

Warum wird im EMS kein „Situational Judgement Test (SJT)“ verwendet? Ergebnis einer Evaluation

Prof. Dr. K.D. Hänsgen

1. Alle uns bekannten Länder, die Eignungstests für die Studienzulassung in der Medizin einsetzen, bemühten sich von Beginn an, **soziale Kompetenzen mit zu erfassen** – z.B. wurden in Deutschland beim TMS mehrere Konzepte erprobt, letztendlich mussten sie mangels Eignung und Güte verworfen werden. Auch bei der Weiterentwicklung des EMS¹ ist dies ein ständiges Thema, tatsächlich verwendbare Konzepte sind rar.
2. In Österreich besteht zusätzlich Handlungsdruck, weil man sich davon die Verringerung des „**Gender Gap**“ **erhofft**: Wenn Frauen hier bessere Leistungen erzielen würden als Männer, könnten vorhandene Unterschiede bei anderen Fähigkeiten in den Zulassungstests ausgeglichen werden. Die Erfüllung dieser Hoffnung ist nicht sicher – nicht alle Befunde belegen bessere Leistungen für Frauen. Insofern sollten hier unbedingt falsche Versprechungen vermieden werden.
3. **Ein mögliches Konzept zur Erfassung sozialer Kompetenzen sind sogenannte „Situational Judgement Tests“ (SJT)**. Es werden soziale Situationen vorgegeben, die man verschieden wahrnehmen und auf die verschieden reagiert werden kann. Es werden Fragen gestellt und aus vorgegebenen Alternativen ist eine richtige Antwort auszuwählen.
4. Der bisher einzige **Praxiseinsatz eines SJT für Studienzulassung** erfolgt nach unserer Kenntnis in **Belgien**. In den ersten Versionen wurden **Videosequenzen** zur Situationsvorgabe verwendet, die dann entsprechend einzuschätzen waren. Weil es während einer Testabnahme zu einem Technikausfall kam und die jährlich neue Produktion solcher Sequenzen auch teuer war, ist man zu **Papierversionen** übergegangen: Situationsbeschreibungen und mögliche Antworten werden jetzt auf Papier vorgegeben.²
5. Die **Evaluation** beider Fassungen in Belgien ist ausführlich erfolgt³: Videobasierte Versionen korrelieren weniger mit rein kognitiven Fähigkeiten, messen wirklich andere Aspekte. Wichtiger ist die **Studienerfolgsprognose** als Voraussetzung einer Verwendung für die Studienzulassung. Die Korrelationen mit zwei **Kriterien** wurden berechnet: (1) Dem „*Grade Point Average*“ (*GPA*) als Punktwert der eher „kognitiv orientierten“ Prüfungsleistung nach dem ersten Jahr sowie (2) der *Note aus „interpersonally orientated courses“*, wo interpersonelle und kommunikative Fähigkeiten die Bewertung beeinflussen.

Die videobasierte Form korreliert zu 0.07 mit GPA und 0.34 mit dem interpersonellen Kriterium (letzteres ist signifikant). Die Papierversion korreliert zu 0.10 mit GPA und nur zu 0.08 mit dem Kriterium „interpersonelle Fähigkeiten“.

Diese Korrelationen rechtfertigen es für die Papierversion nicht, von einer verwendbaren Prognosegüte auszugehen und sie ist der Videoversion unterlegen, wo die Korrelation

¹ Eignungstest zum Medizinstudium (auch in Innsbruck und Wien eingesetzt) siehe www.unifr.ch/ztd/ems

² Information F. Lievens im Workshop zum SJT, Medizinische Universität Wien 6.3.2009

³ Lievens, F, Sackett, P.R. (2006): Video based versus written situational judgement tests: A comparison in terms of Predictive Validity. J. Appl. Psychology. 91 1181-1188
<http://users.ugent.be/~flievens/video.pdf>

mit dem interpersonellen Kriterium einen Einsatz noch rechtfertigte. Zum Vergleich: Die Korrelationen des EMS mit Studienerfolgskriterien liegen in Bereichen um 0.50.

