jahrgang wiederum marginal zugenommen.

In **Veterinärmedizin** bleibt der Anteil der Frauen auf sehr hohem Niveau praktisch konstant, bei einer marginalen Zunahme des Anteils männlicher Kandidaten.

In der **Zahnmedizin** war 2020 eine deutliche Zunahme weiblicher Kandidatinnen zu beobachten.

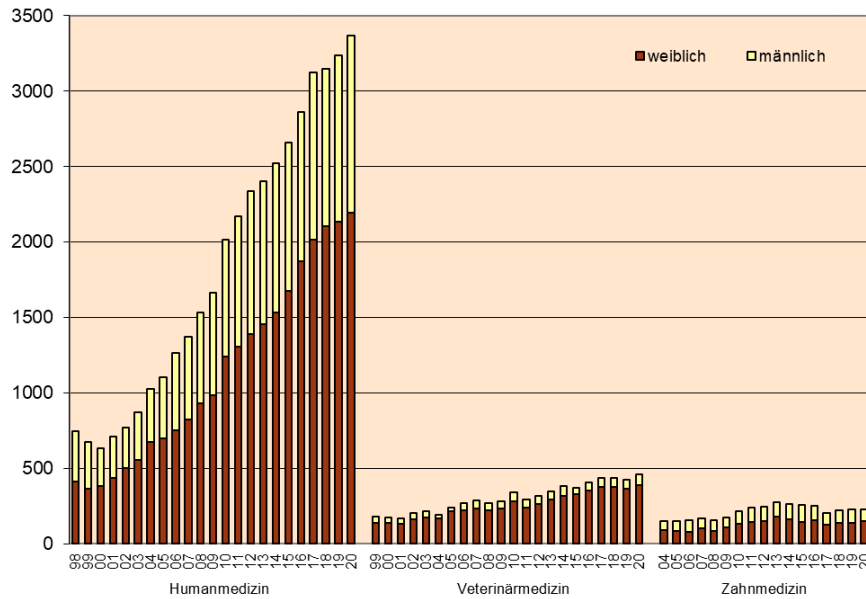


Abbildung 7: Anzahl männlicher Bewerber und weiblicher Bewerberinnen für Humanmedizin (ab 2008 inklusive Chiropraktik), Veterinärmedizin und Zahnmedizin, Jahre mit NC.

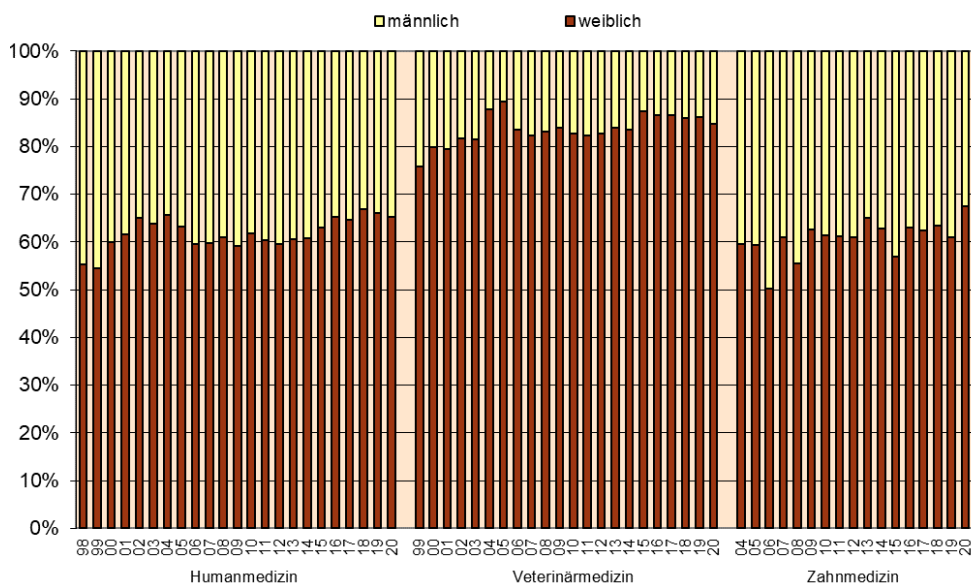


Abbildung 8: Anteil der Geschlechter bei der Bewerbung für Humanmedizin (ab 2008 inklusive Chiropraktik), Veterinärmedizin und Zahnmedizin, Jahre mit NC, bezogen auf 100% pro Jahr und Disziplin.

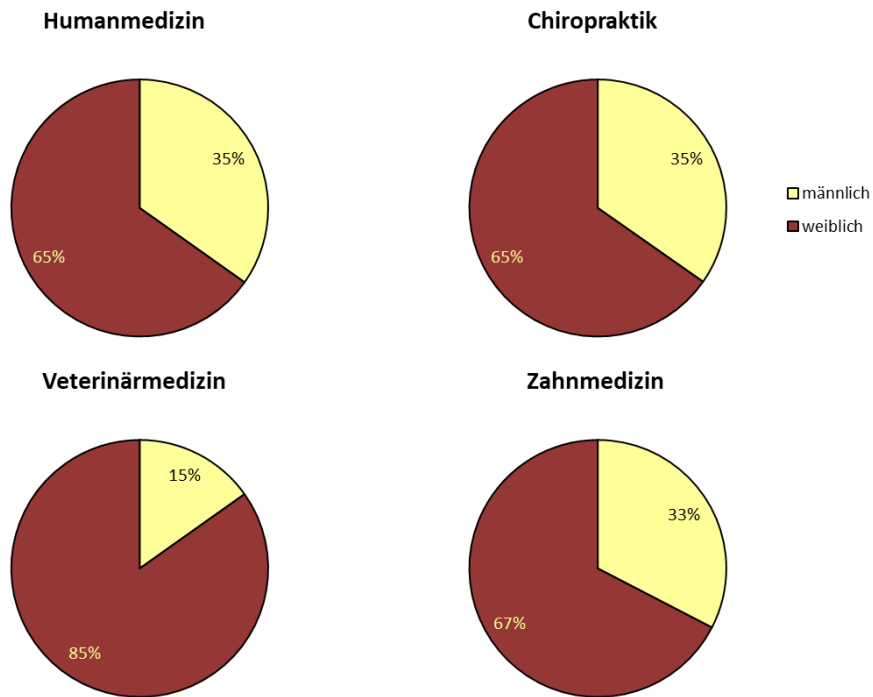


Abbildung 9: Verhältnisse der Geschlechter bei der Bewerbung nach Disziplin 2020.

5.8 Übernahme des Testergebnisses aus dem Vorjahr

Als Alternative zu einem Neuantritt kann ein vorhandenes Ergebnis des Vorjahres übertragen werden (bei einem Neuantritt zählt das neue Ergebnis). In Abhängigkeit der Studienplätze und Teilnehmerzahlen kann dabei ein Test-Prozentrang im Folgejahr nicht mehr für die Zulassung ausreichen, obwohl er im aktuellen Jahr noch genügt hätte. Bei der Entscheidung für diese Variante sollte daher eine gewisse Reserve vorhanden sein.

Viele Personen treten 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 antrat.

In dem Zusammenhang sei 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 bestehen 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
Humanmedizin	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)
	2018	72	20	100	83.7	13.0	66 (6 nicht)
	2019	51	15	100	82.1	14.4	43 (8 nicht)
	2020	73	10	100	80.6	16.8	66 (7 nicht)
Veterinärmedizin	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)
	2018	5	54	81	66.8	12.8	3 (2 nicht)
	2019	10	61	98	80.9	12.8	10 (0 nicht)
	2020	8	17	92	58.5	22.6	5 (3 nicht)
Zahnmedizin	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)
	2018	16	30	96	59.7	13.3	15 (1 nicht)
	2019	13	35	99	60.8	17.0	12 (1 nicht)
	2020	18	27	88	54.8	15.4	15 (3 nicht)

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

2020 haben sich 73 Personen (66 zugelassen) mit einem Ergebnis aus dem Vorjahr für Humanmedizin beworben, 8 (5 zugelassen) für Veterinärmedizin und 18 (15 zugelassen) für Zahnmedizin.

5.9 Sprachgruppen

Nach wie vor treten vor allem deutschsprachige Kandidaten und Kandidatinnen zum EMS an, für die französisch- und italienischsprachigen Kandidaten und Kandidatinnen existieren teilweise Alternativen, das Studium ohne NC in der gewünschten Sprache zu absolvieren. Der Anteil der französisch- und italienischsprachigen Teilnehmenden bleibt über die Jahre vergleichbar und ist in Veterinärmedizin mangels Alternativen höher als in Human- und Zahnmedizin.

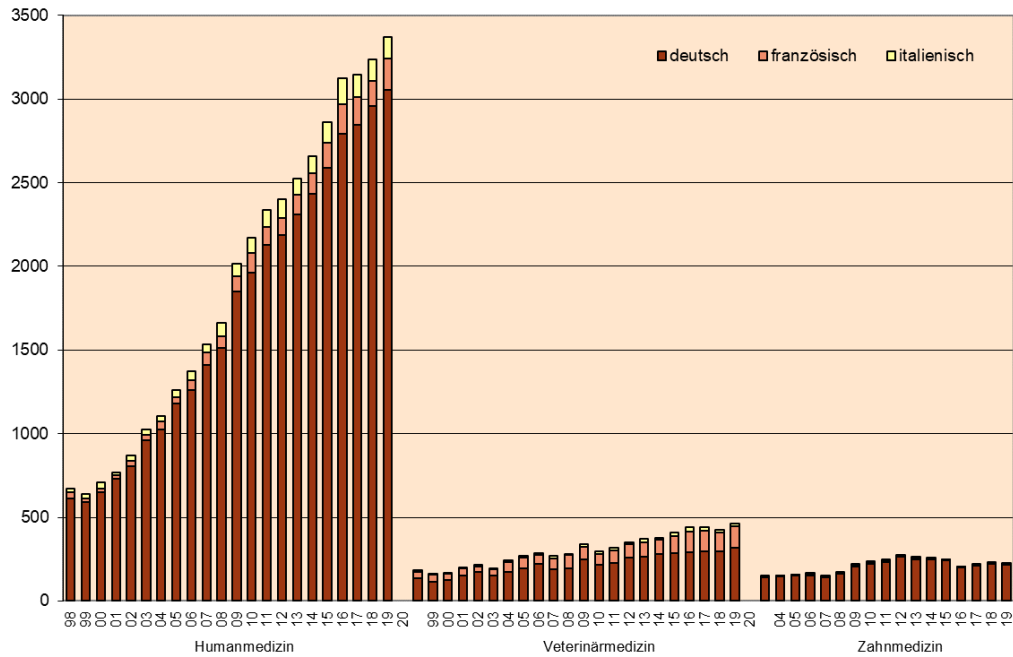


Abbildung 10: Vergleich der Sprachgruppen 1998 bis 2020, Absolutzahlen.

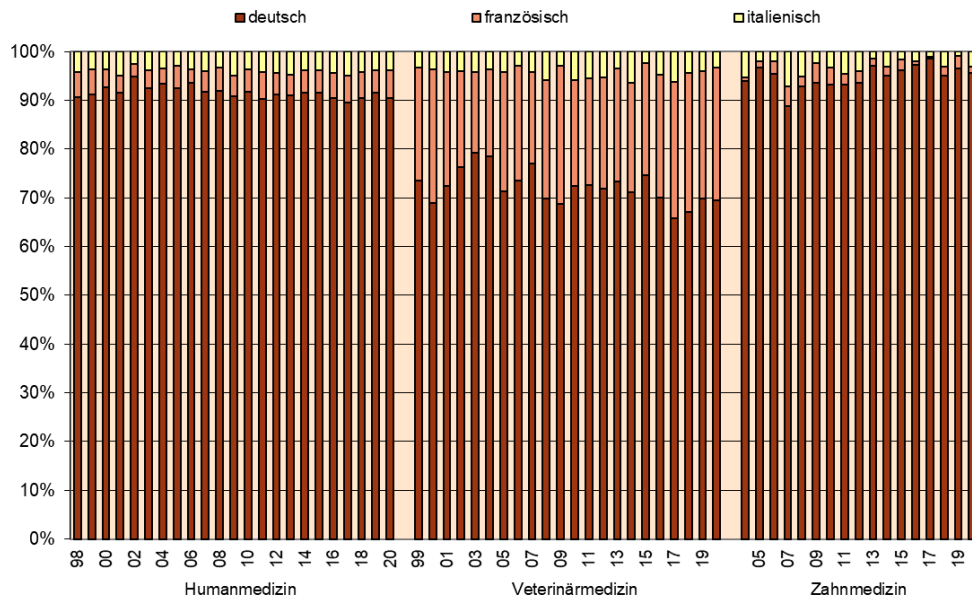


Abbildung 11: Vergleich der Sprachgruppen 1998 bis 2020. Bewerber und Bewerberinnen pro Disziplin und Jahr auf 100% bezogen.

