

Studentische Evaluations-Skala für Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion (SES-Sim)

Zusammenfassung

Zielsetzung: Simulationen von Arzt-Patient-Interaktionen sind zu einer beliebten Methode für das Training medizinischer Kompetenzen, insbesondere kommunikativer Fähigkeiten, geworden. Ein neuer Fragebogen zur Erfassung der studentischen Zufriedenheit mit Lehrveranstaltungen, in denen diese Methode eingesetzt wird, wird vorgestellt, die Studentische Evaluations-Skala für Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion (SES-Sim).

Methodik: Eine Zusammenstellung von Items, die inhaltlich auf die Qualität des Kurses und die zentralen Elemente von Simulationen abzielen, wurde gebildet und 220 Medizinstudierenden vorgelegt, die mit dieser Methode trainiert worden waren.

Ergebnisse: Auf der Basis von Faktorenanalysen wurden 18 Items für die Endversion der Skala ausgewählt, die fünf Dimensionen repräsentieren: Lernerfolg, Schauspielpatienten, Räumlichkeiten, Dozenten und Studierende. Die fünf Dimensionen korrelieren alle bedeutsam mit einer 1-Item-Messung der generellen Zufriedenheit mit dem Kurs.

Schlussfolgerung: Die SES-Sim ermöglicht es Lehrenden, auf ökonomische Weise einschätzen zu können, ob der Kurs den Bedürfnissen der Studierenden entsprach und was verbessert werden kann.

Schlüsselwörter: Medizinische Kompetenzen, Kommunikation, Arzt-Patient-Interaktion, Simulation, studentische Evaluation

Eva Neumann¹
Rainer Obliers¹
Christine Schiessl²
Christoph Stosch³
Christian Albus¹

1 Universität zu Köln, Klinik und Poliklinik für Psychosomatik und Psychotherapie, Köln, Deutschland

2 Universität zu Köln, Zentrum für Palliativmedizin, Köln, Deutschland

3 Universität zu Köln, Medizinische Fakultät, Studiendekanat, Köln, Deutschland

Einleitung

In der medizinischen Ausbildung werden zunehmend innovative Lehrveranstaltungen durchgeführt, in denen als Lehrmethode Simulationen der Arzt-Patient-Interaktion eingesetzt werden [1], [2], [3], [4]. Hierbei stellen Studierende und Schauspieler Interaktionen zwischen Arzt und Patient in Rollenspielen nach. Die Schauspieler folgen dabei einem Skript, das inhaltliche Vorgaben zur Darstellung eines bestimmten Krankheitsbildes enthält und zuvor von ihnen eintrainiert wurde. Auch die Studierenden erhalten meist eine Vorgabe zu der Aufgabe, die sie in der Simulation erfüllen sollen. Simulationen der Arzt-Patient-Interaktion dienen der Einübung praktischer ärztlicher Fertigkeiten. Insbesondere erscheinen sie geeignet zur Verbesserung kommunikativer und sozialer Kompetenzen, doch auch andere Fertigkeiten wie zum Beispiel einige körperliche Untersuchungen können mit dieser Methode praxisnah geübt werden.

Die Vorteile der Methode liegen auf der Hand. Da Schauspieler besser verfügbar sind als reale Patienten, ist die Organisation der Veranstaltung einfacher. Die Suche nach Patienten, die das zu behandelnde Krankheitsbild aufweisen und damit einverstanden sind, an einer Lehrveranstaltung teilzunehmen, entfällt. Die Simulation von Krankheitsbildern durch Schauspieler, die in vorher-

sehbarer Weise eingesetzt werden können, bietet die Möglichkeit, das betreffende Krankheitsbild nachzustellen und auch die Schwere der Symptome zu variieren. Da reale Patienten nicht beteiligt sind, wird vermieden, dass diese emotional belastet werden könnten. Die Studierenden können angstfrei lernen, denn wenn sie Fehler machen, schaden sie damit keinem Patienten. Anders als im echten Patientenkontakt, bei dem es meist keine Rückmeldung gibt, erhalten sie ein Feedback zu ihrer Leistung. Die größte Stärke der Simulationen ist sicherlich die Nähe zur Realität, die vor allem dadurch entsteht, dass die Studierenden in die Interaktion mit einem Menschen eintreten (und beispielsweise nicht an einer Puppe üben). Die Realitätsnähe dürfte den Transfer des Erlernen auf die spätere Berufspraxis erleichtern.

Erfahrungsberichte aus bereits durchgeführten Lehrveranstaltungen zeigen dementsprechend, dass Studierende diese Form der Lehre sehr positiv beurteilen [5], [6], [7]. Die Teilnehmer geben auch an, dass sich durch die Kurse ihre kommunikative Kompetenz erhöht hat [8], [9].

Standardisierte Messinstrumente zur Evaluierung von Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion liegen bislang nicht vor. Allgemeine Fragebögen zur studentischen Evaluation der Lehre erscheinen aus zwei Gründen als nicht geeignet:

- Zum einen sind viele der an Universitäten und Universitätskliniken eingesetzten Evaluationsbögen selbst entwickelt und nicht veröffentlicht, zumindest nicht in einer Zeitschrift mit Peer-Review-Verfahren. Daher kann die Testgüte dieser Verfahren nicht beurteilt werden.
- Zum anderen sind wichtige Aspekte der Simulationen darin nicht enthalten. Da diese Methode in konventionellen Lehrveranstaltungen nicht zum Einsatz kommt, wird in allgemeinen Fragebögen zur studentischen Evaluation nicht nach Variablen wie der Leistung der Schauspieler, der Güte des Feedbacks und der Verbesserung kommunikativer Fähigkeiten gefragt.

Um die studentische Zufriedenheit mit Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion erfassen zu können, wurde in dieser Arbeit ein neuer Fragebogen entwickelt, die **Studentische Evaluations-Skala für Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion (SES-Sim)**.

