

**Zusammenfassung****Abstract****Schlüsselwörter**

Medizinstudium · Prüfungsmethoden · Prüfungsmethoden, schriftliche · Multiple Choice

Key words

Medical education · assessment · written assessment · multiple choice tests

Multiple Choice – Nein, danke!

Unter allen Prüfungsformen, die in der medizinischen Ausbildung im deutschsprachigen Raum angewandt werden, hat sicherlich die Multiple-Choice-Prüfung den bei weitem schlechtesten Ruf. Die Publikationen, die MC-Prüfungen verdammen, füllen ganze Bibliotheken und an Medizinischen Fakultäten gibt es keine einfachere Methode, allgemeine Zustimmung zu erheischen, als eine Polemik gegen die Prüfungen nach dem Multiple-Choice-Verfahren zu starten.

Trotzdem werden Multiple-Choice-Prüfungen weit verbreitet eingesetzt, und nicht zuletzt in all denjenigen Fächern, die sich aktuell unter der neuen Approbationsordnung veranlasst sehen, erstmals oder seit langer Zeit wieder Prüfungen durchzuführen, ist der erste Griff der zu den Multiple-Choice-Fragen aus alten Staatsexamina.

Bei der Auswahl einer Prüfungsmethode sind verschiedene Aspekte zu berücksichtigen, die hier nicht im Einzelnen betrachtet werden sollen. Als Grundsatzüberlegung sollte bei jeder Form der Prüfung beachtet werden:

Institutsangaben

Abteilung Allgemeinmedizin der Universität Ulm

Korrespondenzadresse

Markus Gulich · Abteilung Allgemeinmedizin der Universität Ulm · Helmholtzstraße 20 · 89069 Ulm ·
Tel.: 07 31/50-3 11 03 · Fax: 07 31/50-3 11 09 · E-mail: markus.gulich@medizin.uni-ulm.de

Bibliografie

Z Allg Med 2004; 80: 1–6 © Georg Thieme Verlag Stuttgart · New York · ISSN 0014-336251 ·
DOI 10.1055/s-2004-818750

Jede Prüfung ist ein diagnostischer Test ist zur Therapiekontrolle beim Patienten Student mit der Diagnose Unkenntnis unter der Therapie Studium.

Die Entscheidung für eine bestimmte Prüfungsmethode ist also, genau wie die Entscheidung für einen diagnostischen Test, von verschiedenen unterscheidbaren Faktoren abhängig.

Als erster wichtiger Faktor geht in die Überlegung ein, welche Lehrziele bei welchen Studenten überprüft werden sollen. Selbstverständlich würde – um den Vergleich zu strapazieren – niemand versuchen, eine Depression durch ein Röntgenbild diagnostizieren zu wollen. Trotzdem würde niemand auf die Idee kommen, Röntgen als eine schlechte Methode an sich zu betrachten.

Multiple-Choice wird häufig verdammt, weil es die übergeordneten und komplexen Ziele der Ausbildung zum Arzt nicht so abbildet, wie es sich jeder Lehrende wünschen würden. Multiple-Choice war aber nie angetreten, das zu können. Und selbstredend ist es grob fahrlässig, alle Prüfungen des Medizinstudiums auf eine Methode zu beschränken.

Der Diagnostiker wird – je nach klinischem Verdacht – neben dem Röntgen immer auch andere diagnostische Verfahren einsetzen und so zu einem sinnvollen diagnostischen Schluss gelangen. Die Qualität der Methode hängt in erster Linie davon ab, ob sie für adäquate Fragestellungen eingesetzt wird.

Das Erreichen bestimmter Lehrziele kann nur mittels einer dazu geeigneten Methode überprüft werden. Dazu ist es notwendig, die Lehrziele des Medizinstudiums in geeignete, einfach zugängliche Kategorien einzuteilen. Ein geeignetes Modell, verschiedene Kategorien von Lehrzielen zu unterscheiden, wurde in früheren Artikeln der ZFA [1] erläutert und wird hier schematisch in Abb. 1 dargestellt. Sie beruht darauf, Lehrziele einer Domäne der Bloom'schen Lernzieltaxonomie [2, 3] und verschiedenen Stufen der Miller'schen Kompetenzhierarchie [4] zuzuordnen.