6 Der EMS im Corona-Jahr

Die vom Bund eingeführten und empfohlenen Corona-Massnahmen legten nahe, auch für die (über lange Zeit unsichere) Durchführung des EMS Vorkehrungen zu treffen. Neben organisatorischen (Einlassprozedur...) und ausstattungstechnischen (Sitzabstände, Desinfektionsmittel...) Massnahmen in den Testlokalen wurde auch eine Anpassung des Testablaufs selbst vorgenommen, mit dem Ziel, kritische Konstellationen zu vermeiden und somit zum Schutz der Teilnehmenden und des Testpersonals vor Ort beizutragen.

Zu den kritischsten Zeitpunkten hinsichtlich der Einhaltung des Abstandes von anderen Personen gehören der Einlass zum Test, das Verlassen des Testlokals sowie das Austeilen und Einsammeln von Testmaterialien durch das Testpersonal. Die vorgenommene Verkürzung des EMS sollte dazu beitragen, diese Elemente zu minimieren. Entsprechend wurde auf die Mittagspause verzichtet und es entfielen gewisse Aufgabengruppen.

Dabei soll angemerkt sein, dass es sich bei den durchgeführten Massnahmen um eine Reaktion auf die aktuelle Lage gehandelt hat. Sie sollen damit ausdrücklich nicht als „schlankere“ Lösung auch für zukünftige Testdurchführungen „ohne Not“ interpretiert werden. Wie nachfolgend aufgezeigt, wurde die Berücksichtigung der Corona-Massnahmen durch die Verkürzung des Tests mit den erwarteten Einbussen hinsichtlich der Gütekriterien „erkaufte“ – die angesichts der Situation in ihrem Ausmass aber vertretbar erschienen und die Gültigkeit der Testergebnisse in keiner Weise in Frage stellen.

Die getroffenen Massnahmen sind nachfolgend dargestellt.

6.1 Distanzen

Eine Vergrösserung der Distanzen zwischen den Arbeitsplätzen (mindestens 4m²) erforderte in einigen Testorten die Einrichtung zusätzlicher Testlokale.

6.2 Schutzmasken und Desinfektionsmittel

Ab Betreten des Gebäudes bis nach dem Austeilen der Testmaterialien war das Tragen einer Schutzmaske für die Teilnehmenden obligatorisch (für das Personal auch während des Tests, mit Ausnahme der Testleitung). Im Einladungsschreiben wurde zum Mitbringen einer Maske aufgefordert, zusätzlich wurden ausserhalb der Testlokale Masken verteilt. In allen Testlokalen wurden Desinfektionsstationen eingerichtet.

6.3 Einlass

Um Warteschlangen vor den Testlokalen zu verringern, wurden die Einlasszeiten „gestreckt“ und der Testbeginn auf 9.30 Uhr (statt 9.00 Uhr) verlegt.

6.4 Versetzter Testbeginn in Zürich

Aufgrund der hohen Personenzahlen in Zürich wurde vorgesehen, die Testdurchführung in Zürich zu staffeln – dies nicht nur zur Entlastung des öffentlichen Verkehrs bei der Anreise, sondern auch, um die Situation beim Einlass zu entschärfen. Der zweite, spätere Beginn in Zürich hatte insofern Auswirkungen auf alle anderen Testorte, dass das Verlassen des Testlokals bei einem allfälligen Abbruch bis zu diesem Zeitpunkt nicht gestattet gewesen wäre. Dieser Fall trat jedoch nicht ein.

6.5 Mittagspause

Das Verlassen und der Wiedereintritt der Testlokale führen zu „Flaschenhälsen“. Aus diesem Grund wurde entschieden, 2020 auf diese Prozedur, und damit die Mittagspause, zu verzichten. Dieser Entscheid hatte zur Folge, dass der Test verkürzt werden musste.

6.6 Austeilen von Materialien

Während des normalen Testablaufs wird mehrfach Material ausgeteilt, beziehungsweise eingesammelt. Im Einzelnen sind dies in chronologischer Reihenfolge:

- Austeilen „Hinweise zur Bearbeitung des Tests“
- Austeilen der Testhefte Teil A
- Einsammeln der Testhefte Teil A und Hinweise
- Austeilen der Lernhefte
- Einsammeln der Lernhefte
- Austeilen der Testhefte Teil B
- Einsammeln der Testhefte Teil B
- Einsammeln der Befragung

Ein verkürzter Testablauf hatte auch zum Ziel, diese Arbeitsgänge zu verringern und so den Kontakt zwischen Teilnehmenden und Testpersonal zu minimieren. Kontrollgänge des Testpersonals fanden daneben weiterhin statt – unter Einhaltung der geforderten Abstände. Ebenso durfte nicht darauf verzichtet werden, die Identität der Teilnehmenden sowie zu Beginn des Tests die Beantwortung auf das Antwortblatt (statt in das Testheft) systematisch zu prüfen und gegebenenfalls Personen auf die korrekte Bearbeitung aufmerksam zu machen.

Die Vorgänge der Materiallogistik beschränkten sich mit der verkürzten Testversion auf:

- Austeilen „Hinweise zur Bearbeitung des Tests“ (bereits vor dem Eintreffen der Teilnehmenden)
- Austeilen der Testhefte
- Einsammeln der Testhefte und Hinweise, gleichzeitig Austeilen der Befragung
- Abgabe der Befragung durch die Teilnehmenden in vorgesehenen Behälter (kontaktlos)

6.7 Entlassung der Teilnehmenden

Nach dem Abschluss des Tests (inklusive Durchführung der Befragung während des Nachzählens des Materials) wurden die Teilnehmenden nicht frei, sondern geordnet, abschnittsweise aus dem Raum entlassen, um auch hier grössere Ansammlungen zu vermeiden.

Insgesamt haben sich die Massnahmen bewährt – ein grosser Dank ist an dieser Stelle an alle beteiligten Koordinatorinnen und Koordinatoren, Testleiterinnen und Testleiter, Stellvertreterinnen und Stellvertreter, das Aufsichtspersonal – und nicht zuletzt auch an die Kandidatinnen und Kandidaten zu richten, die unter diesen besonderen Umständen in vorbildlicher Weise zum Schutz aller Beteiligten kooperiert haben.

7 Beschreibung des verwendeten Eignungstests

Der originale EMS besteht aus neun Aufgabengruppen, die zu einem Gesamtwert verrechnet werden.

Die nachfolgend dargestellte Verkürzung hat es erlaubt, den Test ohne Mittagspause durchzuführen, sowie die Arbeitsschritte des Austeilens und Einsammelns von Materialien zu verringern. Tabelle 12 gibt einen Überblick über den originalen und den verkürzten Testaufbau und die Dauer der einzelnen Aufgabengruppen.

Tabelle 12: Struktur und Ablauf des EMS original und verkürzt

Ablauf EMS original		Ablauf „EMS verkürzt“	
Aufgabengruppe	Zeit (Min)	Aufgabengruppe	Zeit (Min)
Muster zuordnen	18	Muster zuordnen	18
Med.-naturwiss. Grundverständnis	50	Med.-naturwiss. Grundverständnis	50
Schlauchfiguren	12	Quantitative und formale Probleme	50
Quantitative und formale Probleme	50	Textverständnis	45
Konz. und sorgf. Arbeiten	8	Diagramme und Tabellen	50
Mittagspause	90	Konz. und sorgf. Arbeiten	8
Einprägephase			
Figuren lernen	4		
Fakten lernen	6		
Textverständnis	45		
Reproduktionsphase			
Figuren lernen	5		
Fakten lernen	7		
Diagramme und Tabellen	50		
Gesamtdauer: Ca. 9.00 bis 16.00 Uhr		Gesamtdauer: Ca. 9.30 bis 13.15 Uhr	

Der Verzicht auf die Aufgabengruppen „Schlauchfiguren“, „Figuren lernen“ und „Fakten lernen“ resultierte in einer Zeitersparnis von (netto) gut 30 Minuten. Die Gesamtdauer des Tests betrug damit noch 221 Minuten (statt 255) und lag beispielsweise unter derjenigen schriftlicher Mathematik-Maturitätsprüfungen in Bern gemäss Matur-Anerkennungsreglement.

Beim EMS handelt es sich um eine Adaptation des deutschen **Tests für Medizinische Studiengänge (TMS)** mit Weiterentwicklungen für die und (durch das ZTD) in der 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.

In jeder Aufgabengruppe können maximal 20 Punkte erzielt werden (Ausnahme: In Textverständnis 18). Die Aufgabengruppen „Muster zuordnen“, „Fakten lernen“ (2020 nicht eingesetzt), „Figuren lernen“ (2020 nicht eingesetzt) sowie „Konzentriertes und sorgfältiges Arbeiten“ werden jährlich durch das Zentrum für Testentwicklung neu erstellt.

Die Vorteile des EMS lassen sich folgendermassen zusammenfassen:

- Die Auswahl der Testanforderungen erfolgt aus einer grösseren Menge möglicher Studienanforderungen aufgrund einer Anforderungsanalyse. Dabei werden ständige Anpassungen an sich verändernde Bedingungen geprüft und wo nötig vorgenommen.
- Der wissenschaftliche Nachweis der Vorhersagbarkeit des Studienerfolgs mittels EMS ist gewährleistet. Dies erlaubt die gesetzlich geforderte Studienzulassung nach Eignung.
- Die Konstruktion der Aufgaben erfolgt durch Experten und es findet eine anschliessende empirische Überprüfung statt, damit die Aufgaben alle Gütekriterien erfüllen, die gewünschte Fähigkeit tatsächlich messen und optimal „schwierig“ sind.
- Es wird beachtet, 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 Lehrbeauftragte sowie Experten und Expertinnen mitarbeiteten. Das Ziel war die Vorhersage des Studienerfolgs, 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. 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. 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 mitbearbeitet, 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.

7.1 Berechnung der Werte 2020

Die Aufgabengruppen des Eignungstests für das Medizinstudium (EMS) liefern Punkte als Summe richtig gelöster Aufgaben – zwischen 0 und 20, bzw. 18 für die Aufgabengruppe Textverständnis. Durch die Verkürzung des Tests konnten 2020 noch maximal 118 Punkte erreicht werden. Aufgrund der Kennwerteanalysen (Schwierigkeiten und Trennschärfen) können alle Aufgaben gewertet werden.

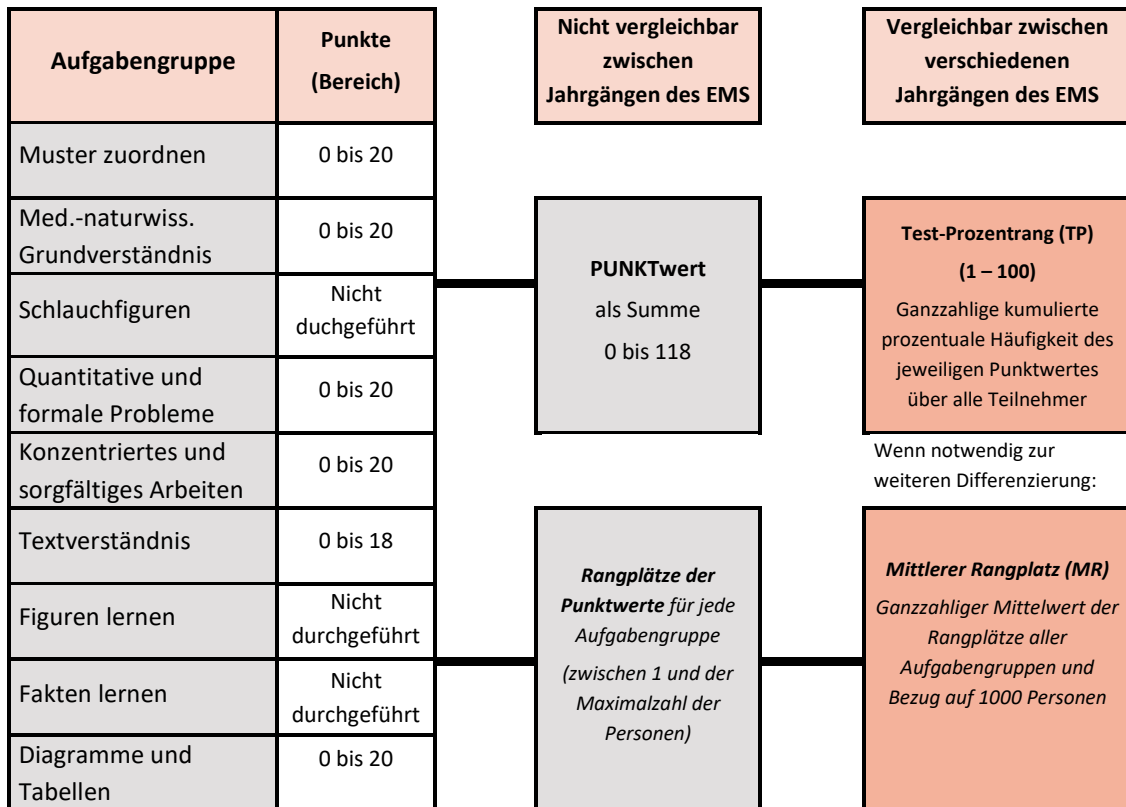


Abbildung 12: Punktwerte der einzelnen Aufgabengruppen 2020 und ihre Zusammenführung über den Punktwert zum Test-Prozentrang und mittlerer Rangplatz.

