

## Aus dem Bericht 19 des ZTD<sup>1</sup>:

### Erläuterungen zur Auswertung des EMS Fassung vom 20.10.2012

Die 10 Aufgabengruppen des Eignungstests für das Medizinstudium (EMS) ausser dem „Konzentrierten und sorgfältigen Arbeiten“ liefern Punkte als Summe richtig gelöster Aufgaben - zwischen 0 und 20 (bzw. 18 für die Aufgabengruppe „Textverständnis“).

Für die Aufgabengruppe „Konzentriertes und sorgfältiges Arbeiten“ müssen seit 2004 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 - 19 zugeordnet. Obwohl sich die Testanforderung zwischen den Jahren ändert, resultieren so vergleichbare und annähernd normalverteilte Punktwertverteilungen.

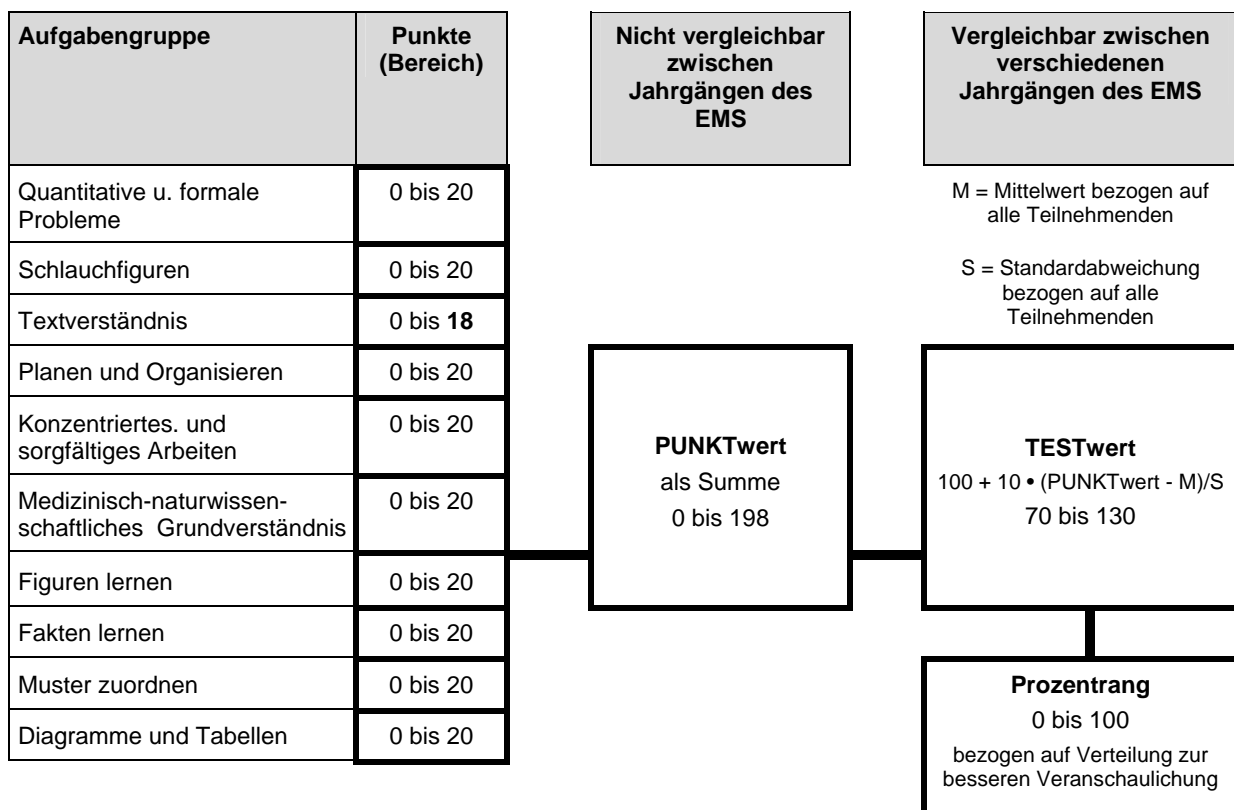


Abbildung1: Punktwerte der einzelnen Aufgabengruppen und ihre Zusammenführung über den Punktwert zum Testwert und Prozentrang.

<sup>1</sup> Berichte des ZTD: siehe <http://www.unifr.ch/ztd/ems/doc/>

Alle Punkte der Aufgabengruppen werden zu einer Summe addiert (**Punktwert**, vgl. Abbildung1). Dieser Wert hat den Nachteil, dass er nicht zwischen Tests verschiedener Jahre vergleichbar ist (diese Vergleichbarkeit ist eine Voraussetzung, dass Testwerte ins Folgejahr übertragen werden können). Deshalb findet eine Standardisierung auf den Mittelwert und die Standardabweichung aller Teilnehmenden statt.

Der resultierende **Testwert** wird auf eine ganze Zahl gerundet und im Intervall zwischen 70 und 130 verwendet (kleiner Werte werden auf 70, grössere Werte auf 130 gesetzt). Der Mittelwert des Testwertes ist in jedem Jahr 100, die Standardabweichung 10. In diesem Jahr konnten alle Aufgaben gewertet werden.