6. Videosequenzen funktionieren besser, weil soziale Kompetenz sich mindestens auf zwei Komponenten bezieht: (1) richtiges **Erkennen** und Bewerten sozialer Situationen und (2) Wissen über richtiges **Reagieren** auf die erkannten Anforderungen. Bei Videopräsentationen dürften die Bewertungsprozesse komplexer sein, man muss z.B. verbale und auch nonverbale Signale richtig erkennen. Eine „Papierbeschreibung“ der Situation ist weniger komplex und nur verbal, die Leistung bezieht sich stärker auf die Wahl der richtigen Antwort, also „soziales Wissen“. Man kann nur vermuten, warum der SJT in Papierform in Belgien trotzdem weiter verwendet wird. Eigentlich wäre sein Einsatz nicht gerechtfertigt, solange man nicht weiss, was der Test eigentlich misst und vorhersagt.
7. 2009 sollte mit den Medizinischen Universitäten Innsbruck und Wien ein **Pilotprojekt** mit einem SJT gestartet werden. Weil für alle Personen eine gleichgute Videopräsentation kaum realisierbar ist, kam nur eine Papierversion in Betracht. Dabei sollten Ergebnisse erst für die Zulassung verwendet werden, wenn **gesichert** ist, was der Test tatsächlich misst. Dazu wurde ein Workshop mit Prof. Dr. F. Lievens, dem Autor des belgischen Tests am 6.3. 2009 in Wien durchgeführt, wo *alle* Medizinischen Universitäten Österreichs vertreten waren. Ausserdem erfolgte die Konsultation eines hinsichtlich psychodiagnostischer Kompetenz international ausgewiesenen potentiellen Kooperationspartners, der Test- und Beratungsstelle des AB Psychologische Diagnostik der Fakultät für Psychologie der Universität Wien (Univ.Prof. Dr. K. D. Kubinger). **Es musste gemeinsam entschieden werden, dass ein solches „Experiment“ aufgrund des vorhandenen Wissens und der konzeptionellen Einwände nicht verantwortbar ist.**
8. Eignungstests als „wettbewerbsorientierte Reihungsverfahren“, mit denen wichtige Entscheidungen getroffen werden, müssen **eindeutig bewertbare Antworten** für jede Frage haben. Wer mehr richtige Antworten gibt, muss über bessere Fähigkeiten und damit Studieneignung verfügen. Wenn Antworten **als „falsch“** oder weniger zutreffend bewertet werden, muss dies **nachweislich** so sein, weil ansonsten jedes Testergebnis angefochten werden kann - **Objektivität bzw. Willkürfreiheit** nicht mehr gegeben wären.
9. Wie werden bei einem SJT die Situationen und Antwortalternativen konstruiert?⁴. Üblicherweise werden **Expertengruppen** für die Entwicklung und Bewertung der Fragen herangezogen und deren **Beurteilerübereinstimmung** ist das wichtigste Kriterium für den Nachweis, dass eine Antwort richtig ist.
 - a. Sind die Situationen speziell auf **Medizin zugeschnitten** und setzen sie Fachwissen voraus, werden Personen ohne dieses medizinische Vorwissen **benachteiligt**. Vorausgesetzt werden darf praktisch nur der Kenntnisstand der Maturität. Der Test würde stark trainierbar, wenn man Fachwissen vorab erwerben müsste.
 - b. Entsprechen die Situationen **alltäglichen Erfahrungen** und sind sich die **Experten** bei der Bewertung der Situationen **einig**, werden praktisch automatisch Antworten nach der sozialen Erwünschtheit einer Reaktion erwartet - unabhängig davon, ob man selbst **tatsächlich** so reagieren würde. Es wäre dann eher ein „Gesinnungstest“ und die Aufgaben würden auch viel zu leicht.
 - c. Wenn man Situationen und Antworten **undurchschaubarer und dadurch mehrdeutiger** gestaltet, sind sich auch die Experten **nicht mehr einig**, was die richtige Antwort ist. In Belgien wird so verfahren und man akzeptiert, dass minimal 60% der Experten eine Antwort als richtig bewerten. Wenn aber bis zu 40% der Experten eine andere Antwort als richtig ansehen, ist die Forderung nach Willkürfreiheit nicht mehr erfüllbar und Testergebnisse wie das gesamte Zulassungsverfahren wären zu Recht anfechtbar.

⁴ F. Lievens Handout zum Workshop zum SJT, Medizinische Universität Wien 6.3.2009

Wenn auch nur ein Experte etwas als richtig ansieht: Wieso sollte die Person, welche diese Antwort im Test dann wählt, dafür nicht auch einen Punkt erhalten? Soziale Sachverhalte sind komplex und es gibt selten nur eine Lösung. Sozialverhalten ist auch kulturabhängig und Bewertungen über die Angemessenheit von Reaktionen können sich schon zwischen Generationen unterscheiden.

10. Das Konzept des SJT ist für **andere** Bereiche bestens geeignet – wenn erwünschte Antworten nicht praktisch gefordert sind wie in einem Zulassungsverfahren, z.B. bei Self-Assessments zur Abklärung, „wie man wirklich ist und etwas einschätzt“. **Die Entscheidung, wer zum Medizinstudium zugelassen wird, erfordert den Einsatz wissenschaftlich abgesicherter und verantwortbarer Methoden. Es ist kein Platz für Experimente – vor allem nicht, wenn man Lösungen versprechen würde, denen aufgrund des vorliegenden Kenntnisstandes jegliche Basis fehlt.**