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 Punktwertevertellungen. Alle Punkte der Aufgabengruppen werden zu einer Summe addiert (**Punkt**wert). 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 Häufigkeits-Verteilung 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 Anmeldungen 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 genügend Reserve gegenüber dem für die Zulassung im Jahr der Testabsolvierung ausreichenden Wert einkalkuliert werden, wenn man das Ergebnis übertragen möchte.

7.2 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 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. Die zweite Aufgabengruppe wird besser bewertet, weil die 20 Punkte dort schwerer zu erzielen waren.

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

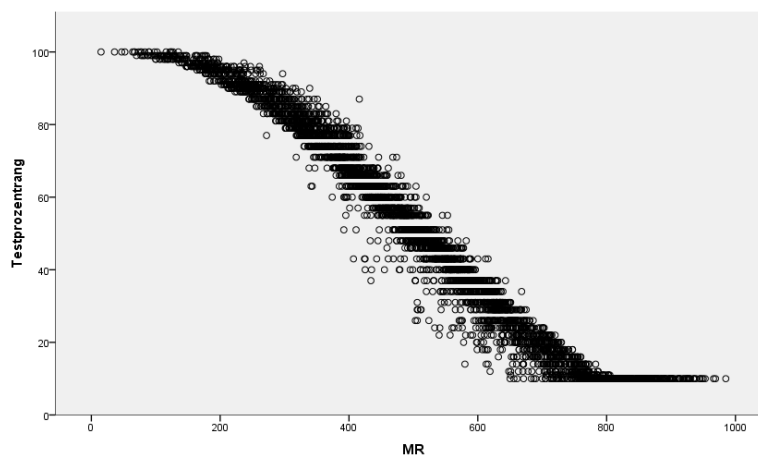


Abbildung 13: Beziehung zwischen Test-Prozentrang und mittlerem Rangplatz für 2020 (Werte kleiner 10 bei TP auf 10 angehoben).

Der mittlere Rangplatz MR bewegt sich 2020 zwischen 15 und 985. Niedrige Werte stehen für bessere Leistungen. Mittlerer Rangplatz und Test-Prozentrang korrelieren 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 und Kandidatinnen die maximale Punktzahl 20. Bei gleicher Punktzahl erhält jede Person den durchschnittlichen Rang dieser Personen (die 1 bis 10 belegen).
- Kandidatin A hat diese maximale Punktzahl erreicht und belegt damit in dieser Aufgabengruppe zusammen mit den anderen 9 Kandidaten und Kandidatinnen Rang 5.5 (Durchschnitt der Ränge 1 bis 10, also $(1+10)/2=5.5$).
- Kandidat B hat, zusammen mit 14 weiteren Kandidaten und Kandidatinnen, 19 Punkte erzielt. Diese Kandidaten und Kandidatinnen 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.
- Kandidatin A erreicht hier 19 Punkte, zusammen mit nur 3 weiteren Kandidaten und Kandidatinnen. Sie alle erhalten in dieser Aufgabengruppe den durchschnittlichen Rang $(2+5)/2=3.5$.

Kandidatin A und Kandidat B erreichten beide damit die gleiche Punktzahl von 39 Punkten (und den gleichen Test-Prozentrang).

Der mittlere Rangplatz für Kandidatin A und Kandidat B berechnet sich wie folgt:

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

Und auf 1000 Personen:

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

Hätten 1000 Personen am Test teilgenommen, resultiert für Kandidatin A im Mittel über beide Aufgabengruppen der 45. Rang, für Kandidat B der 95. Rang. Kandidatin A wird dafür belohnt, in der ersten Aufgabengruppe die maximale, aber vor allem auch in der zweiten Aufgabengruppe zusammen mit nur sehr wenigen Kandidaten und Kandidatinnen 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 Kandidatin A und Kandidat B ist dadurch gerechtfertigt, dass Kandidatin A trotz gleicher Punktzahl über beide Aufgabengruppen mehr Kandidaten und Kandidatinnen übertroffen hat als Kandidat B.

Müsste man zwischen beiden Personen unterscheiden, wäre Kandidatin A die bessere.

8 Testanwendung 2020

8.1 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	58.6	10.4		7.3	9.7			9.6	9.7	11.8
Median	59.0	10.0		7.0	10.0			10.0	10.0	12.0
Modalwert	53	9		8	9			9	9	12
Stand.-abweichung	12.9	3.7		2.4	3.0			2.9	3.0	4.3
Spannweite	93	20		17	19			19	19	20
25. Perzentil	50	8		6	8			8	8	9
75. Perzentil	68	13		9	12			12	12	15
Schiefe	-0.11	0.08		0.15	0.02			-0.12	0.01	-0.56
Kurtosis	-0.14	-0.56		0.07	-0.26			-0.13	-0.36	0.31

Bei der Aufgabengruppe „Textverständnis“ wurden auch 2020 die tiefsten mittleren Punktwerte erreicht (maximal 18 wären möglich). Alle verwendeten Aufgabengruppen differenzieren gut im relevanten Wertebereich.

8.2 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.

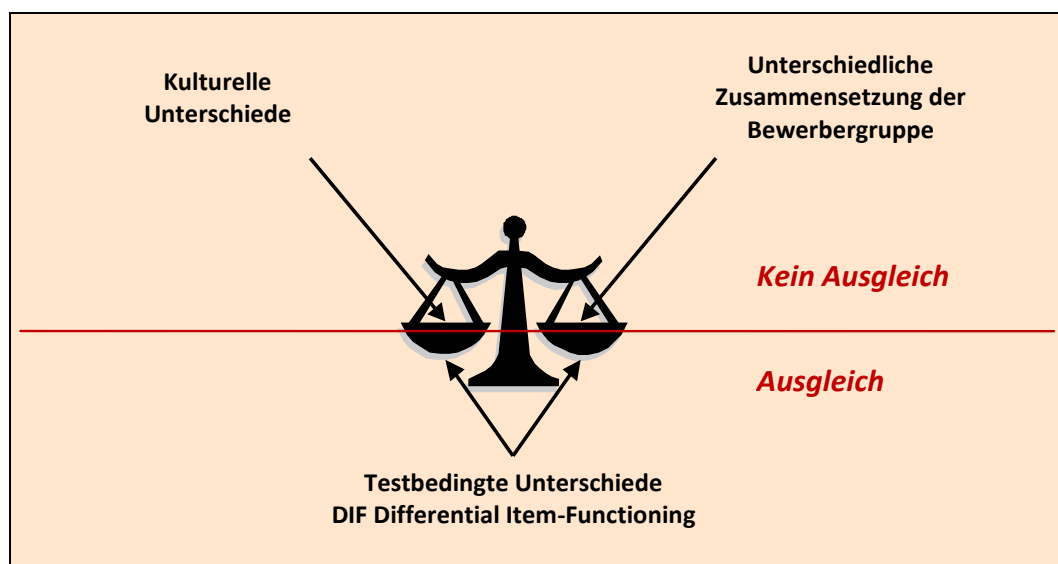


Abbildung 14: Faktoren mit Wirkung auf Unterschiede zwischen den Sprachversionen.

Zwei Ursachen (kulturelle Unterschiede und die unterschiedliche Zusammensetzung der Bewerbergruppe) führen potenziell 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 können durch die Adaptation verursacht sein 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 Teilnehmenden macht dies wahrscheinlich: Die Auswahl kann die „Spitze“ oder das „Ende“ der Leistungsrangreihe aller Maturanden und Maturandinnen 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 (2020: vier) von neun (2020: sechs) Aufgabengruppen. Der Sprachausgleich beschränkt sich demnach auf die stärker sprachabhängigen Aufgabengruppen „Medizinisch-naturwissenschaftliches Grundverständnis“, „Quantitative und formale Probleme“, „Textverständnis“, „Fakten lernen“ (entfiel 2020) 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 sprachabhängigen Aufgabengruppen zu erwarten, während die 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.

8.2.1 Sprachvergleich für die Aufgabengruppen

Auch 2020 treten Unterschiede zwischen den Sprachgruppen in sprachabhängigen und wenig sprachabhängigen Aufgabengruppen auf. Die unkorrigierten Punktzahlen sind nachfolgend abgebildet.

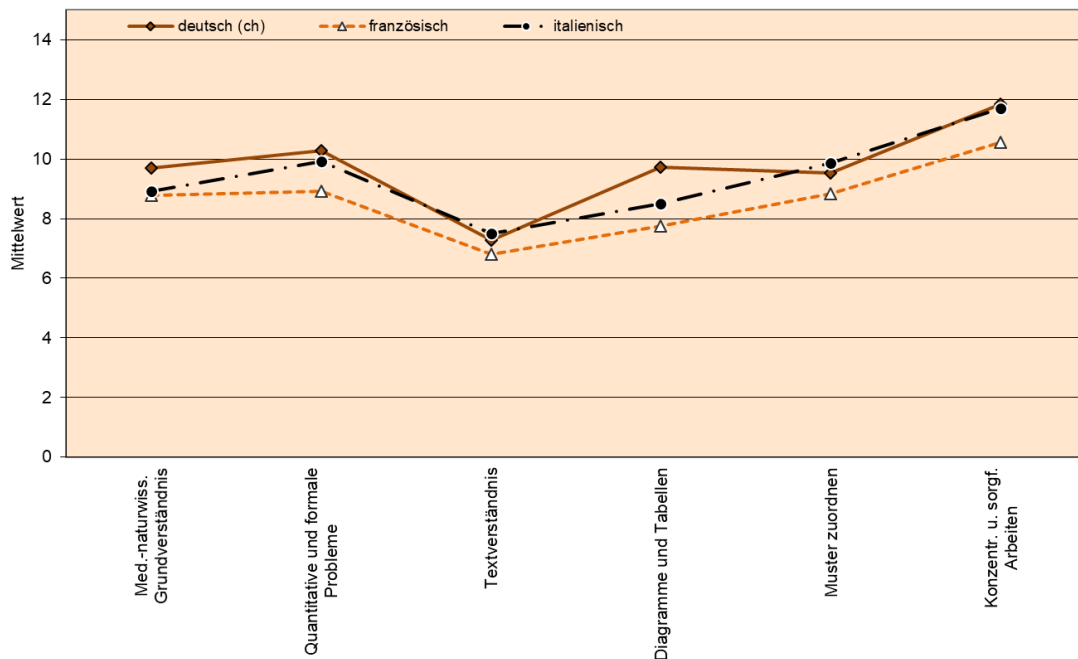


Abbildung 15: Mittelwerte der Punktwerte für die Aufgabengruppen nach Sprachen (unkorrigiert).

Abbildung 16 zeigt den Vergleich zwischen deutsch- und französischsprachigen Testteilnehmenden für das Jahr 2020, in Abbildung 17 ist der Vergleich zwischen deutsch- und italienischsprachigen Personen dargestellt. Die senkrechten Linien bezeichnen Mittelwert und Standardabweichung der Differenzen über die Jahrgänge für die letzten 10 Jahre. Positive Differenzwerte deuten auf bessere Ergebnisse der deutschen Sprachgruppe hin, negative auf solche der jeweils anderen Sprachgruppe.

Für die französische Sprachgruppe sind in „Textverständnis“ eher tiefe Differenzen, in „Konzentriertes und sorgfältiges Arbeiten“ sowie „Quantitative und formale Probleme“ eher höhere Unterschiede zu der deutschen Sprachgruppe zu beobachten.

Die italienischsprachigen Teilnehmenden erzielten in „Muster zuordnen“ und „Textverständnis“ bessere Ergebnisse als die deutsche Sprachgruppe, in der Aufgabengruppe „Medizinisch-naturwissenschaftliches Grundverständnis“ fielen die Differenzen hingegen in entgegengesetzter Richtung etwas höher aus als üblich.