Der Testwert wird auch in einen **Prozentrangwert** umgerechnet. Prozentränge lassen sich einfach veranschaulichen, indem angegeben wird: X Prozent aller Teilnehmenden haben einen schlechteren oder gleichen Testwert erreicht als die entsprechende Person. Zu beachten ist, dass sich dieser Prozentrang auf **alle Teilnehmenden** eines Jahrganges bezieht und daher alle Disziplinen (Human- Zahn-, Veterinärmedizin, Chiropraktik) einschliesst. Deshalb stimmt dieser nicht unbedingt mit den Prozentverhältnissen der Zugelassenen/nicht Zugelassenen pro Disziplin überein. Die Teilnehmenden der einzelnen Disziplinen unterscheiden sich bezüglich des mittleren Testwertes deutlich, zusätzlich werden Personen eingereicht, die ihr Testergebnis aus dem Vorjahr übernehmen. Dies sollte insbesondere beachtet werden, wenn die Chancen einer Testwertübernahme in das Folgejahr abgeschätzt werden.

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

## Mittlerer Rangplatz der Aufgabengruppen

---

Vom Gesetz gefordert ist für die Zulassung ein Wert, welcher die Studieneignung vorhersagt. Dies erfüllt der oben beschriebene Testwert. Die Punkte aller Aufgabengruppen sind dabei gleichwertig. Das Gewicht der Aufgabengruppen in diesem Gesamtwert wird von der Varianz in diesen Aufgabengruppen beeinflusst. Da die für die Studieneignung prognoserelevantesten Aufgabengruppen in der Regel zugleich auch die höchste Streuung aufweisen, erhöht sich dadurch die Prognosekraft des Testwertes.

Eine wichtige Anforderung an des EMS ist die kapazitätsentsprechende Vergabe der Studienplätze, alle vorhandenen Plätze sollen genau ausgenutzt, die Universitäten allerdings auch nicht überlastet werden. Es ist die Regel, dass mehr Personen den niedrigsten für die Zulassung zu berücksichtigenden Testwert aufweisen als Plätze von der Kapazität her noch zur Verfügung stehen. Es hätte die Möglichkeit bestanden, beim Testwert eine Kommastelle zu berücksichtigen (in der Regel umfasst der gleiche Testwert drei Punktwerte). Dies hätte aber folgende Nachteile:

- Minimale Differenzierungen des Punktwertes hätten den Testwert sehr stark beeinflusst. Es wurde infrage gestellt, ob einzelne Punkte der verschiedenen Aufgabengruppen wirklich genügend vergleichbar sind, um auf dieser Basis so fein zu unterscheiden.
- Beim Testwertübertrag zwischen den Jahren wären für korrespondierende Punktwerte keine wirklich identischen Testwerte erreichbar.

- Die maximale Differenzierung des Testwertes reicht für eine punktgenaue Zulassung nicht aus. Bei einer Kommastelle wäre die Zulassung bereits in früheren Jahren mit weniger Anmeldungen nur auf 5-10 Personen genau möglich geworden. 2012 waren in Humanmedizin nach einer Zulassung bis einschliesslich Testwert 107 noch 5 Plätze übrig, 120 Personen hatten aber den Testwert 106. Mit einer Kommastelle hätte man immer noch 40 Personen mit einem Punktwert von 139 und Testwert 106.4 zulassen müssen oder aber die 5 Plätze nicht ausnutzen können.

Aus diesem Grund hat der Beirat EMS im Jahr 2005 beschlossen, bei **gleichem Testwert** die Personen in der Reihenfolge des **mittleren Rangplatzes aller Aufgabengruppen** zu reihen und so einen etwas anderen Auswertaspekt zu berücksichtigen (in durchaus gewollter Analogie zum Prinzip von Sportarten, wo bei gleicher Punktezahl dann das Torverhältnis zählt).

Bis zum vorletzten für die Zulassung berücksichtigten Testwert sind Mittlere Rangplätze unerheblich. Nur wenn die Zahl der Personen mit dem letzten zu berücksichtigten Testwert grösser ist als die verbleibende Kapazität (wie 2012 in Humanmedizin für den Testwert 106), wird der mittlere Rangplatz zur weiteren Differenzierung **innerhalb** dieses Testwertes verwendet. Es werden pro Aufgabengruppe den einzelnen Punktwerten die Rangplätze zugeordnet, die sich aufgrund der Häufigkeitsverteilung aller Teilnehmenden eines Jahrganges ergeben.

Diese Rangplätze werden dann über alle 10 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 Punktezahl ist und je weniger Personen insgesamt einen besseren Punktwert erreicht haben. Rangplätze belohnen gute Leistungen in einer schwierigen Aufgabengruppe stärker, indem dort niedrigere Rangplätze für die besten Leistungen resultieren.

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

In Einzelfällen kommt es daher auch vor, dass eine zugelassene Person einen Punkt mehr erzielt hat als eine nicht zugelassene Person – wenn letztere den besseren mittleren Rangplatz erzielt und die Punkte in schwereren Aufgabengruppen erzielt hat. Die Verwendung der Punktwerte für die Zulassung würde keinen Testwertübertrag aus dem Vorjahr ermöglichen und für die notwendige kapazitätsgenaue Zulassung nicht ausreichend differenzieren.

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

- Der Rangplatz bewegt sich z.B. 2012 zwischen 35 und 977. Niedrige Werte stehen für bessere Leistungen.
- Er korreliert mit dem Testwert über die Jahre zwischen -0.98 und -0.99 was praktisch für eine Äquivalenz spricht.

### **Beispielrechnung**

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

#### **Aufgabengruppe 1:**

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

#### **Aufgabengruppe 2:**

- In dieser Aufgabengruppe erreicht nur Kandidat B die volle Punktzahl von 20 Punkten. Er erhält als einziger den Rang 1.
- Kandidat A erreicht hier 19 Punkte, zusammen mit nur weiteren 4 Kandidaten. Sie alle erhalten in dieser Aufgabengruppe den durchschnittlichen Rang  $(2+5)/2= 3.5$ .

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

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

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

Und auf 1000 Personen:

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

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

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

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

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