Es liegt auf der Hand, dass eine einzelne Prüfungsmethode nicht geeignet ist, alle denkbaren Lehrziele zu überprüfen. So wie der Versuch, eine Depression oder eine Hepatitis durch Röntgenbilder zu diagnostizieren, zum Scheitern verurteilt ist, genau so ist der Versuch ‚manuelle Fertigkeiten‘ oder ‚höherstufige Kenntnis-

Kompetenzstufe nach Miller ↑	Performanz= Realanwendung	angewandte Kenntnisse	Alltagsanwendung	professionelles Alltagshandeln
	„Kompetenz“ = anwendungsbereit	anwendungsbereite Kenntnisse	anwendungsbereite Fertigkeiten	Selbständiges Handeln
	anwendungsorientiert	aktive Kenntnisse (werden produziert)	eigenständiges Wiederholen	Rollenverstehen
	passiv	passive Kenntnisse (werden erkannt)	Imitation	Regelwerk
	→ Domäne nach Bloom	Kenntnisse	Fertigkeiten	Haltungen/ Einstellungen

Abb. 1 Zweidimensionales Modell der Lehrzielbeschreibung (in Anlehnung an Bloom [2, 3] und Miller [4]).

se‘ oder ‚Haltungen‘ durch eine dafür ungeeignete Prüfungsmethode wie Multiple-Choice-Prüfungen zu überprüfen, ebenfalls zum Scheitern verurteilt.

Multiple-Choice-Fragen können Kenntnisse der ersten beiden, oder teilweise auch der dritten Kompetenzstufe und Lehrziele des ersten Niveaus der Haltungen/Einstellungen (Kenntnis und Verstehen von Regelwerken) zuverlässig überprüfen. Wer andere Lehrziele überprüfen will, muss andere, für diese Lehrziele jeweils adäquate Prüfungsmethoden einsetzen.

Die Ausrichtung an den zu überprüfenden Lehrzielen ist einer der Faktoren, der die Auswahl eines Prüfungsinstruments bestimmt. Praktikabilität und Testgütekriterien sind weitere Faktoren.

Es liegt auf der Hand, dass einer der herausragenden Vorteile von MC-Prüfungen ihre Praktikabilität ist. Mit kaum einer anderen Prüfungsform können mit begrenzten Ressourcen in knapp bemessener Zeit eine so große Zahl von Studierenden geprüft und benotet werden.

Das gilt vor allem auch, wenn man nicht nur die reine Praktikabilität der Prüfungsdurchführung betrachtet, sondern auch den Aufwand für die Entwicklung und Auswertung einer Prüfung in die Überlegungen mit einbezieht. In Tab. 1 wird der Versuch unternommen, den Aufwand für Entwicklung, Durchführung und Auswertung von 4 häufig angewandten Prüfungsformen zu vergleichen, wobei für alle hier betrachteten Prüfungsformen angenommen wird, dass sie auf jeweils hohem Qualitätsniveau des state of the art durchgeführt wird, das sich oft erheblich von der Alltagsanwendung unterscheidet.

Auch bei eingehender Betrachtung der Testgütekriterien schneiden Multiple-Choice-Prüfungen relativ gut ab. In Tab. 2 wird vergleichend die Reliabilitätskoeffizienten verschiedener Prüfungsmethoden dargestellt. Die Reliabilitätskoeffizienten sind ein Maß dafür, wie zuverlässig eine Prüfungsform unter gegebenen

Tab. 1 Vergleich des Aufwands für Entwicklung, Durchführung und Auswertung verschiedener Prüfungsformen

Aufwand für ...	MC	OSCE	strukturierte Beobachtung	Kurzantwortfragen
Entwicklung	++	++++	++	++(+)
Durchführung	+	+++	++++	+
Auswertung	+	++	++	++

MC = Multiple-Choice; OSCE = objective structured clinical examination

Tab. 2 Reliabilität verschiedener Prüfungsmethoden

Prüfungsmethode	MC	strukturiert mündlich	OSCE	PMP
Reliabilität	0,93	0,82	0,76	0,69

MC = Multiple-Choice; OSCE = objective structured clinical examination; PMP = patient management problems

Umständen ein reproduzierbares Prüfungsergebnis erzielen. Näherungsweise kann der Reliabilitätskoeffizient auch als ein Indikator für die Fairness einer Prüfung angesehen werden.