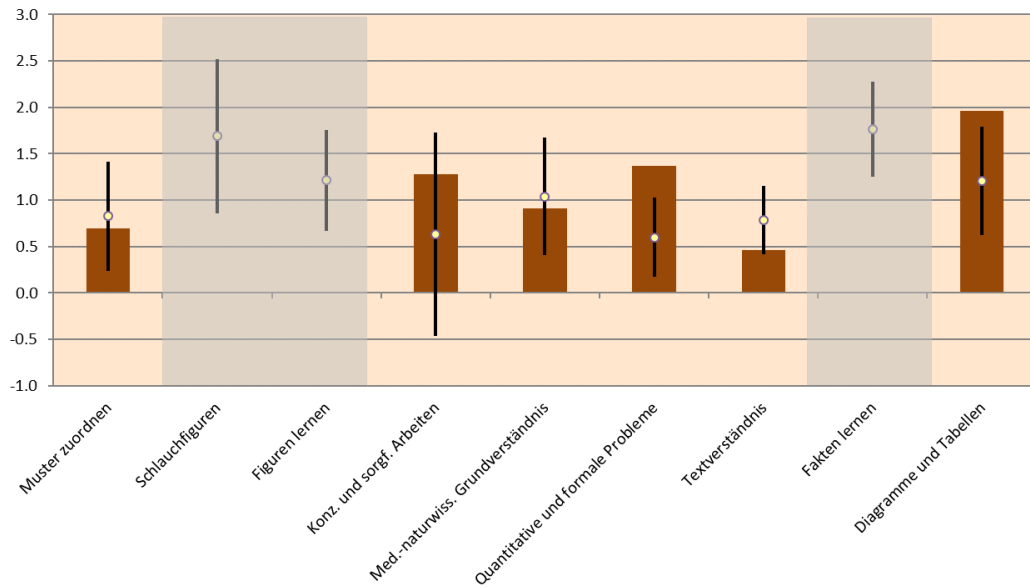


Abbildung 16: Differenz für Aufgabengruppen deutschsprachig-französischsprachig 2020 (unkorrigiert) und Mittelwerte mit Standardabweichungen der Testdurchführungen 2010-2019.

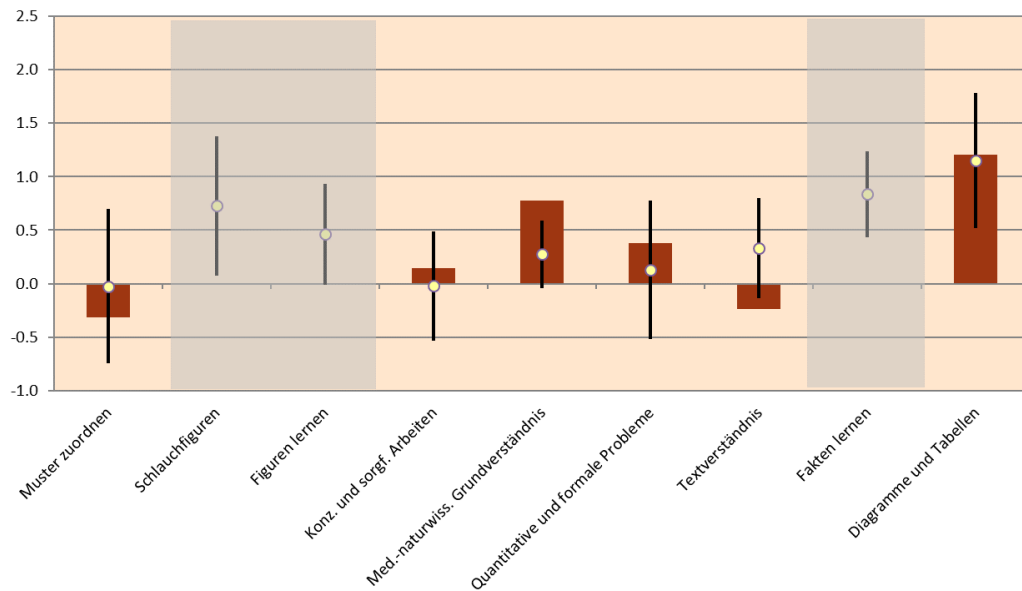


Abbildung 17: Differenz für Aufgabengruppen deutschsprachig-italienischsprachig 2020 (unkorrigiert) und Mittelwerte mit Standardabweichungen der Testdurchführungen 2010-2019.

8.2.2 Darstellung des Korrekturverfahrens

8.2.2.1 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.

8.2.2.2 Welche Aufgabengruppen ausgleichen?

Die folgenden Tabellen zeigen die Sprachvergleiche für die sprachabhängigen Aufgabengruppen. Referenz ist die deutschsprachige Testversion. Die französischsprachigen Teilnehmenden erzielten in allen vier sprachabhängigen Aufgabengruppen signifikant geringere Punktwerte als die deutschsprachigen Teilnehmenden. Für die italienische Sprachgruppe sind „Medizinisch-naturwissenschaftliches Grundverständnis“ sowie „Diagramme und Tabellen“ betroffen.

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

Aufgabengruppe	Sprache	n	m	s	Sig.	DIF-Prüfung
Med.-naturwiss. Grundverständnis	d	3589	9.69	3.022	.000	Ja
	f	319	8.78	3.062		
Quantitative und form. Probleme	d	3589	10.29	3.734	.000	Ja
	f	319	8.92	3.731		
Textverständnis	d	3589	7.27	2.377	.001	Ja
	f	319	6.81	2.258		
Fakten lernen						
Diagramme und Tabellen	d	3589	9.71	2.965	.000	Ja
	f	319	7.75	2.788		

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

Aufgabengruppe	Sprache	n	m	s	Sig.	DIF-Prüfung
Med.-naturwiss. Grundverständnis	d	3589	9.69	3.022	.002	Ja
	i	150	8.91	2.983		
Quantitative und form. Probleme	d	3589	10.29	3.734	.223	Nein
	i	150	9.91	3.534		
Textverständnis	d	3589	7.27	2.377	.241	Nein
	i	150	7.51	2.503		
Fakten lernen						
Diagramme und Tabellen	d	3589	9.71	2.965	.000	Ja
	i	150	8.51	2.792		

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

8.2.3 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 und Probandinnen 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.

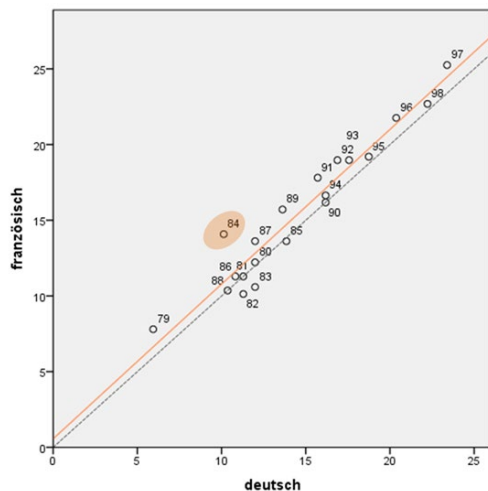


Abbildung 18: Beispiel für einen Delta-Plot.

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.

Auffällige Abweichungen einzelner Punkte von dieser Regressionsgeraden hingegen deuten 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.

8.2.3.1 Medizinisch-naturwissenschaftliches Grundverständnis

Die Aufgaben 27 und 34 liegen für die französische Sprachgruppe über der kritischen Differenz, es werden den Personen, welche diese Aufgaben nicht korrekt gelöst haben, bis zu 0.35 Punkte gutgeschrieben. Bei der italienischen Gruppe sind die Aufgaben 27 und 31 betroffen, auch dort werden bis zu 0.35 Punkte gutgeschrieben.

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

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
21	0.81	0.75	0.73	5.81	7.13	7.60		
22	0.77	0.72	0.59	6.59	7.72	10.54		
23	0.72	0.72	0.74	7.83	7.72	7.32		
24	0.59	0.50	0.62	10.44	12.46	9.84		
25	0.54	0.51	0.47	11.42	12.13	12.92		
26	0.55	0.56	0.53	11.36	11.01	11.66		
27	0.71	0.50	0.51	7.87	12.33	12.08	0.21	0.20
28	0.54	0.51	0.49	11.59	12.06	12.64		
29	0.55	0.53	0.53	11.29	11.67	11.80		
30	0.45	0.39	0.42	13.47	14.69	14.04		
31	0.56	0.46	0.41	11.11	13.25	14.18		0.15
32	0.48	0.40	0.45	12.88	14.50	13.34		
33	0.39	0.43	0.49	14.62	13.84	12.64		
34	0.37	0.23	0.27	15.12	18.11	17.25	0.14	
35	0.37	0.34	0.39	15.04	15.61	14.60		
36	0.31	0.31	0.22	16.42	16.40	18.23		
37	0.27	0.29	0.31	17.27	16.67	16.28		
38	0.25	0.18	0.27	17.60	19.03	17.11		
39	0.24	0.23	0.24	17.85	18.11	17.81		
40	0.22	0.21	0.23	18.20	18.38	18.09		

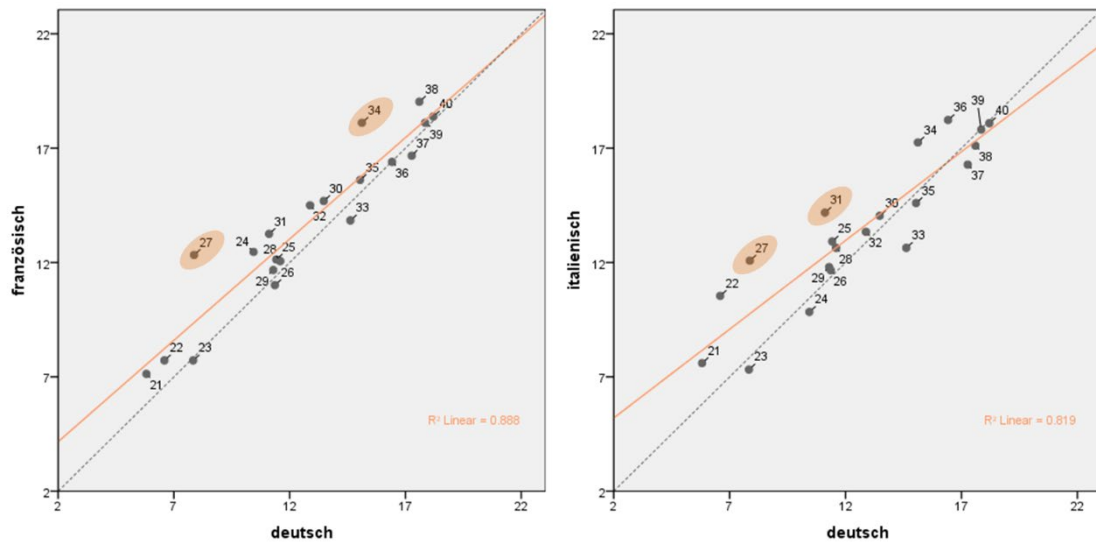


Abbildung 19: Delta-Plot für „Medizinisch-naturwissenschaftliches Grundverständnis“ (deutsch-französisch und deutsch-italienisch).

8.2.3.2 Quantitative und formale Probleme

Alle Aufgaben liegen eng an der Regressionsgeraden. Trotz signifikanter Differenz hinsichtlich des Punktwerts dieser Aufgabengruppe gibt es auf Aufgabenebene keine Hinweise auf übersetzungsbedingte Einflüsse für die französische Sprachgruppe.

Tabelle 17: DIF-Analyse „Quantitative und formale Probleme“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
41	0.85	0.82	0.87	5.06	5.62			
42	0.75	0.61	0.65	7.02	10.16			
43	0.50	0.46	0.47	12.36	13.12			
44	0.62	0.47	0.69	9.91	12.98			
45	0.68	0.65	0.73	8.57	9.30			
46	0.72	0.63	0.62	7.82	9.56			
47	0.65	0.54	0.60	9.24	11.60			
48	0.61	0.49	0.64	10.11	12.52			
49	0.54	0.41	0.53	11.53	14.17			
50	0.54	0.41	0.51	11.46	14.30			
51	0.50	0.49	0.46	12.29	12.59			
52	0.45	0.34	0.43	13.51	15.61			
53	0.43	0.49	0.55	13.91	12.59			
54	0.42	0.34	0.27	14.07	15.75			
55	0.39	0.37	0.38	14.58	15.02			
56	0.39	0.39	0.37	14.67	14.63			
57	0.37	0.31	0.39	15.10	16.27			
58	0.32	0.20	0.19	16.07	18.71			
59	0.31	0.27	0.30	16.40	17.26			
60	0.25	0.23	0.25	17.50	18.11			

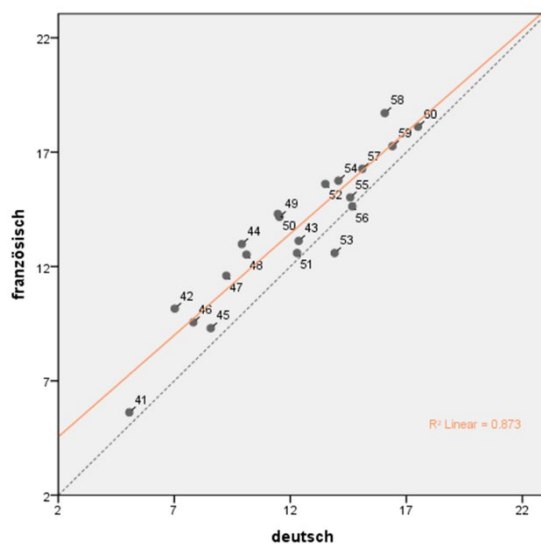


Abbildung 20: Delta-Plot für „Quantitative und formale Probleme“ (deutsch-französisch).

