

Wirken sich unterschiedliche medizinische Curricula auf das selbst eingeschätzte klinische Denken von Studierenden aus?

Zusammenfassung

Zielsetzung: Das klinische Denken ist eine zentrale Grundlage des ärztlichen Handelns und sollte durch das Studium in den humanmedizinischen Studiengängen gefördert werden. Welche Lehr- und Lernformen hierzu vor allem beitragen, ist bis heute allerdings noch nicht abschließend gezeigt. Zur Messung der wissensunabhängigen Komponente des klinischen Denkens wurde das *Diagnostic Thinking Inventory* (DTI) entwickelt. Die vorliegende Pilotstudie untersucht, ob dieses Instrument dazu geeignet ist, Unterschiede im klinischen Denken von Studierenden in unterschiedlichen Ausbildungsabschnitten dreier medizinischer Curricula aufzuzeigen.

Methodik: Das *Diagnostic Thinking Inventory* (DTI) umfasst 41 Items in zwei Subskalen („Flexibilität des Denkens“ und „Strukturierung des Wissens“). Jedes Item enthält eine Aussage oder Feststellung über das klinische Denken in Form eines Stammes und darunter jeweils eine 6-Punkt-Skala an deren beiden Enden gegensätzliche Aussagen stehen. Zwischen diesen beiden Extremen sollen die Probanden ihr eigenes klinisches Denken einordnen. Die deutsche Übersetzung des DTI wurde von 247 freiwilligen Studierenden aus drei Fakultäten und verschiedenen klinischen Semestern ausgefüllt. In einem quasi experimentellen Design waren 219 Teilnehmer aus Regel- und Modellstudiengängen in Nordrheinwestfalen beteiligt, genauer aus dem 5., 6. und 8. Semester des Modellstudiengangs der Universität Witten/Herdecke (UW/H), aus dem Modell- (7. und 9. Semester) und Regelstudiengang (7. Semester) der Ruhr-Universität Bochum (RUB) und aus dem Modellstudiengang (9. Semester) der Universität zu Köln (UzK). Die gewonnenen Daten wurden quantitativ ausgewertet.

Ergebnisse: Die Reliabilität des Gesamt-Fragebogens war gut (Cronbachs alpha zwischen 0,71 und 0,83); die Reliabilität der Subskalen lag zwischen 0,49 und 0,75. Die unterschiedlichen Gruppen wurden mittels Mann-Whitney Test miteinander verglichen. Dabei wurden signifikante Unterschiede sowohl zwischen Semesterkohorten innerhalb einer Fakultät als auch zwischen Studierenden vergleichbarer Jahrgänge unterschiedlicher Fakultäten gemessen. Innerhalb des Modellstudiengangs an der UW/H nahm die Punktzahl vom 5. zum 6. und vom 5. zum 9. Semester zu. Zwischen den einzelnen Kohorten der RUB konnten weder Unterschiede zwischen Modell- und Regelstudiengang noch zwischen dem 7. und dem 9. Semester des Modellstudiengangs gemessen werden. Vergleicht man alle beteiligten höchsten Semester, so erreicht das 8. Semester an der UW/H die größte Punktzahl, die signifikant höher ist als die des 9. Semesters an der RUB oder auch des 9. Semesters an der UzK. Das 9. Semester der RUB liegt signifikant höher als das 9. Semester der UzK.

Diskussion: Die deutsche Fassung des DTI misst selbst eingeschätzte Unterschiede im diagnostischen Denken bei Studierenden aus unterschiedlichen Semestern und aus verschiedenen medizinischen Modell- und Regelstudiengängen mit zufriedenstellender Zuverlässigkeit. Die Ergebnisse lassen sich vor dem Hintergrund der unterschiedlichen Curricula diskutieren. Damit eignet sich der DTI für weiterführende Un-

Kirsten Gehlhar¹

Kathrin Klimke-Jung²

Christoph Stosch³

Martin R. Fischer⁴

1 Carl von Ossietzky Universität Oldenburg, Fakultät für Medizin und Gesundheitswissenschaften, Oldenburg, Deutschland

2 Ruhr-Universität Bochum, Bochum, Deutschland

3 Universität zu Köln, Köln, Deutschland

4 Klinikum der LMU München, Institut für Didaktik und Ausbildungsforschung in der Medizin, München, Deutschland

tersuchungen, die dann mit den unterrichtsmethodischen Charakteristika und Outcomes verschiedener Curricula korreliert werden können.

Schlüsselwörter: clinical thinking, clinical reasoning, PBL, diagnostic thinking inventory

Einleitung

Das klinische Denken ist eine zentrale Komponente ärztlicher Kompetenz. Eine optimale Patientenversorgung hängt von der sorgfältigen Analyse der vom Patienten gegebenen Informationen und der Abwägung zwischen dem Nutzen und den Risiken von diagnostischen Tests und Therapien ab.

Daher hat auch jede universitäre Ausbildung von Medizinstudierenden das Ziel, Absolventen mit guten klinischen Denkfähigkeiten auszubilden. Doch ob eine bestimmte Form des Unterrichts diese Fähigkeiten besonders gut fördert, ist bis heute nicht eindeutig nachgewiesen.

Klinische Probleme unterscheiden sich grundsätzlich von wohlstrukturierten Aufgaben. Zu Beginn des Prozesses liegen noch nicht alle notwendigen Informationen vor, und die Art eines Problems kann sich während des diagnostisch-therapeutischen Prozesses (Anamnese und Untersuchung) dynamisch verändern. Es handelt sich um komplexes Problemlösen, weil es keine standardisierten Vorgehensweisen gibt, die sicher zu einer Lösung führen; stattdessen sind sie für jedes Problem einzigartig, und der Arzt kann nie ganz sicher sein, dass die gefundene Lösung auch tatsächlich die richtige ist [1], [2].

Modelle klinischen Denkens

In der Ausbildungsforschung wird seit den 70er Jahren intensiv auf diesem Gebiet gearbeitet. Das Ziel ist, besser beschreiben zu können, welche Prozesse während des klinischen Denkens ablaufen und wie sich Experten von Novizen unterscheiden. Daraus sollen Schlüsse gezogen werden, die aufzeigen, wie man klinisches Denken besser unterrichten kann.