Für eine detailliertere Darstellung von Testgütekriterien und Bewertungskriterien verschiedener Prüfungsmethoden sei hier auf das Kapitel 2 des gratis aus dem Internet zu beziehenden Handbuchs ‚Kompetent Prüfen‘ der Schweizerischen und Österreichischen Ärztekammern, herausgegeben vom Berner Institut für Aus-, Weiter- und Fortbildung [5], verwiesen.

Verschiedene Formen von Multiple-Choice-Fragen

Einfachwahl 1-aus-5

In Deutschland werden fast ausschließlich die vom Mainzer Institut für Medizinische und Pharmazeutische Prüfungsfragen entworfenen Fragen mit Auswahl genau einer aus fünf Antwortmöglichkeiten als Multiple-Choice-Fragen wahrgenommen. Diese Typ A (bzw. in der negierten Form Typ A_{neg}) bezeichnete Variante zeichnet sich durch hohe Präzision und hohe Reliabilität aus, zieht aber nicht zuletzt wegen der hohen Ratewahrscheinlichkeit Kritik auf sich.

Dieser Fragetyp besteht im Idealfall aus einem Fragenstamm (idealerweise eine Fallvignette), der eigentlichen Frage und fünf Antwortmöglichkeiten, unter denen eine richtige (oder beste) Antwort vier so genannten Distraktoren, also nicht richtigen Antwortmöglichkeiten, gegenübersteht. Bei der Negativauswahl ist die ‚richtige Antwort‘ die nicht (oder am wenigsten) zutreffende, die Distraktoren sind zutreffende Aussagen. Gerade für den hausärztlichen Bereich, der sich ja durch besondere Vielschichtigkeit auszeichnet, gilt, dass Negativauswahlfragen leichter zu erstellen sind, aber oft auch sehr viel leichter zu beantworten sind.

Die Ratewahrscheinlichkeit von 20% lässt sich natürlich durch eine größere Anzahl von Distraktoren reduzieren, allerdings steigt mit der Anzahl der Distraktoren auch die Gefahr, unbeabsichtigt Hilfestellungen, so genannte cues – siehe unten – zu geben.

Durch die weite Verwendung von Fragen des Typs A bzw. A_{neg} durch das IMPP sind andere Varianten von Multiple-Choice-Fragen teilweise außer Gebrauch geraten. Da sie aber für bestimmte Fragestellungen besser geeignet sein können, sollen hier noch die Einfachauswahlfragen (Typ RF oder RF?), die (kombinierten) Zuordnungsfragen (Typ B), die kausalen Verknüpfungsfragen (Typ E) und die Mehrfachentscheidungsfragen (Typ K_{prim}) dargestellt werden.

Einfachauswahl richtig/falsch

Die einfachste Frageform ist die Einfachauswahl (Fragetyp RF). Der Prüfling soll zu einer einfachen Aussage entscheiden, ob diese Frage richtig oder falsch ist, bzw. welche von zwei sich ausschließenden Aussagen zutrifft. Da bei dieser einfachen Form die Ratewahrscheinlichkeit bei 50% oder, wegen möglicher cues in anderen Fragen der gleichen Prüfung, sogar höher liegt, wird im Auswertungsschema im Allgemeinen eine falsche Antwort mit Punktabzug gewertet, um eine Hürde gegen das reine Raten zu errichten. Um einen Prüfling, der eine Frage ehrlicher Weise nicht beantworten kann, nicht zum Raten zu zwingen, wird neben der reinen Einfachauswahl noch eine Möglichkeit gegeben,

eine Frage ergebnisneutral nicht zu beantworten, also neben den Antwortmöglichkeiten ‚stimmt‘ und ‚stimmt nicht‘ noch eine Möglichkeit ‚weiß ich nicht‘ zu bieten. Diese Variante ergibt aber vor allem dann einen Sinn, wenn falsche Antworten negativ sanktioniert werden.

Einfachauswahlfragen haben den großen Vorteil, dass sie sehr einfach zu erstellen, auszuwerten und zu bearbeiten sind, und wegen der großen Zahl von Fragen, die in einer kurzen Zeit bearbeitet werden können, ein großes thematisches Feld zumindest cursorisch überprüft werden kann. Wegen der hohen Ratewahrscheinlichkeit sind für eine aussagekräftige Bewertung oder Benotung aber eben auch eine große Zahl von Fragen notwendig.