8.2.3.3 Textverständnis

In dieser Aufgabengruppe erscheint lediglich ein Item auffällig: Wurden in Aufgabe 73 keine Punkte erzielt, werden den entsprechenden französischsprachigen Personen 0.2 Punkte gutgeschrieben.

Tabelle 18: DIF-Analyse „Textverständnis“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
61	0.71	0.68	0.80	7.94	8.58			
62	0.50	0.48	0.51	12.33	12.72			
63	0.32	0.29	0.27	16.13	16.73			
64	0.41	0.32	0.40	14.31	16.14			
65	0.29	0.32	0.37	16.84	16.14			
66	0.54	0.56	0.63	11.54	11.14			
67	0.88	0.82	0.85	4.29	5.55			
68	0.25	0.20	0.29	17.62	18.57			
69	0.37	0.29	0.45	15.12	16.67			
70	0.43	0.49	0.53	13.74	12.66			
71	0.33	0.31	0.35	15.92	16.34			
72	0.18	0.19	0.26	19.15	18.77			
73	0.72	0.52	0.70	7.69	11.87	0.20		
74	0.27	0.27	0.23	17.28	17.26			
75	0.43	0.38	0.37	13.86	14.96			
76	0.13	0.15	0.09	20.11	19.63			
77	0.26	0.23	0.17	17.49	17.98			
78	0.26	0.29	0.23	17.32	16.67			

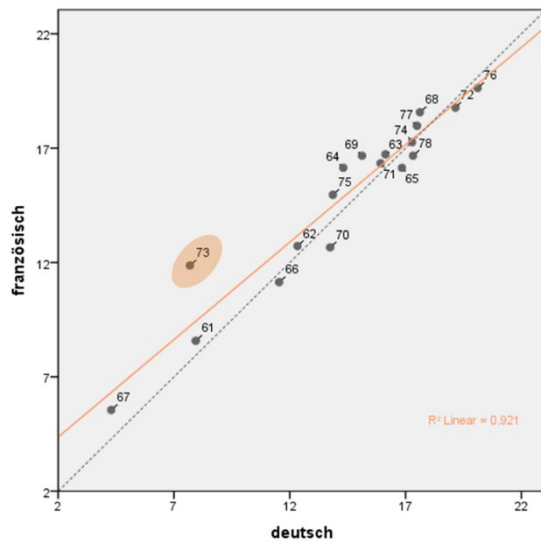


Abbildung 21: Delta-Plot für „Textverständnis“ (deutsch-französisch).

8.2.3.4 Diagramme und Tabellen

In der Aufgabengruppe „Diagramme und Tabellen“ werden der französischen Sprachgruppe bis zu 0.86 Punkte (Items 79, 83, 88) gutgeschrieben, der italienischsprachigen bis zu 0.47 (Items 79 und 85).

Tabelle 19: DIF-Analyse „Diagramme und Tabellen“

Aufgabe	Schwierigkeit			Delta-Wert			Bonus f	Bonus i
	d	f	i	d	f	i		
79	0.55	0.15	0.27	11.25	19.63		0.40	0.28
80	0.81	0.76	0.81	5.85	7.00			
81	0.78	0.68	0.79	6.47	8.51			
82	0.76	0.78	0.87	6.93	6.54			
83	0.77	0.51	0.65	6.73	12.13		0.26	
84	0.58	0.50	0.49	10.64	12.26			
85	0.66	0.53	0.47	8.98	11.73			0.19
86	0.51	0.47	0.43	12.24	12.92			
87	0.50	0.48	0.47	12.30	12.85			
88	0.45	0.26	0.35	13.31	17.46		0.20	
89	0.57	0.38	0.55	10.86	14.89			
90	0.51	0.36	0.43	12.06	15.29			
91	0.49	0.44	0.40	12.53	13.64			
92	0.35	0.31	0.25	15.44	16.40			
93	0.30	0.27	0.37	16.61	17.13			
94	0.25	0.10	0.18	17.59	20.81			
95	0.28	0.29	0.25	16.97	16.73			
96	0.14	0.16	0.17	19.87	19.56			
97	0.26	0.19	0.19	17.49	18.84			
98	0.18	0.13	0.14	19.11	20.02			

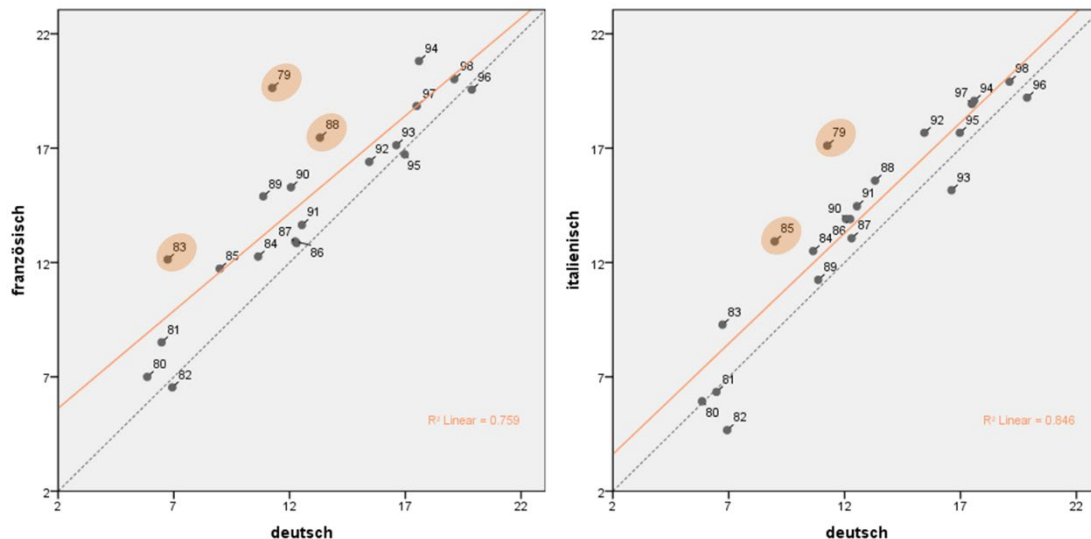


Abbildung 22: Delta-Plot für „Diagramme und Tabellen“ (deutsch-französisch und deutsch-italienisch).

8.2.4 Effekte der Korrektur

Die durch den Ausgleich resultierenden Bonuspunkte können Tabelle 21 entnommen werden. Für die Bestimmung des Punktwertes wird auf ganze Punkte gerundet. Der maximale Bonus beträgt maximal 1 Punkt. Die ungerundeten Punkteanteile wirken sich zusätzlich auf Ebene der Aufgabengruppen auf die Bestimmung des mittleren Rangplatzes aus.

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 Teilnehmenden sind dort in der Form $d^{**} > (f, i)$ angegeben. Das Beispiel würde bedeuten, dass die deutschsprachigen Kandidaten und Kandidatinnen signifikant höhere Werte erzielt haben als die beiden anderen Sprachgruppen, die sich untereinander nicht signifikant unterscheiden.

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

		n	Minimum	Maximum	m	s
französisch	Med.-naturw. Grundverständnis	319	0.00	0.35	0.21	0.12
	Textverständnis	319	0	0.2	0.10	0.10
	Diagramme und Tabellen	319	0.00	0.86	0.61	0.24
	GESAMT (Gerundet)	266	0.00	1.00	0.92	0.27
italienisch	Med.-naturw. Grundverständnis	150	0.00	0.35	0.19	0.12
	Diagramme und Tabellen	150	0.00	0.47	0.30	0.17
	GESAMT (Gerundet)	159	0.00	1.00	0.47	0.50

Nach dem Sprachausgleich bestehen keine signifikanten Unterschiede zwischen deutscher und italienischer Sprachgruppe. Beide erzielen in allen Aufgabengruppen (sprachabhängig und unabhängig) bessere Ergebnisse als die französische Sprachgruppe. Abbildung 23 zeigt die Ausmasse der (ungerundeten) Bonuspunkte für alle bisherigen Testdurchführungen.

Tabelle 21: Test-Prozentrang und Punktwerte inklusive Sprachausgleich.

		Mittelwert	Stand.abw.	Minimum	Maximum
Test-Prozentrang (d,i)** > f	deutsch	52.71	28.11	10	100
	französisch	40.29	26.98	10	99
	italienisch	49.09	27.39	10	100
Quantitative und formale Probleme (d,i)** > f	deutsch	10.29	3.73	0	20
	französisch	8.92	3.73	2	19
	italienisch	9.91	3.53	1	19
Textverständnis (d,i)* > f	deutsch	7.27	2.38	0	17
	französisch	6.91	2.23	1	13
	italienisch	7.51	2.50	1	14
Med.-naturwiss. Grundverständnis $d^{**} > f$	deutsch	9.69	3.02	0	20
	französisch	8.99	3.02	2	17
	italienisch	9.10	2.93	2	18
Muster zuordnen (d,i)** > f	deutsch	9.54	2.92	0	19
	französisch	8.84	2.90	2	17
	italienisch	9.85	2.90	1	16
Diagramme und Tabellen (d,i)** > f	deutsch	9.71	2.96	0	19
	französisch	8.37	2.68	2	16
	italienisch	8.81	2.71	3	15
Konz. und sorgf. Arbeiten (d^{**}, i^*) > f	deutsch	11.83	4.29	0	20
	französisch	10.55	4.40	0	20
	italienisch	11.69	3.82	0	20

n: deutsch 3589 / französisch 319 / italienisch 150

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

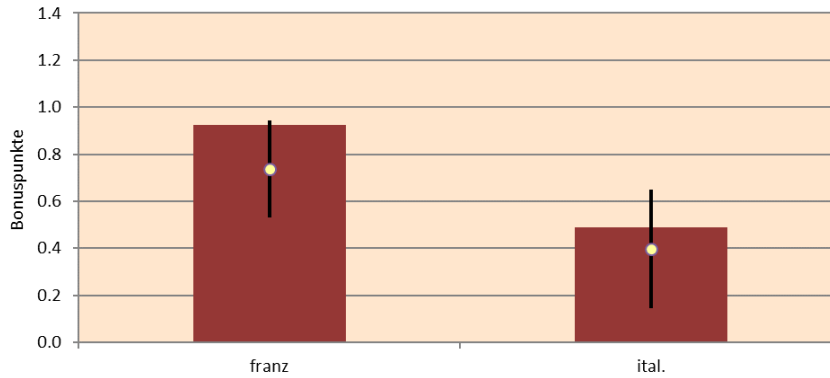


Abbildung 23: Sprachausgleich 2020 (ungerundet) mit Mittelwert und Standardabweichung aller Testdurchführungen 2010-2019

Die Differenzen zwischen der deutschsprachigen Referenzgruppe und den anderssprachigen Teilnehmenden 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.

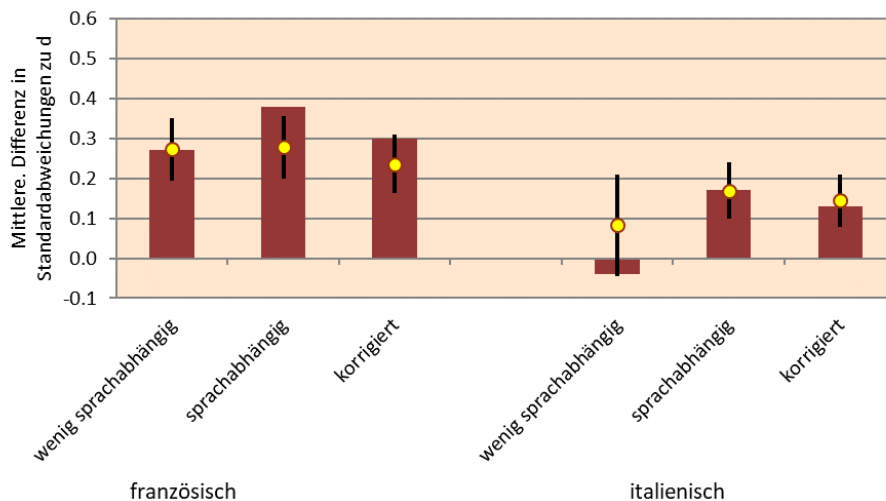


Abbildung 24: Mittlere Abweichung (Delta z) für wenig sprachabhängige und sprachabhängige Aufgabengruppen (korrigiert und unkorrigiert), 2020 und Vergleich zu 2010-2019 (Mittelwert und Streuung).

