

## Rückblick von aussen: Was hat das ZTD so erfolgreich gemacht?

Günter Trost

Geschätzte Festgäste, liebe Kolleginnen und Kollegen

Es ist mir eine Ehre, an diesem Fest-Tag ein paar Beobachtungen und Anmutungen unter der Perspektive des Zuarbeiters und Begleiters aus dem grossen Nachbarkanton beisteuern zu dürfen. Und es ist mir ein Herzensanliegen, dem ZTD zu seiner überaus erfolgreichen Arbeit über ein Vierteljahrhundert zu gratulieren. Wie erfolgreich diese war, will ich an einigen Beispielen belegen.

Die turbulente Vorgeschichte, die ebenso turbulente Anfangszeit und die „institutionelle Geschichte“ des Zentrums insgesamt hat Meinrad Perrez eindrucksvoll geschildert.

Ich richte meinen Blick und lenke den Ihrigen auf die eher operativen Aspekte. Zwei Arten von Herausforderungen hatte und hat das ZTD zu meistern: wissenschaftliche und administrative. Beide waren und sind gleichermaßen entscheidend für den Erfolg des zentralen Projekts dieser Institution: des Eignungstests für das Medizinstudium (EMS).

Unter den zahlreichen **wissenschaftlichen** Herausforderungen will ich drei hervorheben:

- An erster Stelle ist die Verpflichtung zu nennen, die Brauchbarkeit des Tests zu prüfen. Der Druck, die Überlegenheit des EMS gegenüber allen anderen möglichen Auswahlverfahren zu sichern und empirisch nachzuweisen, ist in der Schweiz noch höher als in Deutschland, hat hierzulande der Test als einziges Auswahlinstrument doch ungleich höheres Gewicht gegenüber jenem des Tests für medizinische Studiengänge (TMS), der eines unter mehreren Auswahlverfahren ist. (Die Abiturdurchschnittsnote beispielsweise muss in Deutschland laut Gesetz für die Zulassungsentscheidung stets ein, zumindest minimales, Übergewicht haben.)

Wichtigstes, wenn auch nicht einziges Kriterium der Brauchbarkeit eines Hochschuleingangstests ist die Vorhersagekraft des Testergebnisses bezüglich des Studienerfolgs, seine prognostische Validität also. Das ZTD hat alle Möglichkeiten genutzt, diese Eigenschaft zu evaluieren. Wir wissen aus den Validitätsstudien unseres Hauses, des ITB, wie mühsam dieses Geschäft ist. Mit Schmunzeln erinnere ich mich an die Zeiten, in denen wir in Deutschland Mitarbeiterinnen und Mitarbeiter des ITB an die einzelnen medizinischen Fakultäten entsandten, wo sie auf der Suche nach Examensdaten zum Teil buchstäblich in überwiegend wenig gepflegten Zettel-Kästen mit den individuellen Prüfungsergebnissen wühlen mussten (und durften, was heute undenkbar ist). Nur so konnten damals die Testergebnisse der Einzelpersonen mit deren Prüfungsergebnissen im Studium zusammengeführt werden.

Das ZTD hat alle modernen wissenschaftlichen und organisatorischen Möglichkeiten genutzt, die Vorhersagekraft des EMS zu bestimmen. Fast ideale Untersuchungsbedingungen fanden sich bei der ersten Kohorte von Teilnehmenden am Test des Jahres 1998, von denen 89 Prozent anschliessend zum Studium zugelassen worden waren. So konnten auch die Studienergebnisse jener Personen erhoben werden, die im Test schlechter abgeschnitten hatten. (Später erlangten aufgrund des zunehmenden Selektionsdrucks nur noch Personen mit besseren Testleistungen einen Studienplatz.) Die Studierenden wurden bis zur 2. Vorprüfung in der Humanmedizin „verfolgt“. Die Ergebnisse stehen im Einklang mit jenen, die wir in Deutschland gefunden haben, und brauchen den Vergleich mit Resultaten aus anderen Ländern, die ähnlich anspruchsvolle Auswahltests verwenden, nicht zu scheuen. Kurz gefasst, besagen sie: Wer im Test besser abschneidet, hat deutlich höhere Chancen als Personen mit niedrigeren Testergebnissen, die Vorexamina in der vorgeschriebenen Mindest-Studienzeit abzulegen, die Prüfungen zu bestehen und bessere Examensnoten zu erzielen.