Zuordnungsfragen und kombinierte Zuordnungsfragen

Bei so genannten Zuordnungsfragen (Typ B) wird einer Liste mit Begriffen eine Auswahlliste von Begriffen gegenübergestellt, die entsprechend der Fragestellung zuzuordnen sind. Im Beispiel 2 wird einer Liste von Pharmaka eine Auswahlliste von Kontraindikationen zugeordnet. Geeignete Zuordnungsfragen können eingesetzt werden, um Kenntnisse aus den höheren Kompetenzniveaus zu überprüfen, bergen aber ein hohes Risiko, unbeabsichtigte Hilfestellungen zu geben. Eine besondere Form der Zuordnungsfragen sind die kombinierten Zuordnungsfragen (Beispiel 3), die besondere Anforderungen an die Zusammenstellung der Antwortkombinationen stellen, da oft allein durch Wissen über eine der Aussagen die Ratewahrscheinlichkeit dramatisch steigt, und so die testtheoretische Qualität der Frage drastisch sinkt.

Kausale Verknüpfungsfragen

Einen Sonderfall stellen die kausalen Verknüpfungsfragen vom Typ E dar – Beispiel 4. Sie verbinden eine Aussage 1 über eine kausale Verknüpfung mit einer Aussage 2 und der Prüfling soll beurteilen, ob die Aussagen für sich richtig sind und ob die zweite Aussage die kausale Begründung der ersten Aussage darstellt. Besonders im multifaktoriellen und multidimensionalen hausärztlichen Aufgabenbereich sind Kausalitäten selten so einfach als richtig oder falsch zu beurteilen, deshalb ist von Fragen des Typs E im Großen und Ganzen eher abzuraten, vor allem dort, wo typisch hausärztliche Kenntnisse überprüft werden.

Mehrfachentscheidungsfragen

Als Frageform, die Situationen abbildet, in denen mehrere Sachverhalte, auch mehrere Sachverhalte aus sehr unterschiedlichen Kontexten, von Bedeutung sind, haben sich am ehesten Mehrfachentscheidungsfragen vom Typ K_{prim} herausgestellt. Dieser Fragetyp erlaubt es, verschiedene Aspekte eines Sachverhalts gleichzeitig nebeneinander zu stellen und zu beleuchten, wie in Beispiel 5 verdeutlicht wird. Der Prüfer behält aber im Gegensatz zu so genannten offenen Fragen die Wortwahl in seiner Entscheidung, was Formulierungsprobleme im Gegensatz zur ‚offenen Frage‘ minimiert.

Beim Fragentyp K_{prim} werden falsche Antworten mit Punktabzug bewertet, da sonst die Prüflinge generell möglichst viele oder alle Auswahlmöglichkeiten als Antwort angeben. Besonders wichtig ist es deshalb bei diesem Fragetyp, sich im Voraus mögliche Argumentationsketten bezüglich einzelner Aussagen zu vergegenwärtigen.

Schwierigkeit der Prüfung

Selbstverständlich ist ein wesentlicher Diskussionspunkt einer jeden Prüfung der Schwierigkeitsgrad. Es ist wichtig zu erkennen, dass der Schwierigkeitsgrad einer Multiple-Choice-Frage nicht in erster Linie von der Frage oder der richtigen Antwort abhängt. Der Schwierigkeitsgrad einer Multiple-Choice-Frage wird in erster Linie von der Auswahl der Distraktoren bestimmt. Verdeutlicht wird dieser Sachverhalt im Beispiel 6, bei dem zur gleichen Frage zwei Varianten von Antwortmöglichkeiten zur Auswahl gestellt werden. Während es einem großen Teil gelingen wird, in Variante 2 die richtige Lösung zu finden, gelingt die richtige Lösung in Variante 1 nur einem wirklichen Fachmann auf dem Gebiet.

Cues – unerwünschte Hilfestellungen

Eine weitere Anmerkung bezieht sich auf den wohl häufigsten Grund dafür, dass einzelne Multiple-Choice-Fragen nachträglich annulliert werden müssen oder dafür, dass einzelne Fragen sehr unerwartete Ergebnisse erzielen. Dieser Grund ist, neben reinen Schreib- oder Tippfehlern, das Vorhandensein von so genannten cues, unbeabsichtigten Hilfestellungen in der Wortwahl oder grammatikalischen Ausgestaltung der Fragen oder Antwortmöglichkeiten. Vor allem, da es unter deutschen Medizinstudenten eine weit gehende Kultur des „Knackens“ von Multiple-Choice-Fragen gibt, wird jeder auch noch so kleine Hinweis in den Formulierungen das Ergebnis der Prüfung verändern. Im ungünstigsten Fall werden Prüflinge sich mehr damit beschäftigen, ob eine bestimmte Formulierung als cue anzusehen ist, als sie sich mit der eigentlich zu prüfenden Frage auseinander setzen.