Für 2020 ist zu beachten, dass die sprachunabhängigen Aufgabengruppe lediglich noch durch „Muster zuordnen“ sowie „Konzentriertes und sorgfältiges Arbeiten“ vertreten sind. Die französische Sprachgruppe weist bei den sprachabhängigen Aufgabengruppen eine höhere Differenz aus als bei den wenig sprachabhängigen. Nach der Korrektur wird diese vermindert. Die Unterschiede zwischen den Ergebnissen von italienisch- und deutschsprachigen Personen sind insgesamt gering – sowohl in einer sprachunabhängigen Aufgabengruppe („Muster zuordnen“) wie auch in einer sprachabhängigen („Textverständnis“) erzielt die italienischsprachige Gruppe bessere Ergebnisse als die deutschsprachige.

8.3 Vergleichbarkeit der Testlokale

2020 erfolgte die Testabnahme in 34 Testlokalen/Sektoren verteilt auf acht Testorte.

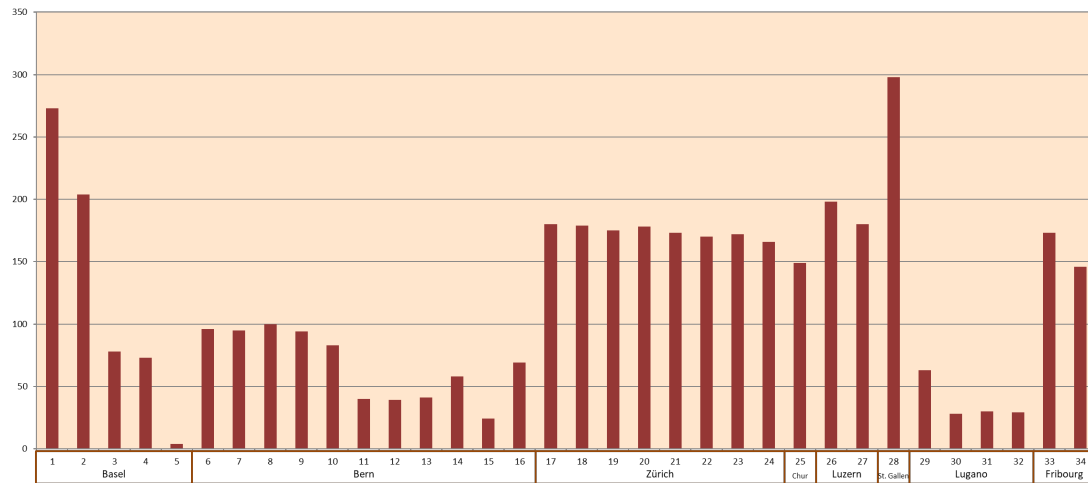


Abbildung 25: Testlokale mit Teilnehmerzahlen.

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

		Quadratsumme	df	Mittl. Quadrat. Abweichung	F	Sig.
Testprozentrang	Zwischen Gruppen	38795.850	27	1436.883	1.830	.006
	Innerhalb Gruppen	2796523.714	3561	785.320		
	Total	2835319.564	3588			
Muster zuordnen	Zwischen Gruppen	280.494	27	10.389	1.221	.199
	Innerhalb Gruppen	30290.529	3561	8.506		
	Total	30571.023	3588			
Med.-naturwiss. Grundverständnis	Zwischen Gruppen	282.325	27	10.456	1.146	.275
	Innerhalb Gruppen	32492.408	3561	9.125		
	Total	32774.733	3588			
Schlauchfiguren	Zwischen Gruppen					
	Innerhalb Gruppen					
	Total					
Quant. und formale Probleme	Zwischen Gruppen	937.531	27	34.723	2.519	.000
	Innerhalb Gruppen	49086.305	3561	13.784		
	Total	50023.836	3588			
Textverständnis	Zwischen Gruppen	128.003	27	4.741	.838	.705
	Innerhalb Gruppen	20145.309	3561	5.657		
	Total	20273.311	3588			
Figuren lernen	Zwischen Gruppen					
	Innerhalb Gruppen					
	Total					
Fakten lernen	Zwischen Gruppen					
	Innerhalb Gruppen					
	Total					
Diagramme und Tabellen	Zwischen Gruppen	365.304	27	13.530	1.546	.036
	Innerhalb Gruppen	31168.067	3561	8.753		
	Total	31533.371	3588			
Konzentr. und sorgf. Arbeiten	Zwischen Gruppen	828.262	27	30.676	1.673	.016
	Innerhalb Gruppen	65308.758	3561	18.340		
	Total	66137.020	3588			

In Basel wurde ein eigenes Testlokal für besonders schutzbedürftige Personen vorgesehen – dort nahmen vier Personen am EMS teil. Aufgrund der besonderen Massnahmen wurden grössere Abstände zwischen den Arbeitsplätzen vorgesehen, was zur Folge hatte, dass mehr und/oder räumlich grössere Lokale nötig waren.

Die Durchführungsbedingungen an den einzelnen Testorten sind standardisiert – aus den Testdurchführungsprotokollen ergeben sich keine Hinweise auf mögliche Ursachen vereinzelt auftretender Unterschiede. Es gibt keinen Anlass, Korrekturen aufgrund der Testlokalzuordnung vorzunehmen.

Tabelle 23: Darstellung der detaillierten Ergebnisse nach Testlokalen.

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	48.6	9.4	9.4		9.5	7.3			9.5	11.4	273
2	50.6	9.4	9.6		10.0	7.4			9.6	11.3	204
3	42.0	8.8	9.0		9.2	6.8			8.8	10.7	78
4	49.0	9.5	9.4		9.6	7.3			9.3	11.8	73
5	53.3	12.0	12.0		10.3	7.3			8.5	12.3	4
6	51.8	9.4	10.0		9.9	7.2			9.3	11.9	96
7	50.9	9.4	9.6		9.9	7.5			9.6	11.4	95
8	48.0	9.4	9.3		10.0	6.9			9.8	10.6	100
9	53.0	9.4	9.8		9.9	7.4			9.5	12.2	94
10	57.1	9.6	10.2		10.4	7.7			9.8	12.1	83
11	44.1	9.1	9.3		9.4	7.0			8.5	10.5	40
12	59.9	9.6	10.1		11.0	7.5			10.5	13.3	39
13	50.3	9.4	9.2		10.4	7.2			9.7	11.4	41
14	51.7	10.1	9.3		10.0	6.9			9.4	11.6	58
15	52.5	10.0	10.2		9.6	7.1			10.0	10.8	24
16	51.1	10.2	9.4		9.6	7.3			9.3	12.0	69
17	53.2	9.6	9.5		10.6	7.2			9.7	12.5	180
18	58.7	9.9	10.2		11.1	7.6			10.3	12.1	179
19	56.7	9.8	10.0		11.0	7.3			10.1	12.2	175
20	54.0	9.6	9.7		10.6	7.2			9.5	11.9	178
21	54.9	9.5	9.8		10.8	7.1			9.6	12.3	173
22	53.7	9.6	9.7		10.3	7.3			10.0	12.1	170
23	54.4	9.3	9.9		10.7	7.5			9.8	11.8	172
24	53.4	9.5	9.5		10.7	7.2			9.9	12.0	166
25	56.4	10.1	9.6		10.9	6.9			10.1	12.3	149
26	52.8	9.4	9.9		10.1	7.5			9.7	11.8	198
27	52.5	9.7	9.7		9.9	7.3			9.7	12.1	180
28	51.8	9.2	9.7		10.4	7.3			9.8	11.8	298
29	53.6	10.1	9.0		10.1	7.8			8.3	12.5	63
30	50.9	10.3	8.8		10.3	7.6			8.7	11.9	28
31	42.0	9.3	8.8		9.6	7.4			8.8	10.1	30
32	44.8	9.6	8.9		9.4	6.8			8.4	11.3	29
33	38.4	8.9	8.5		8.7	6.8			7.8	10.2	173
34	42.5	8.8	9.1		9.1	6.9			7.7	11.0	146
Total	51.6	9.5	9.6		10.2	7.2			9.5	11.7	4058

Markiert: Mittelwerte für Aufgabengruppen im multiplen Mittelwertsvergleich (Tukey), die signifikant höher als in mindestens einem anderen deutschsprachigen Testlokal liegen.

8.3.1 Versetzter Testbeginn in Zürich

Der Test fand in Zürich in vier Hallen statt, die Einhaltung der empfohlenen Abstandsregeln wäre bei 1454 KandidatInnen im Rahmen der gegebenen Lokalitäten, insbesondere während des Einlasses, nicht zu gewährleisten gewesen. Aus diesem Grund wurde entschieden, in Zürich den Test zeitversetzt in zwei Gruppen durchzuführen.

Für die Hallen 5 und 6 war der Testbeginn um 9.45 Uhr vorgesehen, für die Hallen 3 und 4 um 11.15 Uhr. Neben der eigentlichen Einlassprozedur sollten durch die versetzten Einlassungszeiten auch die öffentlichen Verkehrsmittel (z. B. betroffene Buslinien) entlastet werden.

In Halle 3 begann der Test um 11.21 Uhr, in Halle 4 um 11.31 Uhr, in Halle 5 um 10.10 Uhr und in Halle 6 um 10.09 Uhr. Grundsätzlich ist ein Testbeginn dann möglich, wenn alle erwarteten KandidatInnen anwesend und auf ihren Plätzen sind. Nach der geplanten Testzeit muss nicht mehr auf nicht-anwesende Personen gewartet werden.

Bis zum Beginn der späteren Durchführung war es in allen anderen Testorten in der Schweiz (ausser in Notfällen) untersagt, das Testlokal zu verlassen – beispielsweise bei einem Testabbruch. Dadurch wurde sichergestellt, dass keine aufgabenbezogenen Informationen an die Personen mit späterem Testbeginn weitergegeben werden konnten.

Die folgenden Analysen sollen prüfen, ob sich die Testergebnisse der KandidatInnen aus den Testlokalen mit frühem und spätem Testbeginn systematisch unterscheiden – befürchtet wurde insbesondere, dass die Gruppe mit dem späteren Testbeginn möglicherweise benachteiligt sein könnte.

Insgesamt haben 1393 Personen den EMS in Zürich mit einem gültigen Ergebnis absolviert. 2 Personen haben den Test abgebrochen (jeweils eine Person aus den Hallen 3 und 5). 59 Personen haben sich vom Test abgemeldet oder sind am Testtag nicht erschienen. 681 der gültigen Ergebnisse stammen aus Hallen mit frühem Testbeginn, 712 aus den Hallen mit spätem Beginn.

Fasst man die KandidatInnen in den Hallen nach frühem oder spätem Testbeginn zusammen, müsste sich ein allfälliger Einfluss des Testzeitpunkts entsprechend in den Testleistungen manifestieren. Abbildung 26 veranschaulicht die identischen Testergebnisse für die beiden Gruppen, mit tendenziell sogar leicht höheren (sich statistisch nicht signifikant unterscheidenden!) Ergebnissen für die spätere Gruppe.

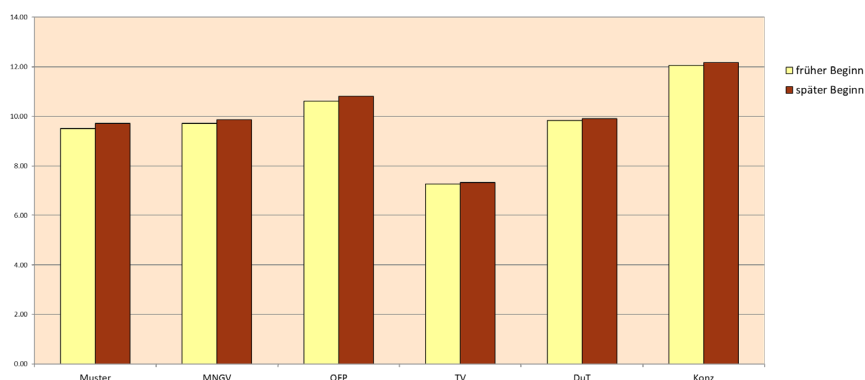


Abbildung 26: Testergebnisse nach frühem und spätem Testbeginn in Zürich.