Die Entwicklung der SES-Sim erfolgte innerhalb des Lehrprojekts „PJ-Start-Block“ (Schlüsselkompetenz-Training und -Anwendung in realitätsnahen Tagesabläufen) an der medizinischen Fakultät der Universität zu Köln. In diesem Projekt wurde eine Lehrveranstaltung mit Simulationen der Arzt-Patient-Interaktion als zentraler Lehrmethode konzipiert und durchgeführt. Es handelt sich um eine einwöchige Blockveranstaltung am Ende des Studiums zur Vorbereitung auf das Praktische Jahr (PJ). Der Alltag auf einer Krankenhausstation wird nachgestellt, wobei die Studierenden in die Rollen von Ärzten schlüpfen und mit Schauspielpatienten in verschiedenen Szenarien interagieren. Nachgestellt werden die Krankheitsbilder Herzinsuffizienz, Appendizitis, Eileiterschwangerschaft, chronische Rückenschmerzen, Diabetes, Harnwegsinfekt mit Fieber und Bronchialkarzinom.

Die Rollenskripte sind meist so konstruiert, dass nicht nur ein somatisches Krankheitsbild präsentiert wird, sondern auch psychosoziale Aspekte auftauchen, die für die Behandlung der Krankheit relevant sind. So wird bei der Simulation der Eileiterschwangerschaft vorgegeben, dass die Patientin aus dem muslimischen Kulturkreis stammt und ein Krankheitsverständnis hat, das mit naturwissenschaftlichen Vorstellungen nicht in Einklang zu bringen ist. Die Rolle der Diabetes-Patientin umfasst Probleme bei der Insulintherapie, die darauf zurückzuführen sind, dass komorbid eine depressive Reaktion auftritt. Die Studierenden stehen daher vor der Aufgabe, nicht nur auf die somatischen Aspekte, sondern auch auf krankheitsrelevante psychosoziale Begleitumstände angemessen einzugehen.

Ergänzend zu den Simulationen werden Übungen zu ärztlichen Aufgaben ohne Patientenkontakt angeboten (z.B. Morgen- und Mittagsbesprechungen im Ärzteteam, das Schreiben von Arztbriefen und die Anforderung von Konsilen).

Methode

Itemzusammenstellung

Zur Entwicklung des neuen Fragebogens wurden 33 Items zusammengestellt, mit denen die zentralen Elemente der Lehrveranstaltung inhaltlich abgebildet werden. Einige Items entstanden in Anlehnung an Items aus den Fragebögen zur studentischen Evaluation von Hochschulveranstaltungen von Diehl [10]. Die meisten Items sind neu formuliert.

Als Antwortformat wurde eine 5-stufige Likert-Skala mit den Endpunkten 1 („stimme überhaupt nicht zu“) und 5 („stimme voll und ganz zu“) festgelegt.

Stichprobe

Die Zusammenstellung der Items wurde den ersten Teilnehmern des PJ-Start-Blocks im Wintersemester 2009/10 und im Sommersemester 2010 vorgelegt, und zwar am letzten Tag der Blockveranstaltung. Insgesamt füllten 220 Studierende der Medizin den Fragebogen aus (66% Frauen, 34% Männer), deren durchschnittliches Alter bei 26 Jahren lag.

Statistische Auswertung

Die Itemauswahl erfolgte auf der Grundlage von drei Faktorenanalysen. Die Daten der Stichproben aus dem Winter- und dem Sommersemester wurden zunächst getrennt voneinander jeweils einer Faktorenanalyse unterzogen. Auf der Grundlage der Ergebnisse beider Faktorenanalysen wurden Items für die Endversion der Skala ausgewählt. Die Itemauswahl wurde statistisch abgesichert, indem die Endversion einer erneuten Faktorenanalyse unterzogen wurde, wobei die Teilnehmer des Winter- und des Sommersemesters hierfür zu einer Gesamtstichprobe zusammengefasst wurden.

Ergebnisse

Bei der 33-Item-Version, die den Studierenden im ersten Durchgang im Wintersemester vorgelegt wurde, zeigen sich in der Faktorenanalyse neun Faktoren mit einem Eigenwert > 1 (Eigenwertverlauf 8.61, 3.14, 2.65, 2.12, 1.57, 1.49, 1.22, 1.14, 1.02, .98 ...). Von den neun Faktoren lassen sich fünf inhaltlich sinnvoll interpretieren; diese beziehen sich inhaltlich auf den Lernerfolg, die Schauspieler, die Räumlichkeiten, die Dozenten und die Studierenden. Die übrigen vier Faktoren bestehen entweder nur aus einem einzelnen Item oder sind inhaltlich heterogen. Auf den fünf interpretierbaren Faktoren laden 23 der 33 Items in eindeutiger Weise, d.h. sie laden auf einem Faktor hoch und auf den jeweils anderen niedrig. Für den zweiten Durchgang im Sommersemester wurde eine gekürzte Version des Fragebogens erstellt, der aus den 23 eindeutig zuzuordnenden Items bestand. In der Faktorenanalyse der 23-Item-Version zeigen sich sieben

Faktoren mit einem Eigenwert > 1 (Eigenwertverlauf: 5.37, 2.33, 1.99, 1.53, 1.49, 1.31, 1.16, .95 ...). Die ersten fünf entsprechen den in der ersten Faktorenanalyse gefundenen interpretierbaren Faktoren. Die beiden anderen Faktoren sind inhaltlich heterogen. Von den 23 Items laden 17 eindeutig auf dem Faktor, dem sie sich auch in der ersten Faktorenanalyse zugeordnet hatten. Die anderen sechs Items weisen entweder Doppelladungen auf oder laden auf einem anderen Faktor als im ersten Durchgang.