Auch ein cue in einem der Distraktoren verfälscht u. U. das Ergebnis der Prüfung, da er die Ratewahrscheinlichkeit ggf. dramatisch erhöhen kann.

Typische cues sind:

- Im Fragenstamm wird grammatikalisch ein Satz begonnen, der nicht in allen Antwortmöglichkeiten grammatikalisch korrekt weitergeführt wird.
Beispiel: Ein dringlicher Hausbesuch ... (A) ... wird ohne Zeitverzögerung, u.U. auch aus der Sprechstunde heraus unternommen. (B) ... stehen dem Hausarzt alle diagnostischen Möglichkeiten zur Verfügung. Antwortmöglichkeit (B) kann nicht stimmen, da der Satz nicht korrekt fortgeführt wird.
- Formulierungen mit Absolutheitscharakter (immer, nie, stets) sind als falsch, bzw. nicht zutreffend anzusehen, weil sich in der Medizin zu praktisch jeder Aussage ein Gegenbeispiel finden ließe.
Beispiel: Pflegende Angehörige ... (A) ... stehen oft in einer symbiontischen Beziehung zur gepflegten Person. (B) ... sind immer starken emotionalen Belastungen ausgesetzt. Antwortmöglichkeit (B) kann nicht stimmen, weil es hin und wieder auch Personen gibt, auf die das nicht zutrifft, wohingegen die Aussage ‚oft‘ in Antwortmöglichkeit (A) nicht bestreitbar ist. Deshalb sind Absolutausdrücke wenn möglich zu meiden.
- Ein zentraler Begriff des Fragenstamms oder der Frage wird nicht gleichmäßig in den Antwortmöglichkeiten wieder aufgegriffen.
Beispiel: Die häufigste Nebenwirkung von Kwoben ist (A) Swell; (B) Carancula; (C) Kwobenödem; (D) Schmalzen; (E) Bermotose. Das Wiederaufgreifen von ‚Kwoben‘ nur in Antwortmöglichkeit (C) macht es sehr wahrscheinlich, dass (C) richtig ist; es sollte

der Begriff entweder in allen oder in keiner Antwortmöglichkeit aufgegriffen werden.

- Eine oder wenige Antwortmöglichkeiten sind auffallend differenzierter oder spezifischer formuliert, als die anderen.
Beispiel: Die familienärztliche Funktion des Hausarztes wird am besten durch welche der folgenden Aussagen beschrieben? (A) Der Sohn darf das Rezept für die Mutter holen. (B) Sie umfasst die hausärztliche und gesundheitliche Betreuung von Familien oder familienähnlichen Gruppen in somatischer, psychischer und sozialer Hinsicht. (C) Ist ein Bonusmodell der AOK. Die sehr viel differenziertere Erklärung in Antwortmöglichkeit (B) sticht ins Auge.
- Zahlenangaben lassen in Größenordnung oder Ziffernfolge Muster erkennen.
Beispiel: Bei einer Frage nach einem Zahlenwert sind die Antwortmöglichkeiten 2,6 – 260 – 350 – 420 – 26000 vorgegeben. Die richtige Antwort wird die Ziffernfolge „26“ beinhalten (das wird am häufigsten angeboten) und wahrscheinlich in der Größenordnung „Hunderte“ liegen (wird ebenfalls dreimal angeboten. Die Schnittmenge „260“ ist dann die wahrscheinlich richtige Lösung. Besser (besser, aber auch schwieriger zu beantworten) wäre es, die Ziffernfolge zu variieren und insgesamt einen engeren Zahlenbereich zu wählen.
- Zahlenbereiche in den Antwortmöglichkeiten überlappen.
Beispiel: Frage nach einem Prozentsatz. Antwortmöglichkeiten: (A) < 15%; (B) 20–30%; (C) > 30%; (D) > 50%; (E) fast 100%. Wenn Antwort (D) oder (E) richtig wäre, wäre logischerweise auch Antwort (C) richtig; deshalb scheiden Antwort (D) und (E) aus, die Ratewahrscheinlichkeit liegt dann bei 1/3 statt 1/5.
- Besonders schwierig vorherzusehen ist es, wenn Studenten begriffliche Widersprüchlichkeiten entdecken, im Fragenstamm ist von einem akuten Ereignis die Rede, bei einigen Antwortmöglichkeiten werden Begriffe wie langsam oder allmählich benutzt.
- Ebenfalls schwierig vorherzusehen aber gelegentlich anzutreffen ist es, dass in einer Frage einer Gesamtprüfung ein eindeutiger Hinweis auf die richtige Lösung einer anderen Frage zu finden ist. Besonders leicht kann ein solcher cue auftreten, wenn mehrfach in einer Prüfung Medikamenten-Schemata oder Dosierungsfragen auftreten.
Beispiel: In einer Frage einer Prüfung, bei der es gar nicht zur Hypertoniebehandlung, sondern zur Langzeitbetreuung geht, ist in der Fallvignette davon die Rede, dass der Patient „... unter der Standardtherapie mit Enalapril und HCT normale Blutdruckwerte hat“. In einer anderen Frage wird nach einem geeigneten Diuretikum als Partner von Enalapril in der Hypertoniebehandlung gefragt.