Sowohl hinsichtlich einzelner Aufgabengruppen wie auch des Gesamtpunktwerts ist kein Einfluss des Testzeitpunkts nachweisbar. Die Unterschiede zwischen den Testlokalen bewegen sich im üblichen Rahmen und geben keinerlei Anlass, für eine der beiden Gruppen systematische Vor- oder Nachteile aufgrund der unterschiedlichen Startzeit zu vermuten.

8.4 Vergleich für die Geschlechter

Die Forderung nach Gleichbehandlung bedeutet, dass bei gleicher Eignung die gleichen Chancen auf eine Zulassung bestehen. Identische Zulassungsquoten für die Geschlechter dürfen folglich bereits aufgrund der unterschiedlichen Repräsentativität der beiden Gruppen nicht erwartet werden.

Nach 2006 hat sich beispielsweise der Männeranteil in Humanmedizin bei 40% stabilisiert, mittlerweile ist er allerdings wieder auf knapp 35% gesunken. 2020 waren 34.8% der Bewerber mit gültigem Testergebnis männlich, 65.2% weiblich.

In der Zahnmedizin sind die Verhältnisse ähnlich, mit 32.6% Bewerbern gegenüber 67.4% Bewerberinnen, während sich die Bewerbergruppe mit gültigem Testergebnis für Veterinärmedizin aus lediglich 15.2% männlichen gegenüber 84.8% weiblichen Personen zusammensetzt.

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

		Testsprache			Wunschdisziplin			Gesamt
		deutsch	französisch	italienisch	HM	VM	ZM	
männlich	m	57.4	45.8	51.6	57.4	46.7	47.4	56.3
	s	28.1	29.6	30.5	28.3	28.7	27.3	28.4
	n	1158	96	63	1173	70	74	1317
weiblich	m	50.5	37.9	47.3	51.2	44.6	35.4	49.4
	s	27.9	25.5	24.9	27.7	27.5	24.9	27.8
	n	2431	223	87	2197	391	153	2741

m: Mittelwert, s: Standardabweichung, n: Personenzahl

Für die zahlenmässig bedeutendste deutschsprachige Teilnehmergruppe wird 2020 ein Unterschied von 6.9 Test-Prozentrangpunkten registriert. Diese Differenz bewegt sich, wie erwartet, im oberen Bereich der bis anhin beobachteten Differenzen.

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

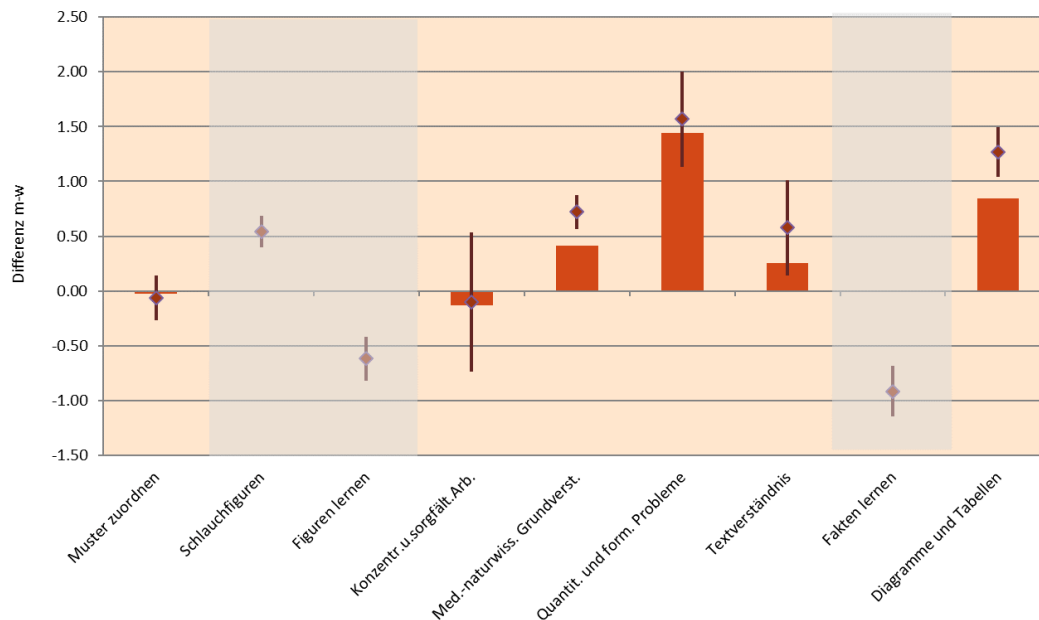
		Humanmedizin			Veterinärmedizin			Zahnmedizin		
		d	f	i	d	f	i	d	f	i
männlich	m	58.2	48.4	53.5	51.9	38.5	43.7	48.6	45.7	22.3
	s	27.9	30.4	31.1	29.2	28.1	11.9	27.8	9.8	11.0
	n	1048	68	57	42	25	3	68	3	3
weiblich	m	51.8	42.3	49.9	49.2	32.7	37.3	35.6	0.0	30.0
	s	27.8	26.3	24.8	27.8	23.6	21.2	24.9	0.0	28.8
	n	2004	122	71	278	101	12	149	0	4

Betrachtet man die Unterschiede disziplin- und sprachspezifisch, zeigt sich für die deutschsprachige Gruppe mit Wunschdisziplin Humanmedizin eine etwas geringere, aber signifikante Differenz von 6.4 Test-Prozentrangpunkten, ebenso für Zahnmedizin (13 Punkte). Alle anderen Unterschiede sind nicht signifikant.

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	0.10	0.76	6.06	3050	0.000
	VM	homogen	0.77	0.38	0.57	318	0.570
	ZM	homogen	1.77	0.18	3.45	215	0.001
Französisch	HM	nicht homogen	4.74	0.03	1.40	123	0.164
Italienisch	HM	nicht homogen	6.38	0.01	0.73	126	0.468

Abbildung 27 verdeutlicht die Vergleichbarkeit der Geschlechterdifferenzen mit den Vorjahren. Die diesjährigen Punktwert-Unterschiede auf Ebene der verbliebenen Aufgabengruppen fallen gegenüber früheren Durchführungen vergleichbar oder geringer aus (insbesondere „Medizinisch-naturwissenschaftliches Grundverständnis“, „Diagramme und Tabellen“).



(positive Werte bedeuten höhere Ergebnisse für die männliche, negative Werte für die weibliche Stichprobe)

Abbildung 27: Geschlechtsspezifische Differenzen nach Aufgabengruppen (deutschsprachig, Humanmedizin) mit Vergleich (Mittelwert und Standardabweichung) für die Vorjahre 2010-2019.

Unter den zugelassenen Personen beträgt der Frauenanteil in Humanmedizin 58.5%, in Veterinärmedizin 82.9%, in Zahnmedizin 59.5%, in Chirporaktik 56.5%.

8.5 Vergleiche nach Wunschhochschulen

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

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

	Wunschhochschule	n	Mittelwert	Standardabw.	Quadratsumme	Df	MQ	F	Sig.
HM	Basel	722	50.4	28.4	37209.9	6	6201.7	8.0	.000
	Bern	807	54.4	27.5					
	ETH	187	65.4	27.2					
	Freiburg	57	50.1	27.1					
	Zürich	1139	54.6	27.6					
	Zürich (LU)	51	55.1	27.6					
	Zürich (SG)	86	48.9	29.4					
VM	Bern	123	50.0	28.8	31.8	1	31.8	0.0	0.840
	Zürich	197	49.3	27.4					
ZM	Basel	54	39.7	27.8	495.4	2	247.7	0.4	0.704
	Bern	65	37.5	24.5					
	Zürich	98	41.1	27.1					

Unter den zugelassenen Personen treten zwischen den Wunschuniversitäten keine Unterschiede mehr auf (Tabelle 28).

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	Basel	216	85.7	9.0	738.5	6	123.1	1.6	.139
	Bern	280	84.2	8.7					
	ETH	102	86.8	7.9					
	Freiburg	16	84.3	10.2					
	Zürich	395	85.7	8.7					
	Zürich (LU)	16	87.8	8.2					
	Zürich (SG)	27	85.3	8.8					
VM	Bern	56	77.9	12.5	104.2	1	104.2	0.7	0.418
	Zürich	87	76.1	12.6					
ZM	Basel	28	62.7	18.3	287.5	2	143.8	0.5	0.633
	Bern	32	59.0	15.3					
	Zürich	52	62.5	18.7					

9 Ergebnisse zur Testgüte

9.1 Zuverlässigkeit

Die zulassungsrelevante Reliabilität (Zuverlässigkeit) über den Gesamtttest 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.

Beide Werte sind, neben der Qualität der eigentlichen Aufgaben, auch abhängig von der Aufgabenzahl (im Fall der Konsistenz des Testprofils also der Anzahl Aufgabengruppen). Aufgrund der vorgenommenen Kürzung war 2020 folglich von einer Verringerung der entsprechenden Werte auszugehen.

Der Zuverlässigkeitswert des Punktwertes erreicht 2020 mit .83 weiterhin einen hohen Wert, das Testprofil mit 0.74 eine angemessene 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 Testbatterie, 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 Zusammenhangs von Studienerfolg und Eignungstest gefunden worden sind.

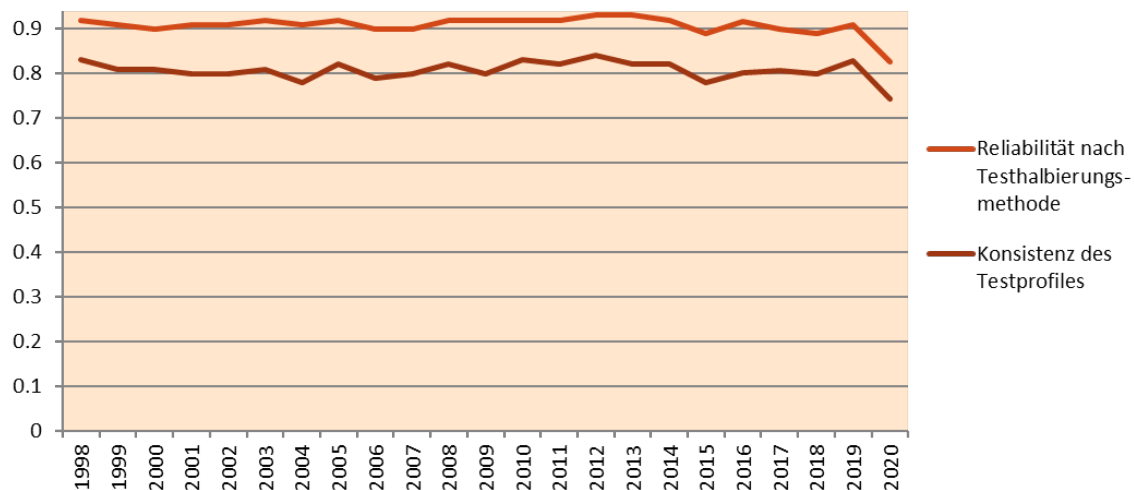


Abbildung 28: Reliabilität des Punktwertes (Split Half) und Konsistenz des Testprofils 1998- 2020.

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. Die einzelnen Aufgaben werden im Zug der Testauswertung auf die Eindeutigkeit der Fragestellung hin geprüft und bezüglich Schwierigkeit und Trennschärfe analysiert. 2020 mussten keine Aufgaben aus Qualitätsgründen aus der Wertung entfernt werden.

Insgesamt wurden aufgrund der besonderen Umstände durch die Corona-Pandemie im Interesse der Gesundheit der Teilnehmenden sowie des Testpersonals durch die Testkürzung geringe Einbussen bezüglich der Test-Reliabilität in Kauf genommen, die aber in ihrem Ausmass vertretbar erscheinen. Die Reliabilität liegt marginal unterhalb der natürlichen, jährlichen Schwankungen im Bereich „gut“, die verwendeten Aufgabengruppen weisen die geforderten Varianzen und Schwierigkeiten auf, um zwischen den KandidatInnen optimal zu differenzieren.

9.2 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.	Quant. u. formale Probl.	Textverständnis	Diagr. und Tabellen	Konzent. u. sorgf. Arbeiten	Punkt-wert
Muster zuordnen	0.24	0.26	0.19	0.25	0.30	0.55
Med.-naturw. Grundverst.		0.49	0.42	0.47	0.26	0.70
Quant. u. form. Probleme			0.37	0.56	0.31	0.76
Textverständnis				0.36	0.20	0.58
Diagramme und Tabellen					0.30	0.73
Konzent. u. sorgf. Arbeiten						0.66

Die Varianzanteile der unrotierten Lösung für die sechs statt neun Aufgabengruppen entsprechen den bekannten Grössenordnungen (Tabelle 30).