- Hervorzuheben ist zweitens die umfassende wissenschaftliche Aufbereitung und die sorgfältige Dokumentation der Ergebnisse eines jeden Testdurchgangs durch das ZTD. Die regelmässigen Untersuchungen schliessen nicht nur Vergleiche der Testleistungen der verschiedensten Gruppen von Teilnehmenden ein, sondern auch die Analyse der Funktionsweise des Gesamttests, der verschiedenen Aufgabengruppen und jeder einzelnen Aufgabe, insbesondere bezüglich des Schwierigkeitsgrads und der Messgenauigkeit. Zudem wird in Strukturanalysen ermittelt, welche übergeordneten kognitiven Fähigkeiten hinter den Leistungen in den einzelnen Aufgabengruppen des EMS stehen. Diese – sämtlich veröffentlichten – Auswertungen dienen nicht nur der Transparenz des Verfahrens. Sie gewährleisten darüber hinaus eine kontinuierliche Qualitätskontrolle und bieten die Garantie, dass etwaige Fehlentwicklungen bei einem Test, der ja jedes Jahr aus komplett neuen Aufgaben besteht, frühzeitig erkannt und korrigiert werden können.

- Erwähnt seien drittens die wichtigen empirischen Studien des ZTD zur Trainierbarkeit des EMS. Über mehrere Jahrgänge wurden die Teilnehmenden am EMS in sehr differenzierter Weise nach Art und Umfang ihrer Vorbereitung auf den Test befragt und die Antworten in Beziehung zum erlangten Testergebnis gesetzt. Die Befunde sind von hoher politischer und praktischer Relevanz. Sie brachten wertvolle Aufschlüsse über sinnvolle und weniger sinnvolle Aktivitäten zur Vorbereitung. Und sie räumten – wie zuvor schon unsere Untersuchungsergebnisse in Deutschland – mit der von interessierter Seite gerne verbreiteten Legende auf, der Besuch eines teuren kommerziellen Trainingskurses garantiere quasi den Erfolg im „Ernsttest“.

Wenn die Repräsentanten des ZTD auf internationaler wissenschaftlicher Bühne Methodik und Ergebnisse ihrer Untersuchungen vorstellten, war ihnen die Aufmerksamkeit der Fachkollegen gewiss und sie gaben „bella figura“ ab.

Nicht weniger anspruchsvoll sind die **administrativen** Herausforderungen, denen sich das Projekt EMS seit seiner Gründung konfrontiert sieht.

- Es ist zum einen die Aufgabe, einen Test, von dessen Ergebnis so viel für die Teilnehmenden wie für die Gesellschaft insgesamt abhängt, an einem bestimmten Tag zur selben Stunde in rund drei Dutzend Testlokalen mit mittlerweile mehr als 4000 Teilnehmenden unter genau gleichen Arbeitsbedingungen durchzuführen. Das bedeutet generalstabsmässige Planung und Organisation, eine minuziöse Schulung der Aufsichtspersonen, präzise, einheitliche Regelungen für alle erdenklichen Zwischenfälle einschliesslich der Kompensation allfälliger individueller Benachteiligungen durch Störungen oder persönliche Beeinträchtigungen. Dass dies über all die Jahre reibungslos funktionierte, ist eine Meisterleistung.
- Eine zweite Herausforderung – von der wir bei unserem deutschen Pendant des Tests verschont sind – besteht in der Notwendigkeit, den EMS in drei Sprachen auszulegen. Hier ist es mit einer wortwörtlichen Übersetzung der von uns gelieferten deutschen Aufgabentexte ins Französische und Italienische nicht getan. Mit einem ausgefuchsten Verfahren gelingt es dem ZTD, die „funktionelle Äquivalenz“, wie ich das gerne nenne, zu wahren, d.h. sicherzustellen, dass die Aufgaben in allen drei Sprachen das Gleiche und dies auf gleichem Schwierigkeitsniveau messen. Wenn trotz aller Vorkehrungen einige wenige Aufgaben in einer der Versionen, aus welchem Grund auch immer, aus dem Rahmen fallen, werden sie dank einer Methode zur Bestimmung des „Differential Item Functioning“ (kurz: DIF) identifiziert. Diese Methode zeigt an, wenn bei einer Aufgabe Unterschiede zwischen den Sprachgruppen in den Ergebnissen nicht auf unterschiedlich ausgeprägte Fähigkeiten zurückzuführen sind, sondern auf eine Besonderheit dieser Aufgabe in einer der drei Sprachversionen. Die so erkannte Abweichung kann mühelos rechnerisch korrigiert werden. Auch hier steht das ZTD im internationalen Vergleich an vorderster Front.
- Die wohl grösste Herausforderung liegt in der Sicherstellung der absoluten Geheimhaltung der Testaufgaben vor und nach deren Einsatz am Testtag. Es würde hier zu weit führen, all die Abschottungen und Kontrollen aufzuführen, die dazu nötig sind. Das Sicherheits-Management des ZTD hat dafür gesorgt, dass von 1998 bis 2019 kein einziges Leck in diesem komplexen System aufgetreten ist. Ein grandioser Erfolg, um den so manches grosse Testinstitut, z.B. in den Vereinigten Staaten, die Schweiz nur beneiden kann. Dass es gleichwohl zu einer Sicherheitspanne kam, hatte seine Ursachen allein in Deutschland.