Die Endauswahl der Items erfolgte nach folgenden Kriterien: 16 der 17 Items, die sich in beiden Faktorenanalysen eindeutig einem der fünf interpretierbaren Faktoren zuordnen, wurden in die Endversion des Fragebogens aufgenommen. Eines der 17 Items wurde trotz eindeutiger Faktorladungen ausgeschlossen, weil es eine starke semantische Ähnlichkeit mit einem anderen Item aufweist. (Beide Items beziehen sich auf den Transfer des Erlernten auf die spätere Berufspraxis; aus ökonomischen Gründen wird darauf verzichtet, dies doppelt einschätzen zu lassen.) Zwei Items, die nach dem Kriterium der eindeutigen Zuordnung in beiden Faktorenanalysen eigentlich hätten ausgeschlossen werden müssen, blieben aus inhaltlichen Gründen in der Endauswahl. Die beiden Items dienen der Bewertung des Feedbacks, einem zentralen Element der Lehrveranstaltung, das auf jeden Fall inhaltlich in dem neuen Fragebogen enthalten sein soll. Insgesamt wurden somit 18 Items für die Endversion des Fragebogens ausgewählt.

Die Faktorenanalyse zur statistischen Absicherung der 18-Item-Version in der Gesamtstichprobe bestätigt die erwartete 5-faktorielle Struktur (Eigenwertverlauf: 4.39, 2.30, 2.07, 1.67, 1.28, .90 ...). Die fünf Faktoren klären zusammen 65% der Varianz auf. Tabelle 1 zeigt die Itemkennwerte.

Die 18 Items ordnen sich den fünf Faktoren eindeutig zu. Alle Items laden auf einem Faktor hoch und auf den jeweils anderen Faktoren niedrig. Die Items erweisen sich weiterhin als trennscharf; die Item-Skala-Korrelationen liegen meist deutlich über dem Grenzwert von .30.

Tabelle 2 zeigt die Kennwerte der fünf Unterskalen der SES-Sim. Die mittlere Zustimmung ist bei allen Skalen hoch, was darauf verweist, dass die Lehrveranstaltung von den Studierenden positiv beurteilt wird. Die internen Konsistenzen fallen zufrieden stellend bis gut aus.

Tabelle 2: Skalenkennwerte

	Zahl der Items	N	M	Md	SD	α
Lernerfolg	5	220	4.16	4.30	.67	.84
Schauspieler/innen	4	220	4.62	4.75	.48	.77
Räumlichkeiten	3	220	4.65	5.00	.60	.83
Dozent/inn/en	3	220	4.91	5.00	.23	.63
Studierende	3	220	4.59	4.67	.45	.66

N = Zahl der Probanden, M = Mittelwert, Md = Median, SD = Standardabweichung, α = Interne Konsistenz (Cronbach's Alpha)

In der Tabelle 3 schließlich ist dargestellt, wie die fünf Unterskalen der SES-Sim mit einer Gesamtbewertung der

Blockveranstaltung zusammenhängen. Die Gesamtbewertung erfolgte in Analogie zu Schulnoten auf einer Skala von 1 („sehr gut“) bis 5 („mangelhaft“). Da die Unterskalen der SES-Sim keine Normalverteilung aufweisen (die Verteilungen sind rechtssteil), wurden non-parametrische Korrelationen berechnet (Spearman-Rho-Korrelationen).

Tabelle 3: Korrelationen der Skalen mit der Gesamtbewertung

	Gesamtbewertung
Lernerfolg	-.49***
Schauspieler/innen	-.29***
Räumlichkeiten	-.24**
Dozent/in	-.38***
Studierende	-.31***

*** $p < .001$, ** $p < .01$

Es zeigt sich, dass das Gesamturteil über die Veranstaltung vor allem mit der Einschätzung des Lernerfolgs zusammenhängt. Das heißt, die Studierenden beurteilen die Veranstaltung um so positiver, je mehr sie den Eindruck haben, dort etwas gelernt zu haben. Auch die Einschätzungen der Leistung aller beteiligten Personengruppen (Dozent/inn/en, Schauspielpatient/inn/en und die Studierenden selbst) und der Räumlichkeiten hängen signifikant mit dem Gesamturteil zusammen. Somit erweisen sich alle Dimensionen der SES-Sim als bedeutsam für die Gesamtbeurteilung der Veranstaltung.

Diskussion

Mit der SES-Sim wird ein Messinstrument vorgelegt, das eine schnelle, ökonomische Erfassung der Zufriedenheit von Studierenden mit Lehrveranstaltungen, in denen die Methode der Simulation von Arzt-Patient-Interaktionen eingesetzt wird, ermöglicht. Kirkpatrick und Kirkpatrick [11], die bei der Evaluation zwischen den vier hierarchisch geordneten Ebenen Reaktion, Lernen, Verhalten und Ergebnis unterscheiden, ordnen diese Form der Evaluation der ersten Ebene zu. Zufriedenheitsmessungen geben demnach Aufschluss über die Reaktion der Teilnehmer auf eine Trainingsmaßnahme, sagen jedoch noch nichts darüber aus, inwieweit ein Wissenszuwachs erreicht wurde, Änderungen des Verhaltens auftreten und die Produktivität bzw. die Qualität der Arbeit ansteigen. Das hohe Maß an Zufriedenheit, das sich in den hohen Mittelwerten der Skalen zeigt, entspricht den Erfahrungen aus Kursen an anderen Universitäten, in denen die Studierenden sich ebenfalls sehr positiv über diese Form von Lehrveranstaltung äußerten [5], [6], [7]. Neben der Praxisnähe der Simulationen spielt hier sicherlich auch der Neuheitswert, den diese Methodik im Medizinstudium noch hat, eine Rolle. Auch die Intensität der Betreuung aufgrund der Arbeit in Kleingruppen, die in anderen Lehrveranstaltungen so nicht gegeben ist, dürfte sich günstig auf die Bewertung auswirken.