Schlussfolgerung

Multiple-Choice-Prüfungen haben in Deutschland einen extrem schlechten Ruf. Inadäquater Einsatz, unrealistische Erwartungen an diese Prüfungsform und der Einsatz qualitativ schlechter Multiple-Choice-Prüfungen können diesen schlechten Ruf großteils erklären.

Multiple-Choice-Prüfungen werden aber aus prüfungstheoretischen Gründen und aus Gründen der Praktikabilität einen dauer-

haften Platz im Methodenarsenal eines jeden Prüfers an den medizinischen Fakultäten behalten.

Ein differenzierter und zielgerichteter Einsatz von Multiple-Choice-Prüfungen ist möglich. Die Entwicklung aussagekräftiger Multiple-Choice-Prüfungen bedarf eines gewissen Aufwands und klarer Überlegungen über die zu überprüfenden Lehr- und Lernziele. Jeder Prüfer sollte sich vergegenwärtigen, dass mit Multiple-Choice-Prüfungen nur das Erreichen ausgewählter Lehrziele, vor allem von kognitiven Lehrzielen überprüft werden kann, so wie mit einer diagnostischen Methode auch nur ein bestimmtes Spektrum von Verdachtsdiagnosen erhärtet werden kann. Zur Überprüfung anderer Lehrziele (Fertigkeiten, Haltungen, Professionalität) müssen andere Prüfungsformen zur Anwendung kommen [1].

Beispiel 1: Einfachwahlfrage Typ A, 1-aus-5

Eine 24-jährige Medizinstudentin (165 cm, 78 kg) gibt an, innerhalb weniger Tage eine Belastungsdyspnoe bekommen zu haben, und hat einmal Blut gehustet. Ihr Examen steht unmittelbar bevor, sie raucht bis zu 40 Zig/Tag und nimmt orale Kontrazeptiva ein. Klin. Untersuchung unauffällig bis auf Tachykardie und Wandruckschmerz rechts. Labor: BSG, Blutbild, Krea, Leberenzyme, Urinstatus unauffällig.

Die wahrscheinlichste Diagnose lautet:

- (A) Hyperventilationstetanie
- (B) Aktive Tuberkulose bei Ansteckung am Krankenbett
- (C) Rezidivierende Lungenembolien
- (D) Akute Sarkoidose
- (E) Goodpasture-Syndrom

richtige Antwort C

Beispiel 2: Zuordnungsfrage Typ B

Nennen Sie für die Antihypertonika der Liste 1 die am ehesten relevante Kontraindikation der Liste 2!