Im Laufe dieser Zeit wechselten viele Konzepte zur Erklärung des Expertise-Erwerbs im klinischen Denken einander ab und wurden weiterentwickelt [3]. Ein Konzept ist z.B. das der sogenannten „illness scripts“ [4]. Letztere sind Repräsentationen von Problemen, z.B. von Krankheiten, Syndromen oder Gruppen von Krankheiten zusammen mit den Bedingungen, unter denen sie auftreten, ihren Manifestationen, Diagnosen und Therapiekonzepten, ebenso wie ihre pathophysiologischen Grundlagen. Andere Modelle betonen, dass medizinische Expertise besonders durch intensive Praxiserfahrungen erreicht wird [5], [6]. Es konnte empirisch gezeigt werden, dass Fähigkeiten, die sich dem klinischen Problemlösen zuordnen lassen, bei Experten elaborierter sind als bei Novizen [7]. Dabei zeigte sich z.B., dass Expertenwissen v.a. durch Wahrscheinlichkeiten verknüpft ist [4], oder dass es geprägt ist von Mustern, die spontan erkannt und dann wiederum hypothetisch-deduktiv bestätigt werden [2].

Des Weiteren sind Experten besser in der Anwendung von sogenannten „semantic qualifiers“, die ein Symptom auf einer bipolaren Skala einordnen oder in der Qualität von in einer Patientengeschichte erkannten Schlüsselmerkmalen [8].

Daneben ist das klinische Denken aber auch von Inhalt und Kontext abhängig [9] und Expertise auf einem Gebiet bedeutet nicht ein vergleichbares Können in einer anderen Spezialdisziplin oder in einem anderen Patientenfall. Selbst innerhalb eines Gebietes ist Expertise nicht generalisierbar [10], [11].

All diese Ansätze werden nicht mehr als sich gegenseitig ausschließend angesehen [12]. Experten nutzen demzufolge parallel sowohl analytische (deduktive, kontrollierte) als auch nicht-analytische Prozesse (unbewusste, spontane) zur Lösung ihrer Patientenfälle, denn das effektive klinische Problemlösen basiert einerseits auf klaren Abläufen der Informationssammlung, Hypothesenbildung und Hypothesentestung. Andererseits ist zur Diagnosestellung ein Kenntnis der zugrundeliegenden Mechanismen notwendig, und dieses Wissen muss problembezogen integriert und mental organisiert werden. Das Wiedererkennen von Mustern beschleunigt das Wiederabrufen. Beide Prozesse ergänzen sich und es findet ein bidirektionaler Austausch oder Informationsfluss statt [13].

Unterricht im klinischen Denken?

Die Frage, wie man Studierende in diesen komplexen Fähigkeiten optimal unterrichten kann, ist bis heute aus gutem Grund nicht eindeutig beantwortet.

Studierende, denen beispielsweise die Methode der Heuristik für konkurrierende Hypothesen (Bayes' Theorem) beigebracht wurde, konnten diese Methode nach der Unterrichtseinheit zwar gut anwenden, scheiterten später in der Klinik jedoch an der Übertragung dieses abstrakten Modells in die Praxis [14]. In anderen Studien konnte gezeigt werden, dass sich die Arbeit an didaktisch ausgewählten Patientenfällen mit Fehlern und elaboriertem Feedback oder an Fallbeispielen mit Instruktion positiv auf das klinische Denken in der jeweiligen Versuchssituation auswirkt [15], [16], [17].

Zwei Lehr- und Lernumgebungen haben sich bisher fast übereinstimmend als effektiv zur Förderung des klinischen Denkens erwiesen: einerseits problemorientiertes Lernen (PbL) und klinische Praktika/klinische Erfahrung. PbL soll deshalb dazu beitragen, dass Studierende besser das klinische Denken erlernen [18], [19], [20], da im PbL der Problemlöseprozess besonders gut erlernt und geübt wird [21]. PbL verbindet Inhalt mit Kontext, und die Anwendung von Wissen auf klinische Probleme fördert die Entwicklung von zusammenhängenden, pathophysiologischen Konzepten [21], [22]. Eine Meta-Analyse [23] zur

Frage, welche Unterrichtsformen das klinische Denken fördern können, fand zwei Untersuchungen, die zeigen, dass sich Studierende durch PbL im kritischen Denken verbessern [24], [25], und dass sie gegenüber anderen Studierenden akkuratere, zusammenhängende und umfassendere Erklärungen für medizinische Probleme liefern [22]. Allerdings ist kritisches Denken nicht notwendigerweise mit klinischem Denken gleichzusetzen.

Die konkrete klinische Erfahrung mit Patienten ist ein weiterer unverzichtbarer Faktor für die Entwicklung klinischen Denkens. Erstens müssen Studierende ihre erlernten Fähigkeiten auch üben [26], zweitens wird klinisches Denken stark durch Erfahrung beeinflusst [12] und ist eine Konsequenz von sich daraus entwickelndem multidimensionalem Wissen. Es hat sich gezeigt, dass der Entwicklungsstand von Studierenden oft eher ihrer klinischen Erfahrung z.B. durch Famulaturen als dem Studienjahr entspricht [27], [28].

Auf Basis dieser Erkenntnisse lässt sich vermuten, dass traditionelle Curricula, in denen Fächer oft isoliert unterrichtet werden, kein oder nur wenig PbL eingesetzt wird und klinische Erfahrung erst später und oft in geringerem Umfang stattfindet, die Studierenden weniger Gelegenheit haben, das klinische Denken zu lernen und zu üben.