Vertrauensvolle Zusammenarbeit wie die zwischen dem ZTD und seinem siamesischen Zwilling ITB bewährt sich vor allem in Krisen. Eine Krise gab es in der Tat, und zwar in einem Ausmass, das um ein Haar die Hoffnung auf eine Zukunft des EMS und damit möglicherweise auch auf den Fortbestand des ZTD zunichte gemacht hätte. Dank eines couragierten Whistleblowers, der an einem der kommerziellen Trainings zum EMS und anschliessend am Ernsttest teilgenommen hatte, wurde bekannt und belegbar,

dass der betreffende Trainingsanbieter Aufgaben des Originaltests verwendet hatte. Diese konnten mithin im Ernsttest nicht gewertet werden. Der Vorfall löste ein gewaltiges Echo in den Medien, der Öffentlichkeit und der Politik aus. Blitzschnelles Handeln auf Seiten des ZTD und des ITB, das den beschädigten Test ja geliefert hatte, war das Gebot der Stunde. Das ZTD entwickelte sogleich ein korrigiertes Auswertungsprogramm und Modelle, mit denen die Auswirkungen der Nicht-Wertung der bekannt gewordenen Aufgaben ermittelt werden konnten. Die Abstimmung mit dem ITB erfolgte fast im Stundentakt. Wir eruierten zunächst, woher die widerrechtlich im Trainingskurs verwendeten Aufgaben stammten. Sie waren 1996, am letzten Testtermin des damals gültigen Mediziner-Zulassungsverfahrens in Deutschland (inzwischen wird der Test wieder verwendet), trotz aller Sicherheitsmassnahmen unter bis heute ungeklärten Umständen entwendet worden, und zwar in Form einzelner ganzer Seiten des Testhefts. Zum Glück hatten wir aus jener Version dem ZTD nur eine kleine Zahl an Aufgaben als Ersatz für veraltete Aufgaben überlassen. Der konkrete Schaden hielt sich somit in Grenzen, der Imageschaden war ungleich grösser.

Zu allem Unglück zeigte sich im Jahr darauf, dass weitere Originalaufgaben aus dem TMS, wenn auch in fehlerhafter Wiedergabe, in kommerziellen Trainingskursen verwendet worden waren. In fieberhafter Tag- und Nacharbeit spürten wir die Herkunft jener Aufgaben auf, wobei wir zusätzlich jedes Trainingsheft der kommerziellen Anbieter, dessen wir habhaft werden konnten, Aufgabe für Aufgabe auf Übereinstimmungen mit „echten“ TMS- bzw. EMS-Aufgaben prüften. Das gemeinsame Krisenmanagement funktionierte, beschädigte Aufgaben wurden identifiziert und Ausgleichsmassnahmen mit swissuniversities abgestimmt. Als wichtigste Konsequenz aus diesen Vorfällen wurde beschlossen, dass die Aufgaben im EMS jeweils nur noch ein einziges Mal verwendet werden.

Mit welcher Dreistigkeit einzelne Trainingsanbieter vorzugehen versuchen, zeigt die folgende Anekdote: Vor einigen Jahren stellten wir fest, dass einer dieser Anbieter – wir kennen sie alle - sich zur Teilnahme am TMS angemeldet hatte. Im Testlokal, dem er zugewiesen war, sorgten wir dafür, dass ein ITB-Mitarbeiter am Tisch neben ihm sass. Dem fiel auf, dass der oberste Hemdenknopf des Beobachteten etwas anders aussah als die übrigen. Es war eine Knopfkamera. Die Staatsanwaltschaft tat, was zu tun war.

An einem Tag wie diesem soll die **Perspektive der Bewerberinnen und Bewerber** nicht zu kurz kommen. Wie das ZTD die Interessenten und Interessentinnen für ein medizinisches Studium im Vorfeld der Bewerbung unterstützt, ist vorbildlich: Drei komplette Originalversionen des Tests – die wir auch in Deutschland verwenden – sind veröffentlicht, und zwar nicht nur in deutscher, sondern in der Schweiz auch in französischer und italienischer Sprache. Ein ausführliches Test-Info, ebenfalls in drei Sprachen, enthält alle wichtigen, z.T. sehr praktischen und konkreten Tipps zur Vorbereitung auf den Test und zum sinnvollen Vorgehen bei der Testbearbeitung. Ferner hält das ZTD detaillierte Informationen über die medizinischen Studiengänge im Internet bereit. Und schliesslich bietet das kostenlose Online-Self-Assessment den Ratsuchenden