Die Items der SES-Sim beziehen sich unmittelbar auf die zentralen Elemente von Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion. Außerdem wurde

Tabelle 1: Itemkennwerte

		a ¹	a ²	a ³	a ⁴	a ⁵	r _{it}
1	Im Kurs habe ich viel Neues gelernt.	.64	.00	.11	.20	.08	.54
2	Durch den Kurs bin ich in Bezug auf die ärztliche Tätigkeit sicherer geworden.	.85	.05	.07	.10	.02	.75
3	Durch den Kurs hat sich meine ärztliche Kompetenz erhöht.	.81	.01	.15	.05	.04	.69
4	Durch den Kurs hat sich meine Kompetenz in Bezug auf die ärztliche Gesprächsführung erhöht.	.77	.21	.02	.13	.01	.65
5	Das in diesem Kurs Gelernte kann ich in der späteren Berufspraxis gut gebrauchen.	.68	.18	-.02	.18	.09	.60
6	Die Schauspieler/innen schienen auf ihre Rolle als Patient/in gut vorbereitet zu sein.	.15	.83	-.10	.00	.10	.66
7	Die Schauspieler/innen gaben ein hilfreiches Feedback.	.05	.87	.05	-.03	-.01	.70
8	Die Schauspieler/innen waren in ihrer Rolle als Patient/in glaubwürdig.	.08	.61	.11	.10	.06	.45
9	Die Kontakte mit Schauspielpatient/inn/en kamen der Realität nahe.	.07	.78	.19	.00	.00	.59
10	Die Räumlichkeiten fand ich angemessen.	.18	.02	.87	.01	.03	.72
11	Die Ausstattung der Räumlichkeiten fand ich angemessen.	.15	.10	.88	-.03	.11	.76
12	Die Räumlichkeiten und die Ausstattung sahen echt aus.	-.05	.14	.80	.13	.10	.61
13	Der Dozent/die Dozentin schien gut vorbereitet zu sein.	.12	-.03	.18	-.17	.74	.42
14	Der Dozent/die Dozentin moderierte den Kurs kompetent.	.06	.14	.14	.12	.84	.58
15	Der Dozent/die Dozentin war am Lernerfolg der Studierenden interessiert.	.02	.04	-.08	.29	.70	.36
16	Die Studierenden arbeiteten aktiv mit.	.20	-.06	-.04	.76	.08	.48
17	Die Studierenden gaben ein hilfreiches Feedback.	.17	-.04	.11	.83	-.02	.59
18	Die Studierenden wirkten aufmerksam und interessiert.	.21	.23	.06	.59	.15	.36

a¹⁻⁵ Ladung auf Faktor 1-5 nach Varimax-Rotation, r_{it} = Trennschärfe

darauf geachtet, dass sie inhaltlich eindeutig formuliert sind. Daher kann dem neuen Fragebogen sicherlich Augenscheinvalidität zugesprochen werden.

Eine Konstruktvalidierung über ein Außenkriterium wäre allerdings schwierig durchzuführen. Ein Grund dafür liegt darin, dass mit der Methode der Simulation komplexe Fertigkeiten trainiert werden, die mehrere Dimensionen umfassen. Wenn also beispielsweise der Lernfortschritt in der kommunikativen Kompetenz erhoben werden sollte, was nach Kirkpatrick und Kirkpatrick [11] eine Evaluierung auf der zweiten der vier Ebenen darstellen würde, so müssten im ersten Schritt die Dimensionen, aus denen diese Fähigkeit besteht, identifiziert werden, und im zweiten müssten für jede Dimension passende Operationalisierungen gefunden werden. Dieses Vorgehen wäre aufwendig und mit einer großen Ungenauigkeit behaftet. So wäre es bereits schwierig festzustellen, ob alle für die kommunikative Kompetenz relevanten Dimensionen identifiziert worden sind. Auf eine Konstruktvalidierung der SES-Sim wurde in dieser Arbeit daher verzichtet. Eine weitere Begrenzung, die nicht nur auf den hier vorgestellten Fragebogen, sondern auf alle Messinstrumente zur studentischen Evaluation von Lehrveranstaltungen zutrifft, besteht darin, dass die Zufriedenheit von Studierenden nur bedingt die tatsächliche Güte der Lehrveranstaltung widerspiegelt [12]. So zeigt sich beispielsweise,

dass Studierende der Medizin Veranstaltungen zu Grundlagenfächern schlechter beurteilen als solche zu klinischen Fächern [13]. Bei diesen Urteilen lassen sie sich offensichtlich weniger von der didaktischen Qualität der Veranstaltungen leiten, sondern urteilen eher danach, was sie als nützlich für die spätere Berufstätigkeit ansehen. Aus der Psychologie ist bekannt, dass Vorlesungen und Seminare zur Statistik regelmäßig schlechter abschneiden als die Veranstaltungen zu anderen Fächern, weil dieses Fach den Vorstellungen vor allem von Studienanfängern über Psychologie zuwider läuft [10]. Studentische Evaluationen sind in hohem Maße von der generellen Beliebtheit des Fachs abhängig; weniger populäre Fächer werden regelmäßig schlechter beurteilt als populäre.

Daher wäre es problematisch, sich allein auf studentische Evaluationen zu stützen, wenn die Güte von Lehrveranstaltungen eingeschätzt werden soll; in diesem Fall sollten weitere Kriterien wie zum Beispiel Erhebungen des Lernfortschritts herangezogen werden. Doch wenn sich Lehrende eine Orientierung darüber verschaffen wollen, wie ihre Veranstaltung bei den Studierenden ankam und was möglicherweise verbessert werden kann, dann können Fragebögen wie die SES-Sim hilfreich sein.

Danksagung

Für hilfreiche Anmerkungen zu dieser Arbeit Dank an Dr. Valentin Goede, Houda Hallal, Wencke Johannsen, Ortrun Kliche, Dr. Sabine Teschendorf und Christian Thrien, Universitätsklinikum Köln und Universität zu Köln.