Messung des klinischen Denkens

Um das klinische Denken per Selbsteinschätzung zu messen, wurde 1990 von Bordage und Marsden ein entsprechendes Inventar entwickelt [29], das sog. Diagnostic Thinking Inventory (DTI). Bis dahin wurden Inventare verwendet, die z.B. kritisches (nicht klinisches) Denken gemessen haben. Probanden mussten dabei mündlich oder schriftlich ihre Denkprozesse erläutern oder es wurde die Lösung von konkreten klinischen Problemen getestet. Das DTI ist dagegen ein Instrument, das unabhängig vom Kontext die selbst eingeschätzte Flexibilität des Denkens und die Strukturierung des Wissens misst und zwischen Medizinern verschiedener Ausbildungsstufen differenzieren kann. Es besteht aus 41 Items, in denen die Probanden ihr Denken in vorgegebenen Situationen auf einer Skala einschätzen. Die Antworten wiederum repräsentieren eine bestimmte Stufe des klinischen Denkens. Dies wurde ursprünglich anhand von 270 Versuchspersonen in unterschiedlichen Ausbildungsstadien von Erstsemester-Studierenden bis hin zu erfahrenen Ärzten untersucht. Es konnten signifikante Unterschiede zwischen den Studierenden und den Ärzten aufgezeigt werden, die Unterschiede innerhalb der Ärzte mit verschieden großer Erfahrung waren dagegen nicht signifikant.

Das DTI-Inventar wurde in den nachfolgenden Jahren in verschiedenen Studien eingesetzt und validiert [30] und auch im Zusammenhang mit anderen kognitiven oder psychometrischen Tests untersucht [31].

Insgesamt sind die bisherigen mit dem DTI erlangten Studienergebnisse nicht einheitlich. In der Mehrheit der Studien verbessern sich allerdings durch Unterricht zum klinischen Problemlösen, zu Fehlern bei der Diagnosefin-

dung, durch Bearbeitung von Patientenfällen, Teilnahme an diagnostischen Fallbesprechungen oder zunehmende Studiendauer die DTI-Ergebnisse von Studierenden signifikant [32], [33], [34], [35], [36]. In zwei anderen Studien zeigte sich dagegen keine Verbesserung im DTI durch die vorgenommenen Interventionen [26], [37].

Das Inventar misst die Selbsteinschätzung der Probanden über Art und Struktur ihres klinischen Denkens, nicht jedoch ihre tatsächliche Diagnosefähigkeit. Daher ist die Korrelation zwischen Ergebnissen des DTI und gelösten Fällen oder erstellten Diagnosen oft nur gering [6], [32]. Zur Erstellung der korrekten Diagnose muss zusätzlich zur Fähigkeit zum klinischen Denken auch eine entsprechende Wissensbasis im jeweiligen Fachgebiet und klinische Erfahrung kommen.

Zielsetzung

Die Medizincurricula der verschiedenen Fakultäten unterscheiden sich deutlich in ihren Zielsetzungen und Schwerpunkten. Alle wollen jedoch – auf jeweils unterschiedlichen Wegen – gute Ärztinnen und Ärzte ausbilden. Längerfristig muss daher gezeigt werden, inwieweit sich verschiedene Curricula auf den Erwerb des klinischen Denkens als ärztliche Kernkompetenz auswirken. In einem ersten Schritt wurde in dieser Pilotstudie an drei verschiedenen Medizinischen Fakultäten in NRW untersucht, inwiefern Unterschiede zwischen Studierenden verschiedener Curricula einerseits und zwischen unterschiedlichen Ausbildungsstufen andererseits zu messen sind.

Folgende Fragestellungen sollten beantwortet werden:

1. Kann mittels DTI nachgewiesen werden, ob die Kompetenz des klinischen Denkens in der Selbsteinschätzung der Studierenden im Laufe des Studiums zunimmt?
2. Bestehen Unterschiede in der Selbsteinschätzung bezüglich des klinischen Denkens zwischen Studierenden verschiedener Studiengänge und Curricula?

Es wäre zu vermuten, dass die Kompetenz des klinischen Denkens im Laufe des Studienfortschritts, also mit steigender Semesterzahl zunimmt. Unterschiede bei Studierenden unterschiedlicher Curricula würden sich zeigen, wenn sich z.B. die Anteile der Elemente wie PbL oder klinische Praktika unterscheiden. Die Studierenden der UW/H und des Modellstudiengangs der RUB durchlaufen Modellcurricula, die u.a. durch problemorientiertes Lernen (PbL) von Beginn an und – besonders an der UW/H – durch lange und zahlreiche klinische Praktika gekennzeichnet sind. An der UW/H ist PbL das zentrale Element in den ersten vier Semestern. Zusätzlich ist an der UW/H der Anteil an curricular verankerter klinischer Erfahrung von allen Studiengängen am höchsten. Ab der zweiten Hälfte des vierten Semesters bis zum Praktischen Jahr werden allein 46 Wochen in klinischen Blockpraktika und sechs Wochen in allgemeinmedizinischen Hospitationen verbracht.

Die beiden Studiengänge an der RUB unterscheiden sich hinsichtlich ihres Curriculums. Der Regelstudiengang hat ein fächerbezogenes traditionelles Curriculum mit einer sechswöchigen PbL-Insel im 4. Semester. Der Modellstudiengang, ist themenbasiert aufgebaut und praxisbezogen. PbL ist ein strukturierendes Studienelement in den ersten vier Semestern und wird danach studienbegleitend in den Semestern 5, 8 und 9 angeboten. Auch bezüglich ihrer Studierendenzahl unterscheiden sich die beiden Studiengänge. Während pro Studienjahr 42 Studierende in den Modellstudiengang aufgenommen werden, studieren im Regelstudiengang ca. 200 Studierende pro Jahr. Die Befragung schloss jeweils alle anwesenden Studierenden im 7. und 9. Semester des Modellstudiengangs ein sowie zwei Seminargruppen der Lehrveranstaltung für Allgemeinmedizin im Regelstudiengang des 7. Semesters.

Der Modellstudiengang in Köln bietet PbL nur in einem (ersten) Semester und später nur in einzelnen Fächern an.