Liste 1

- (1) Verapamil
- (2) Enalapril
- (C) Hyperlipoproteinämie

Liste 2

- (A) Stabile Angina pectoris
- (B) AV-Block III. Grades
- (D) pAVK
- (E) Hereditäres Angioödem

richtige Lösungen: 1 B; 2 E

Beispiel 3: Kombinierte Zuordnungsfragen

Folgende Zustände sind bei zu Hause lebenden alten Menschen mit einem deutlich erhöhten Sturzrisiko verbunden:

- (1) Einnahme von mehr als 4 Medikamenten
- (2) Sehstörungen
- (3) regelmäßige Benutzung einer Gehhilfe
- (4) Sturz in den vergangenen 6 Monaten
- (A) Nur (4) ist richtig
- (B) Nur (2) und (4) sind richtig
- (C) Nur (1), (3) und (4) sind richtig
- (D) Nur (2), (3) und (4) sind richtig
- (E) Alle sind richtig

richtige Lösung E (alle Zuordnungen sind richtig)

Beispiel 4: kausale Verknüpfungsfrage Typ E

- (1) Fragen vom Typ E sind oft problematisch, weil
- (2) Kausalitäten nur selten eindeutig zu beurteilen sind
- (A) Beide Aussagen treffen zu, Verknüpfung ist berechtigt
- (B) Beide Aussagen treffen zu, Verknüpfung nicht berechtigt
- (C) Nur Aussage (1) trifft zu, Aussage (2) trifft nicht zu
- (D) Aussage (1) trifft nicht zu, nur Aussage (2) trifft zu,
- (E) Beide Aussagen sind falsch

richtige Antwort A: Aussage (1) trifft zu, Aussage (2) trifft zu; Aussage (2) ist (eine der) Begründungen für Aussage (1)

Beispiel 5: Mehrfachentscheidungsfrage Typ K_{prim}

Bei welchen der folgenden Personengruppen besteht eine offizielle Impfpflicht (STIKO) für eine Grippeimpfung?

- (A) Personen über 60 Jahre
- (B) Bewohner von Pflegeeinrichtungen
- (C) Personen in Erziehungs- und Pflegeberufen
- (D) Forstarbeiter und Jäger
- (E) Beschäftigte in der Wasserver- und -entsorgung
- (F) Kinder unter 16 Jahren
- (G) Beschäftigte in medizinischen Laboratorien

richtige Antworten: (A), (B) und (C)

Beispiel 6: Schwierigkeitsgrad der Frage

Ein Marathonlauf geht über eine Strecke von

Variante 1

- (A) 42,195 km
- (B) 42,951 km
- (C) 41,925 km
- (D) 41,592 km
- (E) 42,519 km

Variante 2

- (A) 42,195 km
- (B) 18,453 km
- (C) 21,459 km
- (D) 84,394 km
- (E) 62,567 km

Literatur

- ¹ Gulich M. Prüfungen zur Beurteilung komplexer Lehrziele. ZfA 2003; 79: 507–511
- ² Bloom BS(Hrsg). Taxonomy of educational objectives, I Cognitive domain. McKay, New York 1956
- ³ Krathwohl DR(Hrsg). Taxonomy of educational objectives, II Affective domain. McKay, New York 1964
- ⁴ Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990; 65: 63–67
- ⁵ Bloch R. Kompetent Prüfen – Handbuch zur Planung, Durchführung und Auswertung von Facharztprüfungen Bern, Wien 1999. Download unter: http://www.iawf.unibe.ch/aae/documents/methodik/kompetent_pruefen.pdf

Empfohlene weiterführende Lektüre zum Entwerfen einzelner Multiple-Choice-Fragen und zum Erstellen einer Prüfung aus MC-Fragen:
Krebs R. Anleitung zur Herstellung von MC-Fragen und MC-Prüfungen für die ärztliche Ausbildung. Bern, IAWF 2003. Download unter: http://www.iawf.unibe.ch/aae/documents/methodik/mc_anl_med.pdf

Zur Person



Dr. med. Markus Sebastian Gulich
Geboren 1959 in Nürnberg. Medizinstudium in Würzburg und Ulm, Approbation 1987, Promotion 1989, Facharzt für Allgemeinmedizin. Wissenschaftliche und ärztliche Tätigkeit am Institut für Arbeits-, Sozial und Umweltmedizin der Universität Ulm, an der medizinischen Universitätsklinik Ulm und am Kreiskrankenhaus Geislingen/Steige. Postgraduierten-Studiengang medical education am University of Wales College of Medicine in Cardiff, GB, Abschluss master of science in medical education 1994. Seit 1995 wissenschaftlicher Assistent der Abteilung Allgemeinmedizin der Uni Ulm und Dauerassistent in allgemeinmedizinischer Praxis in Kuchen/Fils.