PJ-Start-Block ist ein Lehrprojekt der Universität zu Köln mit Beteiligung folgender Institutionen: Medizinische Fakultät: Studiendekanat und Kölner Interprofessionelles Skills Lab und Simulationszentrum (Prof. Dr. Dr. Lehmann, Dr. Boldt, Dr. h.c. (RUS) Stosch), Institut für Geschichte und Ethik der Medizin (Prof. Dr. Karenberg, Prof. Dr. Dr. Schäfer), Institut für Pharmakologie (Prof. Dr. Herzig, PD Dr. Matthes), Klinik und Poliklinik für Psychosomatik und Psychotherapie (PD Dr. Albus, Prof. Dr. Obliers, Dr. Koerfer), Zentrum für Palliativmedizin (Prof. Dr. Voltz, PD Dr. Schiessl) und Humanwissenschaftliche Fakultät: Institut für vergleichende Bildungsforschung und Sozialwissenschaften (Prof. Dr. Allemann-Ghionda)

Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenskonflikte im Zusammenhang mit diesem Artikel haben.

Literatur

- Swanson DB, Stillman PL. Use of standardized patients for teaching and assessing clinical skills. *Eval Health Prof.* 1990;13:79-103.
- Ainsworth MA, Rogers LP, Markus JF, Dorsey NK, Blackwell TA, Petrusa ER. Standardized patient encounters. A method for teaching and evaluation. *J Amer Med Assoc.* 1991;266:1390-1396.
- Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med.* 1993;68(6):443-453.
- Ortwein H, Fröhmel A, Burger W. Einsatz von Simulationspatienten als Lehr-, Lern- und Prüfungsform. *Psychother Psych Med.* 2006;56:23-29.
- Bachmann C, Barzel A, Dunkelberg S, Schrom K, Ehrhardt M, van den Bussche H. Fachübergreifendes Kommunikationstraining mit Simulationspatienten: ein Pilotprojekt ins Curriculum. *GMS Z Med Ausbild.* 2008;25(1):Doc58. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2008-25/zma000542.shtml>
- Koerfer A, Köhle K, Obliers R, Sonntag B, Thomas W, Albus C. Training und Prüfung kommunikativer Kompetenz. Aus- und Fortbildungskonzepte zur ärztlichen Gesprächsführung. *Gesprächsforsch Z verbal Interaktion.* 2008;9:34-78.
- Nikendei C, Zipfel S, Roth C, Löwe B, Herzog W, Jünger J. Kommunikations- und Interaktionstraining im psychosomatischen Praktikum: Einsatz von standardisierten Patienten. *Psychother Psych Med.* 2003;53:440-445.

- Schildmann J, Härlein J, Burchardi N, Schlögl M, Vollmann J. Breaking bad news: evaluation study on self-perceived competences and views of medical and nursing students taking part in a collaborative study. *Support Care Cancer.* 2006;14(11):1157-1161.
- Simmenroth-Nayda A, Weiß C, Chenot J-F, Scherer M, Fischer T, Kochen MM, Himmel W. Verbessern Anamneseübungen die kommunikativen Fähigkeiten von Studierenden? Ein Prä-Post-Vergleich. *GMS Z Med Ausbild.* 2007;24(1):Doc22. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2007-24/zma000316.shtml>
- Diehl JM. Normierung zweier Fragebögen zur studentischen Beurteilung von Vorlesungen und Seminaren. *Psychol Erz Unterr.* 2003;50:27-42.
- Kirkpatrick DL, Kirkpatrick JD. Evaluating training programs: The four levels. 3rd ed. San Francisco: Berrett-Koehler Publishers; 2006.
- Kromrey H. Wie erkennt man „gute Lehre“? Was studentische Vorlesungsbefragungen (nicht) aussagen. *Empirische Pädagogik.* 1994;8:153-168.
- van den Bussche H, Weidtmann K, Kohler N, Frost M, Kaduszkiewicz H. Evaluation der ärztlichen Ausbildung: Methodische Probleme der Durchführung und der Interpretation von Ergebnissen. *GMS Z Med Ausbild.* 2006;23(2):Doc37. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2006-23/zma000256.shtml>

Korrespondenzadresse:

Dr. Eva Neumann
Universität zu Köln, Klinik und Poliklinik für Psychosomatik und Psychotherapie, Kerpener Straße 61, 50931 Köln, Deutschland, Tel.: +49 (0)221/478-4103, Fax: +49 (0)221/478-3103
eva.neumann@uni-duesseldorf.de

Bitte zitieren als

Neumann E, Obliers R, Schiessl C, Stosch C, Albus C. Studentische Evaluations-Skala für Lehrveranstaltungen mit Simulationen der Arzt-Patient-Interaktion (SES-Sim). *GMS Z Med Ausbild.* 2011;28(4):Doc56.
DOI: 10.3205/zma000768, URN: <urn:nbn:de:0183-zma0007688>

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2011-28/zma000768.shtml>

Eingereicht: 01.12.2010

Überarbeitet: 14.07.2011

Angenommen: 27.07.2011

Veröffentlicht: 15.11.2011

Copyright

©2011 Neumann et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.

Student Evaluation Scale for Medical Courses with Simulations of the Doctor-Patient Interaction (SES-Sim)

Abstract

Objective: Simulations of doctor-patient interactions have become a popular method for the training of medical skills, primarily communication skills. A new questionnaire for the measurement of students' satisfaction with medical courses using this technique is presented, the Student Evaluation Scale for Medical Courses with Simulations of the Doctor-Patient Interaction (SES-Sim).

Method: A set of items focusing on the course quality and the core elements of simulations was created and presented to 220 medical students who had been trained with this method.

Results: Based on factor-analyses 18 items were selected for the final version of the scale, which represent five dimensions: learning success, actors, premises, tutors and students. The five dimensions are all significantly correlated with a 1-item-measure of the general satisfaction with the course.

Conclusion: The SES-Sim enables tutors to assess in an economic way whether the course has met the students' needs and what can be done better.

Keywords: Medical skills, communication, doctor-patient interaction, simulation, student evaluation

Eva Neumann¹
Rainer Obliers¹
Christine Schiessl²
Christoph Stosch³
Christian Albus¹

1 University of Cologne,
Department of
Psychosomatics and
Psychotherapy, Cologne,
Germany

2 University of Cologne, Centre
for Palliative Medicine,
Cologne, Germany

3 University of Cologne, Faculty
of Medicine, Dekanat of
Study, Cologne, Germany

Introduction

In medical education, innovative lectures are increasingly used, employing teaching methods simulating the doctor-patient interaction [1], [2], [3], [4]. In these simulations, students and actors enact interactions between doctors and patients as role play. The actors follow a script which prescribes the content requirements for representing a particular disease which they train beforehand. Students also usually receive guidance on the role they are to fulfil in the simulation. Simulations of doctor-patient interaction serve to practice practical medical skills. They appear particularly suitable for improving communication and social skills, but certain other skills such as physical examinations can also be practiced using this hands-on approach.