Methodik

Teilnehmer, Erhebungsinstrument

Zur Messung der Selbsteinschätzung der wissensunabhängigen Komponente des klinischen Denkens wurde das DTI (Diagnostic Thinking Inventory) herangezogen. Dazu wurde der ins Deutsche übersetzte Fragebogen, der freundlicherweise von Dr. Stieger [36] zur Verfügung gestellt wurde, in Modellstudiengängen an drei Fakultäten in Nordrhein-Westfalen (Bochum, Köln und Witten/Herdecke) und einem Regelstudiengang (Bochum) in maschinenlesbarer Form eingesetzt.

Zur Beantwortung der Frage nach der Verbesserung des klinischen Denkens im Laufe des Studiums wurden unterschiedliche Semester eines Studiengangs mit einbezogen. Dies war für den Modellstudiengang der UW/H und im Modellstudiengang an der RUB der Fall. Zur Messung der Unterschiede im selbst eingeschätzten klinischen Denken zwischen Studierenden verschiedener Studiengänge und Curricula wurden fortgeschrittene Studierende an allen drei Standorten vor dem Praktischen Jahr befragt. An der UW/H stand zum Zeitpunkt der Studie nur ein 8. Semester als höchstes Semester vor PJ-Beginn zur Verfügung, da diese Studierenden zu einer Zeit ihr Studium begonnen hatten, als nur zum Sommersemester Einschreibungen möglich waren. An der RUB und an der UzK konnten Studierenden des 9. Semesters einbezogen werden (Studienbeginn zum Wintersemester).

Die Fragebögen wurden an den beteiligten Universitäten im Zeitraum von Oktober bis Dezember 2010 zu jeweils einem Zeitpunkt in Präsenzveranstaltungen ausgeteilt und direkt wieder eingesammelt, in denen das jeweilige Semester möglichst vollständig anwesend war.

Die Studierenden wurden direkt vor dem Austeilen des Fragebogens über den Zweck des Projektes aufgeklärt, die Teilnahme erfolgte freiwillig und die Befragung war

anonym. Eine gesonderte Einwilligungserklärung wurde von den Teilnehmern nicht ausgefüllt.

Auf Wunsch wurde den Studierenden nach der Auswertung ihre erreichte Punktzahl mit entsprechenden Erläuterungen mitgeteilt. Dazu mussten sie die Nummer des Fragebogens angeben, den sie ausgefüllt hatten, die Ergebnisse wurden dann unter der Fragebogennummer zur Verfügung gestellt.

Instrument

Der DTI-Fragebogen misst die Selbsteinschätzung der wissensunabhängigen Komponente des klinischen Denkens und besteht aus insgesamt 41 Fragen (Beispielfragen siehe Tabelle 1). Dabei lassen sich 20 Fragen der Subskala „Flexibilität des Denkens“ und 21 Fragen der Subskala „Strukturierung des klinischen Denkens“ zuordnen (Beispiele s. unten). Die Flexibilität des Denkens bildet die Fähigkeit des Teilnehmers ab, während des Diagnoseprozesses auf die richtige Diagnose zu kommen und flexibel neue Informationen zu verarbeiten. Die Struktur des Denkens spiegelt wider, wie gut das klinische Wissen, auf das während des Diagnoseprozesses zurückgegriffen wird, organisiert und verfügbar ist. Jede Frage besteht aus einem Fragenstamm (meist einer Feststellung) und einer Antwortskala. Die Antwortskala bietet zwei entgegengesetzte Antworten/Aussagen zu der Eingangsfeststellung mit sechs Auswahlfeldern zum Ankreuzen dazwischen. Die Teilnehmer sollen das Feld ankreuzen, das am besten ihre Einstellung auf der Skala zwischen den Antwortmöglichkeiten widerspiegelt. Die Fragen sind willkürlich „linksbündig“ bzw. „rechtsbündig“, d.h. die Antwort, die ein weiter fortgeschrittenes klinisches Denken wiedergibt, ist unterschiedlich mal links oder rechts aufgeführt.

Beispiel für eine Frage aus der Subskala Flexibilität (siehe Abbildung 1).

Beim Anamnesegespräch

bleibe ich oft bei einer möglichen Diagnose hängen.

fällt es mir leicht, unterschiedliche Diagnoseansätze abzuklären.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Abbildung 1

Beispiel für eine Frage aus der Subskala Struktur (siehe Abbildung 2).

Während ich Informationen zur Anamnese sammle,

kann ich diese bereits umfassend einteilen und klassifizieren.

fällt es mir schwer, diese zu bewerten und den Bezug zueinander herzustellen.

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Abbildung 2

Tabelle 1: Anzahl und Zuordnung der Fragebögen

Universität	Art des Studiengangs	Semester	Zahl der ausgeteilten Fragebögen	Zahl der ausgefüllten Fragebögen	Zahl der auswertbaren Fragebögen	Verteilung weiblich/männlich [%]
Witten/Herdecke	Modell	5	42	33	31	58,6 / 41,4
Witten/Herdecke	Modell	6	42	25	24	66,6 / 33,4
Witten/Herdecke	Modell	8	42	25	23	56,5 / 43,3
Ruhr-Universität	Modell	7	30	28	20	50,0 / 50,0
Ruhr-Universität	Regel	7	40	31	20	53,5 / 46,5
Ruhr-Universität	Modell	9	37	36	29	68,0 / 32,0
Universität zu Köln	Modell	9	150	141	101	58,5 / 41,5

Auswertung

Alle ausgefüllten Fragebögen wurden gescannt und mittels Analysesoftware (FormPro, Version 2.5) eingelesen. Zur Auswertung wurden nur diejenigen Fragebögen einbezogen, bei denen alle Fragen eindeutig beantwortet worden waren.

Die erreichten Punkte wurden ausgerechnet, indem den Antwortfeldern die Werte 1 bis 6 zugeordnet wurden. Die Antwort, die das ausgeprägteste klinische Denken darstellte, erhielt jeweils die höchste Punktzahl. Es wurden sowohl die erreichte Punktzahl (maximal waren 246 Punkte erreichbar) als auch die Subskalen Flexibilität (max. 120 Punkte) und Struktur (max. 126 Punkte) für jeden Teilnehmer berechnet.