Tabelle 30: Faktorenanalyse: Verlauf der unrotierten Lösung (2016-2020).

Faktor	Unrotierte Lösung														
	Eigenwert					% Varianz					% Kumuliert				
	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020
1	3.64	3.61	3.58	3.84	2.73	40.40	40.14	39.72	42.66	45.52	40.40	40.14	39.72	42.66	45.52
2	1.44	1.31	1.12	1.40	0.94	15.96	14.59	12.43	15.51	15.63	56.37	54.72	52.15	58.17	61.15
3	0.92	0.77	0.82	0.71	0.71	10.23	8.50	9.14	7.93	11.86	66.60	63.22	61.29	66.10	73.01
4	0.71	0.71	0.71	0.63	0.67	7.89	7.91	7.92	7.05	11.17	74.49	71.13	69.20	73.16	84.17
5	0.58	0.63	0.67	0.56	0.52	6.41	6.96	7.41	6.27	8.60	80.90	78.09	76.61	79.43	92.77
6	0.50	0.56	0.58	0.52	0.43	5.60	6.18	6.41	5.75	7.23	86.51	84.27	83.02	85.17	100
7	0.48	0.53	0.56	0.50		5.32	5.83	6.23	5.56		91.82	90.10	89.25	90.73	
8	0.37	0.46	0.53	0.43		4.16	5.10	5.86	4.82		95.99	95.20	95.10	95.55	
9	0.36	0.43	0.44	0.40		4.01	4.80	4.90	4.45		100	100	100	100	

Tabelle 31: Varianzanteile der 2-Faktorenlösung 2016-2020 (rotiert).

Varimax-rotierte Lösungen														
Eigenwert					% Varianz					% Kumuliert				
2016	2017	2018	2019	2020	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020
2.8	2.5	2.4	2.7	2.3	30.9	28.0	27.0	29.8	37.7	30.9	28.0	27.0	29.8	37.7
2.3	2.4	2.3	2.6	1.4	25.4	26.7	25.2	28.3	23.5	56.4	54.7	52.1	58.2	61.2

Ein allgemeiner Fähigkeitsfaktor klärt 46% der Varianz auf – es gibt einen grossen „gemeinsamen Nenner“ in allen Aufgabengruppen. Mit der Zwei-Faktorenlösung können über 60% der Varianz aufgeklärt werden. Die Aufgabe „Konzentriertes und sorgfältiges Arbeiten“ ist 2020 deutlich dem ansonsten von den Kürzungen betroffenen Faktor „Werkzeugfunktionen“ zuzuordnen.

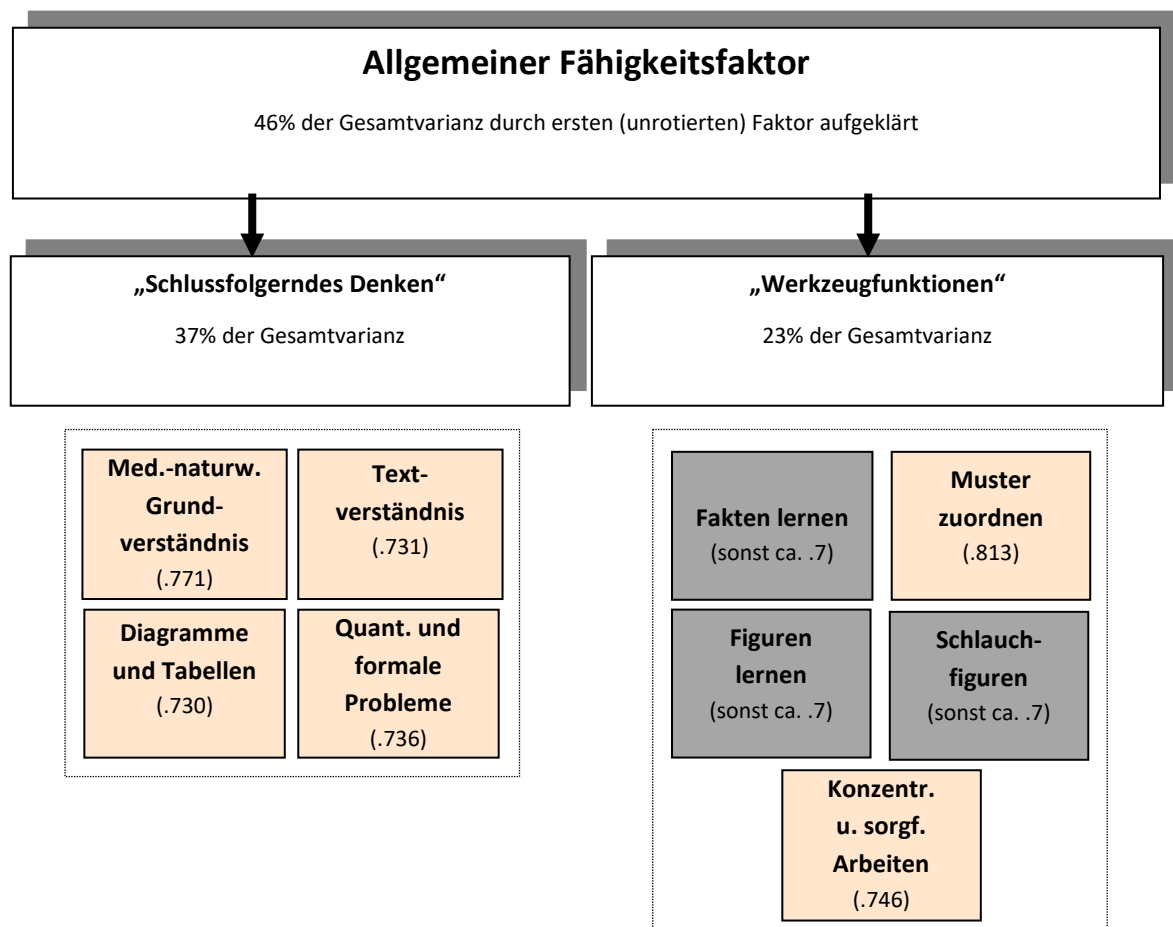


Abbildung 29: Struktur der Aufgabengruppen des EMS, hierarchische Faktorenstruktur 2020 mit Faktorladungen der Aufgabengruppen auf rotierten Faktoren.

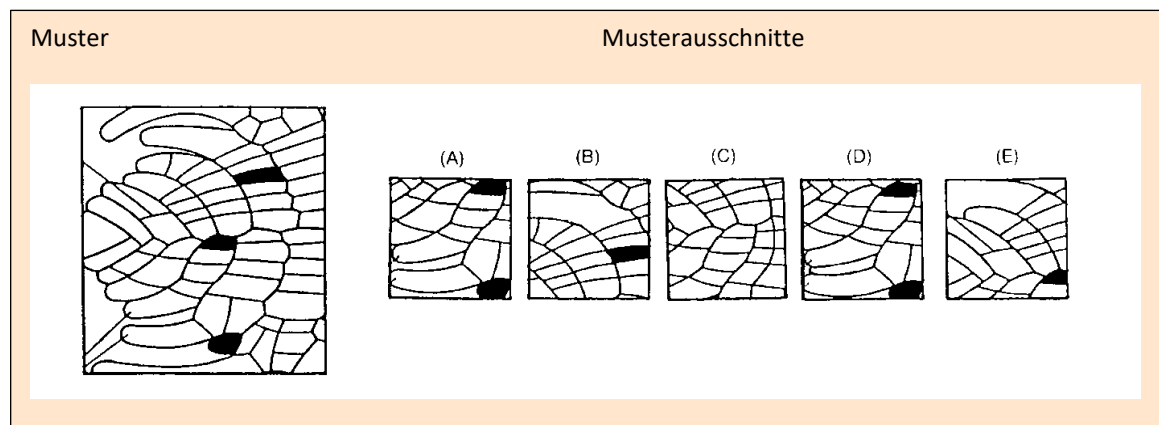
10 Beispielaufgaben für die Aufgabengruppen

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.

10.1 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.

10.2 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.

10.3 Schlauchfiguren (2020 nicht enthalten)

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 ...?

10.4 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.

10.5 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 (beispielsweise ein b mit 2 Strichen) oder eine Zeichenfolge (wenn ein u auf ein m folgt). Bei diesem Test hatte sich gezeigt, dass exzessives Üben zu verbesserten Leistungen führt. Da ein „Fleisstest“ nicht intendiert ist, werden seit 2004 die Zeichen und die Regel vor dem Test nicht mehr bekannt gegeben. Übungseffekte haben dadurch nachweislich einen geringen Einfluss.

10.6 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?

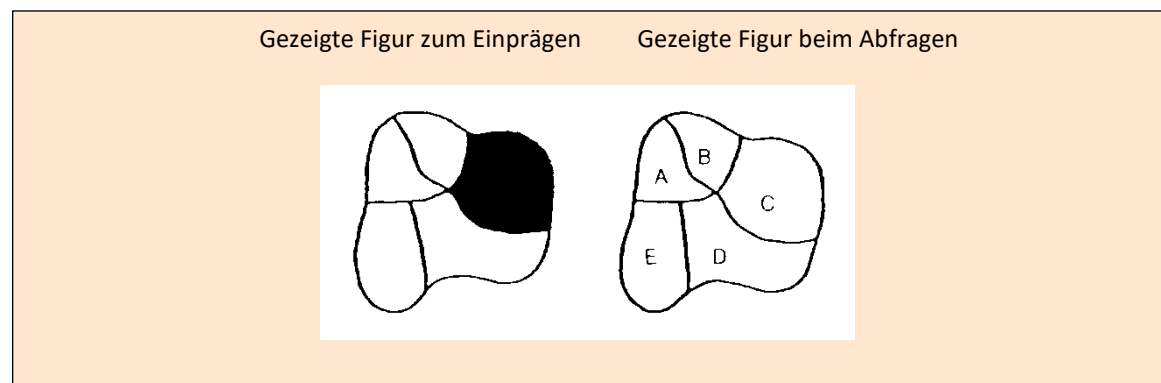
- (A) Transport von Jod aus den Epithelzellen in die Follikel
- (B) Umwandlung von Jod in Jodid in den Follikeln
- (C) Transport von Jodid aus dem Blut in die Epithelzellen
- (D) Verknüpfung von MIT- und DIT-Resten in den Follikeln
- (E) Verknüpfung von Jod und Tyrosin-Resten in den Follikeln

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.

10.7 Figuren lernen (2020 nicht enthalten)

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.

10.8 Fakten lernen (2020 nicht enthalten)

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 und Patientinnen vorgestellt, von denen jeweils der Name, die Altersgruppe, Beruf und Geschlecht, ein weiteres Beschreibungsmerkmal (z.B. Familienstand) sowie die Diagnose angegeben wird. Ein Beispiel für eine derartige Fallbeschreibung ist:

Lemke, 30 Jahre, Dachdecker, ledig, Schädelbasisbruch

Eine Frage zum obigen Beispiel könnte z.B. lauten:

Der Patient mit dem Schädelbasisbruch ist von Beruf...

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

10.9 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.

11 Literatur

Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In Holland, P.W., Wainer, H. (Eds.), *Differential Item Functioning*. Hillsdale, NJ: Erlbaum.

Beller, M., Gafni, N. (1995). Translated Versions of Israel's interuniversity Psychometric Entrance Test (PET). In T. Oakland & R.K. Hambleton (Eds.), *International Perspectives of Academic Assessment*, S.207-218. Boston: Kluwer.

Beller, M. (1996). Translating, equating and validating Scholastic Aptitude Tests: The Israeli Case. In Hänsgen, K.-D., Hofer, R. & Ruefli, D. (Hrsg). *Eignungsdiagnostik und Medizinstudium*, (S. 14-29), Bericht 2. Freiburg/Schweiz: Zentrum für Testentwicklung und Diagnostik.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (Second Edition)*. Hillsdale, NJ: Erlbaum.

Cook, L.L. (1998). *Can Scores Obtained on Test Given in Different Languages to Examinees of Different Cultures be Equally Valid?* ICAP: San Francisco.

Deidesheimer Kreis (1997). *Hochschulzulassung und Studieneignungstests: 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.

11.1 Originaltest zur Information und Vorbereitung

Es existieren in allen drei Sprachen jeweils drei 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

11.2 Frühere Berichte des ZTD

Über EMS-Seite: www.unifr.ch/ztd

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