Gelegenheit, sich über zentrale Aspekte ihrer Persönlichkeit, über ihre Motive und Interessen sowie über ihre Erwartungen – oder auch Fehl-Erwartungen – an das Medizinstudium klar zu werden. Eine mehr als 30-seitige individuelle und sehr differenzierte Rückmeldung gibt das Ergebnis der Auswertung wieder. Nur wenige andere Länder bieten den Ratsuchenden so viel Unterstützung von offizieller Seite; vielfach wird dies kommerziellen Vorbereitungs-Instituten überlassen.

Ich habe die Qualität der Arbeit des ZTD gewürdigt. Nun sollte ich wohl ein paar Worte darüber verlieren, warum es genau jenes Verfahren ist (eben der EMS), auf das sich dieses erfolgreiche Wirken richtete. Schliesslich gibt es eine Vielfalt an Verfahren und Instrumenten für die Auswahl bei der Zulassung zur Hochschule. In der hitzigen politischen Diskussion Mitte der neunziger Jahre wurden denn auch alle möglichen Modelle aufs Tapet gebracht, wobei lästige Fragen nach der tatsächlichen Aussagekraft der propagierten Verfahren oder nach deren Durchführbarkeit zum Teil souverän ignoriert wurden. Ich erspare Ihnen die Aufzählung all dieser Optionen; sie sind in den Schriften des ZTD gut dokumentiert. Doch schlussendlich setzten sich Augenmass, Besonnenheit und die Berücksichtigung wissenschaftlich-empirischer Evidenz durch, und die Entscheidung fiel zugunsten eines studienfeldspezifischen Studierfähigkeitstests. Weil es einen solchen (den TMS, Test für medizinische Studiengänge) im Nachbarland bereits gab, nahm man diesen genauer unter die Lupe. Und stellte fest, dass er zu diesem Zeitpunkt schon in einem für die deutsche Bildungsgeschichte einmaligen Umfang beforscht worden war. Das heisst, es lagen Untersuchungsergebnisse vor – grossenteils von unabhängigen externen Experten ermittelt – zu praktisch allen kritischen Fragen, die an ein derartiges Verfahren gestellt werden müssen: Ist der Test trainierbar? Diskriminiert er bestimmte Bewerbergruppen wie etwa Frauen? Misst er tatsächlich etwas Anderes als die Schulnoten? Lässt er sich unter exakt gleichen Bedingungen mit grossen Teilnehmerzahlen durchführen? Vor allem aber: Hilft er wirklich, diejenigen Bewerberinnen und Bewerber zu identifizieren, die anschliessend auch die besseren Studienleistungen erbringen? Antworten auf all diese und etliche weitere Fragen lagen vor und überzeugten auch die Schweizer Verantwortlichen. So konnte das ZTD auf ein bewährtes Verfahren zurückgreifen – das es mittlerweile signifikant weiterentwickelt hat –, und eine fruchtbare Zusammenarbeit mit dem deutschen Anbieter, ITB, nahm ihren Anfang.

Diese Zusammenarbeit mit dem ZTD, hier vor allem mit Klaus-Dieter Hänsgen, Benjamin Spicher und Tanja Läng/Cruchaud, ebenso wie mit den Verantwortlichen in SHK, CRUS, SUK und swissuniversities und mit den Mitgliedern des Beirats des EMS gehört zu den schönsten und den besonders bereichernden Erfahrungen meiner beruflichen Laufbahn. Sie war geprägt von der gemeinsamen Begeisterung für die Sache, von Verantwortung gegenüber den betroffenen Personen und Institutionen, von Kollegialität und, nicht zuletzt, von grossem wechselseitigen Vertrauen.

Natürlich gab es in den 25 Jahren der Kooperation auch kleine Konflikte zwischen ZTD und ITB – wie könnte es anders sein? In den wenigen Fällen, in denen ein solcher aufbrach, lag

es an einer Inkongruenz Schweizer und deutscher Interessen. Stets aber wurde er mit Fairness, gegenseitigem Respekt und Bereitschaft zum Kompromiss ausgetragen und entsprechend schnell gelöst.

Liebes Zentrum für Testentwicklung und Diagnostik, lass mich dir danken für die wunderbare Zusammenarbeit über ein Vierteljahrhundert und dir viele, viele weitere Jahre erfolgreichen Wirkens wünschen. Deine Zukunft liegt, davon bin ich überzeugt, unter Petra Klumbs und Benjamin Spichers kundiger Führung in den allerbesten Händen.