Für die Ermittlung der internen Konsistenz des DTI wurde Cronbachs alpha berechnet. Die Effektstärke wurde als Cohens d berechnet.

Alle Daten wurden mittels SPSS 19.0 statistisch ausgewertet. Da die Daten nicht normalverteilt waren (Überprüfung mittels Kolmogorov-Smirnov Anpassungstest), wurden die einzelnen Gruppen gegeneinander mittels des Mann-Whitney-Test auf Unterschiede geprüft. Für alle Tests wurde ein Signifikanzniveau von 5% gewählt. Eine Bonferroni-Korrektur des alpha-Fehlers war bei der Zahl der durchgeführten Vergleiche nicht notwendig.

Ergebnisse

Fragebögen

An den Umfragen beteiligten sich zwischen 48% und 78% der befragten Studierenden. Zwischen 71% und 95% der ausgefüllten Fragebögen waren vollständig und korrekt ausgefüllt und auswertbar.

Die auswertbaren Fragebögen der Testkohorten setzten sich wie folgt zusammen (siehe Tabelle 1 und 2).

Intrafakultärer Vergleich

In einer ersten Auswertung der drei Jahrgänge der Universität Witten/Herdecke (UW/H) und der drei Kohorten der Ruhr-Universität Bochum sollte überprüft werden, ob mittels des DTI signifikante Unterschiede in der Selbsteinschätzung des klinischen Denkens zwischen verschiedenen Jahrgängen eines Curriculums bzw. gleichen Jahrgängen verschiedener Curricula gemessen werden können. Im Idealfall sollten die Fähigkeiten im klinischen Denken mit steigender Semesterzahl zunehmen. Bei den Studierenden der UW/H (siehe Abbildung 3) wird so eine Zunahme im mittels DTI selbst eingeschätzten klinischen Denken mit steigender Semesterzahl deutlich, der Anstieg vom 5. zum 8. Semester ist signifikant und hat eine hohe Effektstärke. Eine Zunahme vom 5. zum 6. Semester mit mittlerer Effektstärke zeigt sich ebenfalls, während sich kein signifikanter Unterschied zwischen dem 6. und dem 8. Semester messen lässt.

Wenn die Veranstaltungen eines Modellstudiengangs das klinische Denken besser fördern als die in einem Regelstudiengang, dann sollten sich Unterschiede zwischen Kohorten nach gleicher Studiendauer in unterschiedlichen Curricula zeigen.

Ein Vergleich der Studierenden im 7. Semester an der RUB konnte jedoch keinen Unterschied zwischen Studierenden im Regel- und solchen im Modellstudiengang nachweisen. Es gab auch keine Zunahme der erreichten Punkte im DTI bei Studierenden im Modellstudiengang vom 7. zum 9. Semester (siehe Abbildung 4).

Interfakultärer Vergleich

In einem weiteren Vergleich wurden die Studierenden der drei beteiligten Modellstudiengänge miteinander verglichen, und zwar jeweils das 9. Semester der Universitäten Köln und Bochum und das 8. Semester der UW/H. Die Ergebnisse im mittels DTI selbst eingeschätzten klinischen Denken der drei untersuchten Kohorten unterschied-

Tabelle 2: Angabe des Cronbachs alpha für alle Fakultäten getrennt sowohl für die Gesamtpunktzahl als auch die beiden Subskalen.

Reliabilität des Fragebogens

Gruppe	Summe aller Fragen	Skala Flexibilität	Skala Struktur
UWH	0,82	0,58	0,75
UzK	0,71	0,49	0,60
RUB	0,83	0,74	0,73

Gesamtpunktzahl UWH

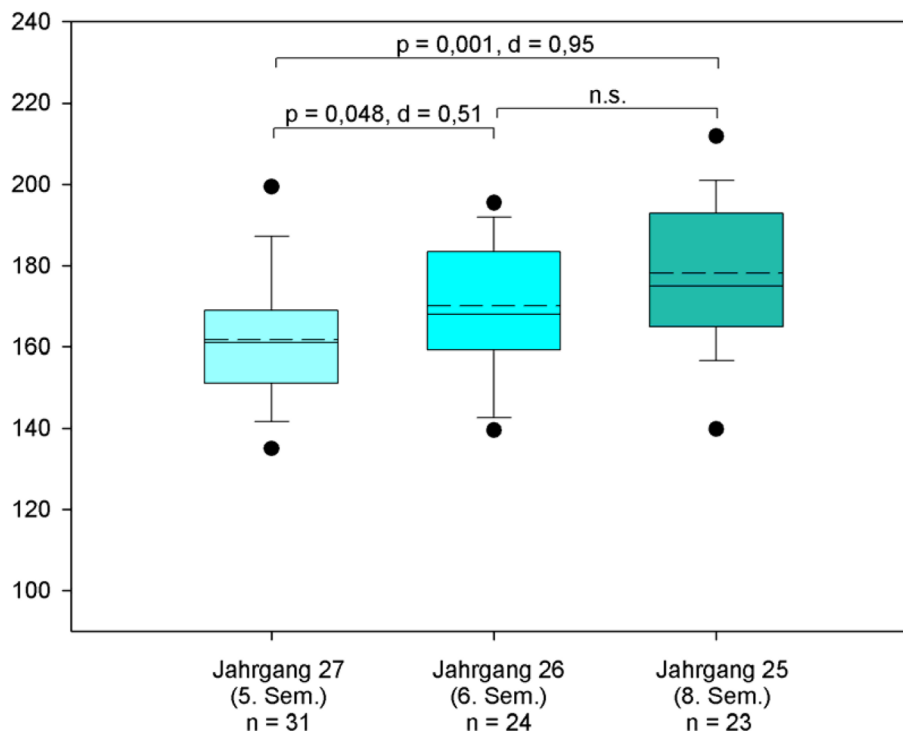


Abbildung 3: Vergleich der Semester 5, 6 und 8 der Universität Witten/Herdecke. Dargestellt sind jeweils die in den Fragebögen erreichte Maximalsumme als Boxplot mit Median (durchgezogene Linie), Mittelwert (gestrichelte Linie), oberem und unterem Quartil (Box), den 5%/95% Perzentilen (Whiskers) und den Minimal- bzw. Maximalwerten (schwarzer Kreis). Oberhalb der Vergleiche sind jeweils der Signifikanzwert (p) und die Effektstärke (d) angegeben.

den sich signifikant voneinander. Die Studierenden der UW/H erreichten höhere Punktzahlen als die Studierenden der anderen beiden Fakultäten, und die Studierenden der RUB erzielten höhere Gesamtpunkte als die der UzK (geringere Effektstärke bei diesem Vergleich), siehe Abbildung 5.