The advantages of this method are obvious. Because actors are more readily available than real patients, organising such a session is easy. The need to search for a patient who has the required disease and who will agree to participate in a teaching session is thereby removed. Using actors to simulate clinical conditions who can be employed in a predictable way offers the possibility for simulating the relevant disease pattern and for varying the severity of symptoms. As real patients are not involved any emotional stress is also avoided. Students are able

to learn without fears as they will not harm a patient if they make a mistake. Unlike with real patient contact where students usually receive no feedback, simulations allow for feedback on their performance. The greatest strength of simulations is certainly their closeness to reality, caused primarily by the fact that students interact with a human (and not practicing with a doll, for example). This closeness to reality is likely to facilitate the transfer of skills to future professional practice.

Reports from previous courses accordingly show that students assess this form of teaching very positively [5], [6], [7]. The participants also state that the courses have increased their communication skills [8], [9].

Standardised measurement instruments for evaluating teaching with simulations of the doctor-patient interaction are not yet available. General questionnaires for student evaluation of teaching seem to be unsuitable for two reasons:

- On the one hand, many of the forms were developed by universities and university hospitals themselves and have not been made public, at least not through journals with peer-review process. Therefore, the test accuracy of these methods cannot be evaluated.
- Secondly, important aspects of simulations are not covered. Because this method is not used in conventional courses, student evaluation questionnaires generally do not query variables such as the perform-

ance of actors, the quality of feedback and improvement of communication skills.

In order to measure the students' satisfaction with courses simulating the doctor-patient interaction, this project developed a new questionnaire, the **Student Evaluation Scale for Courses with Simulations of the Doctor-Patient Interaction (SES-Sim)**.

SES-Sim was developed as part of the educational "PJ-Start-Block" project (key skills training and application in realistic daily routines) at the Medical Faculty of the University of Cologne. A course using simulations of the doctor-patient interaction as a central teaching method was designed and implemented in this project. It consists of a one-week block course at the end of the undergraduate degree in preparation for the Practical Year (PJ). It mimics everyday life on a hospital ward, with students assuming the roles of doctors who interact with actor-patients in different scenarios. The disease patterns of heart disease, appendicitis, ectopic pregnancy, chronic back pain, diabetes, urinary tract infection with fever and lung cancer are presented.

The role scripts are usually designed not only to present somatic symptoms but also psychosocial aspects which are relevant to the treatment of diseases. For example, the simulation of an ectopic pregnancy specifies that the patient comes from a Muslim background and has an understanding of diseases which cannot be reconciled with scientific concepts. The role of the female diabetic patient includes problems with insulin therapy which are due to the fact that a comorbid depressive reaction occurs. The students are therefore faced with the task of responding appropriately, not only to the somatic aspects but also to disease-related psychosocial issues.

In addition to the simulations, exercises are offered about medical tasks without patient contact (e.g. morning and afternoon meetings of the medical team, writing medical reports and requesting consultations).

Methods

Item Composition

To develop the new questionnaire, 33 items were compiled which mirror the central elements of the course contents. Some items were adapted from items taken from the questionnaires for student evaluation of university courses by Diehl [10]. Most items use new wording. As a response format, a 5-point Likert scale with endpoints 1 ("strongly disagree") and 5 ("strongly agree") was defined.

Sample

The compilation of the items was presented to the first participants of the PJ-Start-Block in the winter semester 2009/10 and the summer semester 2010 on the last day of the block course. A total of 220 medical students

filled in the questionnaire (66% women, 34% men), the average age being 26.

Statistical Analysis

Item selection was based on three factor analyses. The data of the samples from the winter and the summer semester were first separately subjected to a factor analysis. Based on the results of both factor analyses, items were selected for the final version of the scale. The item selection was statistically confirmed by subjecting the final version to another factor analysis, having combined the participants from the winter and summer semester into a single sample.

Results

(Note on the English version of this paper: Statistical analysis was done for the original version of SES-Sim which is in German. All statistical parameters presented in this paper are therefore valid for the German version, but not for the translation of items presented in Table 1). In the 33 item version which was presented to the students in the first round during the winter semester, nine factors with an eigenvalue > 1 appear in the factor analysis (eigenvalue distribution 8.61, 3.14, 2.65, 2.12, 1.57, 1.49, 1.22, 1.14, 1.02, .98 ...). Of these nine factors, five can be meaningfully interpreted as relating to content; their content relates to the learning success, the actors, the premises, the tutors and the students. The remaining four factors either consist of a single item or have heterogeneous contents. For the five interpretable factors, 23 of the 33 items load in a unique way, i.e. they load high on one factor and low on the other.

For the second round during the summer semester, an abridged version of the questionnaire was created which consisted of the 23 items which could be mapped unambiguously. The factor analysis of the 23 item version showed seven factors with an eigenvalue > 1 (eigenvalue distribution: 5.37, 2.33, 1.99, 1.53, 1.49, 1.31, 1.16, .95 ...). The first five correspond to the interpretable factors found in the first factor analysis. The other two factors are of heterogeneous content. Of the 23 items, 17 load uniquely on the factor to which they had been mapped to in the first factor analysis. The other six items either have double loads or load on a different factor compared to the first round.