Diskussion

Das Ziel dieses Projekts war es, zu prüfen, ob mit Hilfe des Diagnostic Thinking Inventory (DTI) sowohl Unterschiede in der Selbsteinschätzung des klinischen Denkens von Studierenden verschiedener Fakultäten als auch zwischen Studierenden unterschiedlicher Semester innerhalb einer Fakultät gemessen werden können.

Die interne Konsistenz des Fragebogens für den Gesamtfragebogen war befriedigend bis gut, die der Subskalen

dagegen nur gering. Daher wurde in allen Gruppen jeweils nur die erreichte Gesamtpunktzahl verglichen.

Es ließen sich signifikante Unterschiede sowohl zwischen einzelnen Jahrgängen einer Universität (zunehmende Punktzahl im selbst eingeschätzten klinischen Denken bei Studierenden der UW/H) als auch zwischen vergleichbaren Studiensemestern einzelner Universitäten messen. Für die Studierenden der UW/H ergeben sich damit Ergebnisse, die der Eingangs-Hypothese entsprechen, dass mit zunehmendem Studienfortschritt das klinische Denken zunimmt. Beide Elemente, die laut Literatur dazu beitragen können (PbL [24], [25] und klinische Erfahrung in Form von Blockpraktika oder Famulaturen [27], [28]), sind im Curriculum der UW/H stark vertreten. Wenn PbL allein eine Auswirkung auf das klinische Denken hätte, würde man allerdings auch erwarten, dass die Studierenden im Modellstudiengang der RUB höhere Werte erzielen

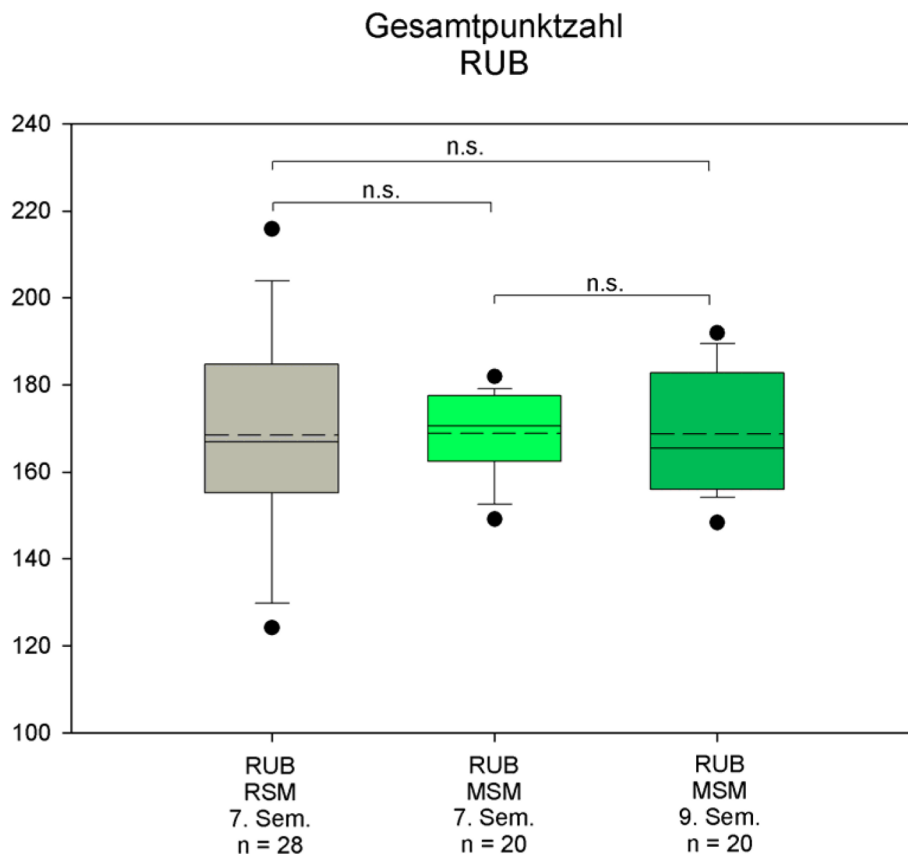


Abbildung 4: Vergleich der Semester 7 des Regelstudiengangs Medizin (RSM) sowie die Semester 7 und 9 des Modellstudiengangs Medizin (MSM) der Ruhr Universität Bochum. Dargestellt sind jeweils die in den Fragebögen erreichte Maximalsumme als Boxplot mit Median (durchgezogene Linie), Mittelwert (gestrichelte Linie), oberem und unterem Quartil (Box), den 5%/95% Perzentilen (Whiskers) und den Minimal- bzw. Maximalwerten (schwarzer Kreis). Oberhalb der Vergleiche ist jeweils der Signifikanzwert (p) angegeben.

als ihre Kommilitonen aus dem Regelstudiengang. Dies ist in dieser Studie nicht der Fall.

Wenn man die am weitesten fortgeschrittenen Jahrgänge aller untersuchten Fakultäten vergleicht, so zeigt sich auch hier, dass diejenigen aus der Fakultät mit den größten Anteilen an PbL und klinischen Praktika die höchsten Punkte erreichen, diejenigen, aus Curricula mit geringeren Anteilen an diesen Elementen, erzielen die schlechteren Ergebnisse.

Um die gemessenen Daten nicht nur innerhalb der untersuchten Gruppen zu verwenden und die erreichten Punktzahlen auf Plausibilität zu überprüfen, wurden sie mit den von Bordage [29] gemessenen Werten verglichen. Die niedrigste in dieser Studie gemessene Gesamtpunktzahl einer Gruppe (Studierenden der UzK, 9. Semester mit 150,5 Punkten) liegen unterhalb der von Bordage gemessenen Werte für Studierende im 3. Studienjahr (158,3 Punkt), während die Gruppe mit der höchsten erreichten Punktzahl aus unserer Studie (UW/H, 8. Semester mit 178,2 Punkten) oberhalb der für Assistenzärzte in der Inneren Medizin (Senior House Officers: 168,4 Punkte) und Hausärzte (General Practitioners 172,3 Punkte) gemessenen Werte liegt. Dies legt nahe, dass sich der DTI besonders für innerfakultäre Vergleiche eignet, da eine Kalibrierung zwischen verschiedenen Systemen schwierig ist. In der Studie von Bordage war von den Studierenden im ersten Studienjahr (153,9 Punkte) bis

hin zu Registrars (180,2 Punkte) zwar ein kontinuierlicher Zuwachs an Punkten zu verzeichnen, die Unterschiede in unserer Studie sind zwischen den gemessenen Gruppen jedoch größer.

Der Fragebogen wurde bei freiwilligen Teilnehmern in nicht-standardisierter Umgebung aus drei verschiedenen medizinischen Fakultäten in Nordrhein-Westfalen eingesetzt. Dies bedeutet, dass die Ergebnisse nicht zwingend für die jeweiligen Kohorten repräsentativ sind. Besonders bei den für den Regelstudiengang (7. Semester, RUB) erhobenen Daten ist die Streuung der Werte sehr groß, und die befragten Teilnehmer (40) sind nur eine kleine, zufällig ausgesuchte Gruppe der Gesamtkohorte. Daher ist es fraglich, wie repräsentativ die Daten dieser Stichprobe für ihren Jahrgang sind.

In Köln und an der UW/H wurden die Testbögen zudem im Anschluss an den Progress-Test ausgegeben, was zum Teil für die recht geringe Rücklaufquote, die niedrige Reliabilität und die niedrigen Gesamtergebnisse verantwortlich sein kann. Die Reliabilität und die Gesamtpunktzahlen sind an der UW/H höher. Hier kann sich motivationssteigernd ausgewirkt haben, dass der soziale Druck teilzunehmen, in der kleinen Kohorte größer ist. Auch der Anteil der nicht auswertbaren Fragebögen ist an der UW/H geringer. Die Bögen der UzK konnten etwa in einem Drittel der Fälle nicht ausgewertet werden; daher muss der Vergleich mit den Werten der UzK kritisch betrachtet

Vergleich UWH : UzK : RUB Gesamtpunktzahl Semester 8 bzw. 9

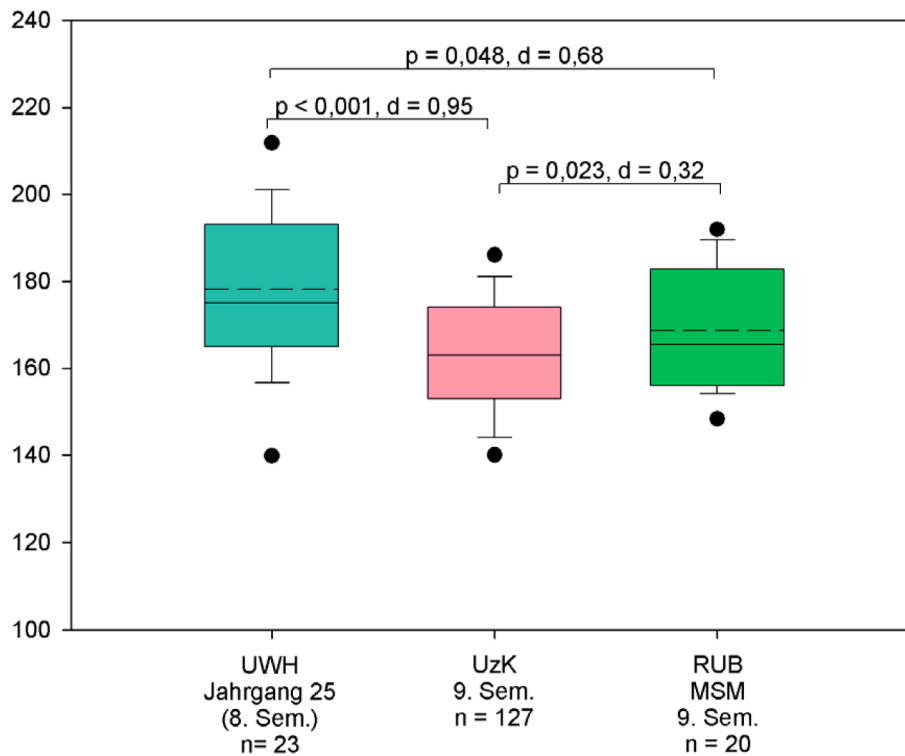


Abbildung 5: Vergleich der Semester 8 bzw. 9 der Universitäten Witten/Herdecke (UWH), Universität zu Köln (UzK) und Ruhr-Universität Bochum (RUB). Dargestellt sind jeweils die in den Fragebögen erreichte Maximalsumme als Boxplot mit Median (durchgezogene Linie), Mittelwert (gestrichelte Linie), oberem und unterem Quartil (Box), den 5%/95% Perzentilen (Whiskers) und den Minimal- bzw. Maximalwerten (schwarzer Kreis).