The final selection of items was carried out according to the following criteria: 16 of the 17 items which in both factor analyses could be clearly mapped to one of the five interpretable factors were included in the final version of the questionnaire. One of the 17 items was excluded despite clear factor loads because it has a strong semantic similarity with another item. (Both items are related to the transfer of skills and future professional practice; for reasons of economy, a decision was made not to evaluate the same thing twice.) Two items which according to the criterion of clear association in both factor

Table 1: Item Parameters

		a ¹	a ²	a ³	a ⁴	a ⁵	r _{it}
1	I learned a lot of new things on the course.	.64	.00	.11	.20	.08	.54
2	I gained confidence in terms of medical practice through this course.	.85	.05	.07	.10	.02	.75
3	The course has increased my medical competence.	.81	.01	.15	.05	.04	.69
4	The course has increased my competence in conducting medical patient-doctor communication.	.77	.21	.02	.13	.01	.65
5	What I learned in this course will come in handy in the future when practicing medicine.	.68	.18	-.02	.18	.09	.60
6	The actors seemed to be well prepared for their roles as patients.	.15	.83	-.10	.00	.10	.66
7	The actors gave helpful feedback.	.05	.87	.05	-.03	-.01	.70
8	The actors were credible in their role as patients.	.08	.61	.11	.10	.06	.45
9	Contacts with the patient-actors were very realistic.	.07	.78	.19	.00	.00	.59
10	The premises were adequate.	.18	.02	.87	.01	.03	.72
11	The premises were equipped adequately.	.15	.10	.88	-.03	.11	.76
12	The premises and equipment appeared real.	-.05	.14	.80	.13	.10	.61
13	The tutor appeared to be well prepared.	.12	-.03	.18	-.17	.74	.42
14	The tutor monitored the course competently.	.06	.14	.14	.12	.84	.58
15	The tutor was interested in the students' success.	.02	.04	-.08	.29	.70	.36
16	The students participated actively.	.20	-.06	-.04	.76	.08	.48
17	The students gave useful feedback.	.17	-.04	.11	.83	-.02	.59
18	The students appeared alert and interested.	.21	.23	.06	.59	.15	.36

a¹⁻⁵ Load on factor 1-5 after varimax rotation, r_{it} = discriminatory power

analyses should have been excluded were left in the final selection for substantive reasons. The two items are used to evaluate the feedback, a central element of the course which should definitely be included in the new questionnaire. So overall 18 items were selected for the final version of the questionnaire.

The factor analysis for statistical confirmation of the 18-item version for the total sample confirmed the expected 5-factor structure (eigenvalue distribution: 4.39, 2.30, 2.07, 1.67, 1.28, .90 ...). The five factors together explain 65% of the variance. Table 1 shows the item parameters. The 18 items are clearly mapped to the five factors. All items load high onto one factor and low to the other factors. The items continue to be selective, the item-scale correlations are usually well above the limit of .30.

Table 2 shows the parameters of the five subscales of SES-Sim. The mean agreement is high on all scales, which points to the course being assessed positively by the students. The internal consistencies range from satisfactory to good.

Table 2: Scale Parameters

	No of Items	N	M	Md	SD	α
Learning success	5	220	4.16	4.30	.67	.84
Actors	4	220	4.62	4.75	.48	.77
Premises	3	220	4.65	5.00	.60	.83
Tutor	3	220	4.91	5.00	.23	.63
Students	3	220	4.59	4.67	.45	.66

N = Number of study participants, M = Mean, Md = Median, SD = Standard deviation, α = Internal consistency (Cronbach's Alpha)

Table 3 shows how the five subscales of SES-Sim are related with the overall assessment of the block course. The overall evaluation was performed using German school grades on a scale from 1 ("very good") to 5 ("poor"). Because the subscales of SES-Sim do not exhibit normal distribution (the distributions are right-skewed), non-parametric correlations were calculated (Spearman's rank correlation coefficient).

Table 3: Correlation of scales with the overall evaluation

	Overall evaluation
Learning success	-.49***
Actors	-.29***
Premises	-.24**
Tutors	-.38***
Students	-.31***

*** p < .001, ** p < .01

It turns out that the overall evaluation of the course is above all related to the perceived learning success. This means that the more students have the impression of having learned something, the more they judge the course to be positive. The evaluations of the performance of all participating groups (tutors, actors and the students themselves) including the premises also significantly relate to the overall result. Thus, all dimensions of SES-Sim prove to be significant for the overall evaluation of the event.

Discussion

SES-Sim represents a measuring instrument which enables fast and economical surveys of student satisfaction with courses which use simulations of doctor-patient interactions. Kirkpatrick and Kirkpatrick [11] who distinguish between the four hierarchical levels of reaction, learning, behaviour and results in evaluations assign this form of evaluation to level one. Satisfaction measurement therefore gives information about the response of the participants to a training measure, but still says nothing about the extent of the knowledge increase, behaviour change and increased productivity or the quality of work. The high degree of satisfaction which is reflected in the high average values of the scale corresponds to the experiences from other universities where students also rated this type of course very positively [5], [6], [7]. In addition to the simulations being highly practical, the novelty factor of this methodology in medical school is also bound to play a role. The intensity of support, due to small group sizes, which is not the case in other courses, is also likely to have a favourable impact on the evaluation.

The SES-Sim items relate directly to the central elements of courses using simulations of doctor-patient interactions. In addition, it was ensured that the content was formulated clearly. Therefore on the face of it, the new questionnaire can certainly be regarded as valid.

But construct validation through an external criterion would be difficult to perform. One reason for this is that the method of simulation trains complex skills which span multiple dimensions. For example, if one wanted to evaluate the learning progress in communicative competence, which according to Kirkpatrick and Kirkpatrick [11] represents an evaluation on the second of the four levels, the first step would require identification of the dimensions which constitute this ability. In the second step, appropriate operationalisations would have to be found

for each dimension. This approach would be expensive and fraught with great uncertainty. Even determining whether all dimensions relevant for communicative competence have been identified would be difficult. For this reason it was decided not to conduct a construct validation of SES-Sim in this project.

Another limitation which not only applies to this questionnaire but to all measuring instruments for student evaluation of courses is that the satisfaction of students only partially reflects the actual quality of a course [12]. It turns out, for example, that medical students assess courses on basic subjects less positively than clinical subjects [13]. These judgements appear to be guided less by the didactic quality of the event but rather by what is viewed as useful for a future career. It is known in psychology that lectures and seminars on statistics regularly are judged less positively than classes in other subjects because this subject runs counter the concepts of psychology of freshmen in particular [10]. Student evaluations are highly dependent on the general popularity of a subject; less popular subjects are regularly assessed less positively than popular ones.

It would therefore be unwise to rely solely on student evaluation in order to assess the value of a course; in such cases additional criteria should be used, for example surveys of student progress. But if teaching staff wants to get a quick overview as to how their classes are judged by the students and what could be improved, then surveys such as SES-Sim can be helpful.

Acknowledgements

Our thanks Dr. Valentin Goede, Houda Hallal, Wencke Johannsen, Ortrun Kliche, Dr. Sabine Teschendorf and Christian Thrien, University Hospital Cologne and University of Cologne for their helpful feedback on this paper. PJ-Start-Block is a teaching project by the University of Cologne in collaboration with the following bodies: Medical Faculty: Office of the Dean of Studies and Inter-professional Skills Lab and Simulation Centre Cologne (Prof. Lehmann, Dr. Boldt, Dr. h.c. (RUS) Stosch), Institute for History and Ethics of Medicine (Prof. Karenberg, Prof. Schäfer), Institute for Pharmacology (Prof. Herzig, PD Dr. Matthes), Clinic and Policlinic for Psychosomatics and Psychotherapy (PD Dr. Albus, Prof. Obliers, Dr. Koerfer), Centre for Palliative Medicine (Prof. Voltz, PD Dr. Schiessl) and Faculty of Humanities: Institute for Comparative Research in Education and Social Sciences (Prof. Allemann-Ghionda)

Competing interests

The authors declare that they have no competing interests.

References

1. Swanson DB, Stillman PL. Use of standardized patients for teaching and assessing clinical skills. *Eval Health Prof.* 1990;13:79-103.
2. Ainsworth MA, Rogers LP, Markus JF, Dorsey NK, Blackwell TA, Petrusa ER. Standardized patient encounters. A method for teaching and evaluation. *J Amer Med Assoc.* 1991;226:1390-1396.
3. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med.* 1993;68(6):443-453.
4. Ortwein H, Fröhmel A, Burger W. Einsatz von Simulationspatienten als Lehr-, Lern- und Prüfungsform. *Psychother Psych Med.* 2006;56:23-29.
5. Bachmann C, Barzel A, Dunkelberg S, Schrom K, Ehrhardt M, van den Bussche H. Fachübergreifendes Kommunikationstraining mit Simulationspatienten: ein Pilotprojekt ins Curriculum. *GMS Z Med Ausbild.* 2008;25(1):Doc58. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2008-25/zma000542.shtml>
6. Koerfer A, Köhle K, Obliers R, Sonntag B, Thomas W, Albus C. Training und Prüfung kommunikativer Kompetenz. Aus- und Fortbildungskonzepte zur ärztlichen Gesprächsführung. *Gesprächsforsch Z verbal Interaktion.* 2008;9:34-78.
7. Nikendei C, Zipfel S, Roth C, Löwe B, Herzog W, Jünger J. Kommunikations- und Interaktionstraining im psychosomatischen Praktikum: Einsatz von standardisierten Patienten. *Psychother Psych Med.* 2003;53:440-445.
8. Schildmann J, Härlein J, Burchardi N, Schlögl M, Vollmann J. Breaking bad news: evaluation study on self-perceived competences and views of medical and nursing students taking part in a collaborative study. *Support Care Cancer.* 2006;14(11):1157-1161.
9. Simmenroth-Nayda A, Weiß C, Chenot J-F, Scherer M, Fischer T, Kochen MM, Himmel W. Verbessern Anamneseübungen die kommunikativen Fähigkeiten von Studierenden? Ein Prä-Post-Vergleich. *GMS Z Med Ausbild.* 2007;24(1):Doc22. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2007-24/zma000316.shtml>
10. Diehl JM. Normierung zweier Fragebögen zur studentischen Beurteilung von Vorlesungen und Seminaren. *Psychol Erz Unterr.* 2003;50:27-42.
11. Kirkpatrick DL, Kirkpatrick JD. Evaluating training programs: The four levels. 3rd ed. San Francisco: Berrett-Koehler Publishers; 2006.
12. Kromrey H. Wie erkennt man „gute Lehre“? Was studentische Vorlesungsbefragungen (nicht) aussagen. *Empirische Pädagogik.* 1994;8:153-168.
13. van den Bussche H, Weidtmann K, Kohler N, Frost M, Kaduszkiewicz H. Evaluation der ärztlichen Ausbildung: Methodische Probleme der Durchführung und der Interpretation von Ergebnissen. *GMS Z Med Ausbild.* 2006;23(2):Doc37. Zugänglich unter: <http://www.egms.de/static/de/journals/zma/2006-23/zma000256.shtml>

Corresponding author:

Dr. Eva Neumann
University of Cologne, Department of Psychosomatics and Psychotherapy, Kerpener Straße 61, 50931 Cologne, Germany, Tel.: +49 (0)221/478-4103, Fax: +49 (0)221/478-3103
eva.neumann@uni-duesseldorf.de

Please cite as

Neumann E, Obliers R, Schiessl C, Stosch C, Albus C. Student Evaluation Scale for Medical Courses with Simulations of the Doctor-Patient Interaction (SES-Sim). *GMS Z Med Ausbild.* 2011;28(4):Doc56. DOI: 10.3205/zma000768, URN: [urn:nbn:de:0183-zma0007688](http://nbn-resolving.org/urn:nbn:de:0183-zma0007688)

This article is freely available from

<http://www.egms.de/en/journals/zma/2011-28/zma000768.shtml>

Received: 2010-12-01

Revised: 2011-07-14

Accepted: 2011-07-27

Published: 2011-11-15

Copyright

©2011 Neumann et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en>). You are free to Share – to copy, distribute and transmit the work, provided the original author and source are credited.