Oberhalb der Vergleiche sind jeweils der Signifikanzwert (p) und die Effektstärke (d) angegeben.

werden. Die Fragebögen, die nicht in allen Fragen eine eindeutige Antwort aufwiesen, wurden in der Auswertung nicht berücksichtigt. In vielen der Fragebögen war trotz der Aufforderung, immer ein Kästchen zwischen den Aussagen anzukreuzen, die Trennlinie zwischen zwei Kästchen gewählt worden, da sich die Studierenden offensichtlich nicht für einen Wert der Skala entscheiden wollten. Dies war auch auf den Bögen zu beobachten, die von den Studierenden der RUB ausgefüllt wurden und ist der Grund für den geringen Anteil auswertbarer Fragebögen. Des Weiteren muss man bei der Interpretation der Daten berücksichtigen, dass es sich beim DTI zwar um ein etabliertes Instrument handelt, dass es allerdings nicht objektiv das klinische Denken sondern lediglich die Selbsteinschätzung der Studierenden in dieser Disziplin misst. Eine (durchaus erwünschte) insgesamt kritische Haltung im Hinblick auf die eigenen Kompetenzen könnte nämlich die Selbsteinschätzung zu klinischem Denken negativ beeinflussen. Auch in dem Maß der Selbstkritik können sich die untersuchten Kohorten durchaus unterscheiden. Hierzu wurde kein weiteres Instrument eingesetzt, das zur Standardisierung in diesem Punkt hätte dienen können. Man kann beispielsweise davon ausgehen, dass sich männliche Studierende eher in ihren

Kompetenzen überschätzen [38]. In der Verteilung männlich/weiblich unterscheiden sich die untersuchten Gruppen jedoch nicht signifikant voneinander. Daher sollten die in dieser Studie gefundenen Unterschiede nicht durch eine geschlechts-spezifische Überschätzung der Studierenden der UW/H und eine Unterschätzung derer in Köln bedingt sein.

Zusammenfassung

Zusammenfassend kann man aus dieser Pilotstudie schließen, dass der DTI als Instrument durchaus geeignet ist, das selbsteingeschätzte klinische Denken bei Studierenden unterschiedlicher Curricula vergleichend zu erheben.

Die Studie hat aber auch gezeigt, dass weitere Schritte unternommen werden müssen, wenn die gemessenen Unterschiede auf die Verteilung unterschiedlicher curricularer Elemente wie z.B. PbL oder klinisch-praktische Erfahrung zurückgeführt werden sollen.

Grundsätzlich muss gewährleistet werden, dass die Befragung unter vergleichbaren Bedingungen stattfindet, wobei eine höhere Rücklaufquote erreicht werden muss.

Zusätzlich erscheint es notwendig, ein weiteres Inventar einzusetzen, um zu überprüfen, ob die Studierenden zu Selbstüberschätzung/-unterschätzung neigen.

Der DTI selbst sollte durch weitere fachbezogene objektive Tests zum klinischen Denken und Verhalten weitergehend validiert werden [39], z.B. im Sinne eines Key Feature [40], Script Concordance [41] oder Situational Judgment Tests [42].

Interessenkonflikt

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

Literatur

- Barrows HS, Feltovich PJ. The clinical reasoning process. *Med Educ.* 1987;21(2):86-91. DOI: 1365-2923.1987.tb00671.x
- Pelaccia T, Tardif J, Tribi E, Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Med Educ Online.* 2011;14:16. DOI: 10.3402/meo.v16i0.5890
- Norman G. Research in clinical reasoning: past history and current trends. *Med Educ.* 2005;39(4):418-427. DOI: 10.1111/j.1365-2929.2005.02127.x
- Charlin B, Tardif J, Boshuizen HPA. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med.* 2000;75(2):182-190. DOI: 10.1097/00001888-200002000-00020
- Schmidt HG, Norman GR, Boshuizen HP. A Cognitive Perspective on Medical Expertise: Theory and Implications. *Acad Med.* 1990;65(10):611-621. DOI: 10.1097/00001888-199010000-00001
- Schmidt HG, Boshuizen HP. On acquiring expertise in medicine. *Educ Psychol Rev.* 1993;5(3):205-221. DOI: 10.1007/BF01323044
- Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. *Med Teach.* 2003; 25(3):308-313. DOI: 10.1080/0142159031000100427
- Grant J, Marsden P. The structure of memorized knowledge in students and clinicians: an explanation for diagnostic expertise. *Med Educ.* 1987;21(2):92-98. DOI: 10.1111/j.1365-2923.1987.tb00672.x
- Eva KW, Neville AJ, Norman GR. Exploring the Etiology of Content Specificity: Factors Influencing Analogic Transfer and Problem Solving. *Acad Med.* 1998;73(10):S1-S5. DOI: 10.1097/00001888-199810000-00028
- Dory C, Gagnon R, Charlin B. Is case-specificity content-specificity? An analysis of data from extended-matching questions. *Adv Health Sci Educ.* 2010;15(1):55-63. DOI: 10.1007/s10459-009-9169-z
- Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ.* 2006;40(7):618-623. DOI: 10.1111/j.1365-2929.2006.02511.x
- Kassirer JP. Teaching clinical reasoning: Case-based and coached. *Acad Med.* 2010;85(7):1118-1124. DOI: 10.1097/ACM.0b013e3181d5dd0d
- Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ.* 2004;39(1):98-106. DOI: 10.1111/j.1365-2929.2004.01972.x
- Wolf FM, Gruppen LD, Billi JE. Use of the competing-hypotheses heuristic to reduce 'pseudodiagnosticity'. *Acad Med.* 1988;63(7):548-554. DOI: 10.1097/00001888-198807000-00006
- Kopp V, Möltner A, Fischer MR. Key-feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. *GMS Z Med Ausbild.* 2006;23(3):Doc50. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2006-23/zma000269.shtml>
- Kopp V, Stark R, Fischer MR. Förderung von Diagnose-Kompetenz in der medizinischen Ausbildung durch Implementation eines Ansatzes zum fallbasierten Lernen aus Lösungsbeispielen. *GMS Z Med Ausbild.* 2007;24(2):Doc107. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2007-24/zma000401.shtml>
- Gräsel C, Mandl H. Förderung des Erwerbs diagnostischer Strategien in fallbasierten Lernumgebungen. *Unterrichtswissenschaft.* 1993;21:355-370.
- Maudsley G, Strivers J. Promoting professional knowledge, experiential learning and critical thinking for medical students. *Med Educ.* 2000;34(7):535-544. DOI: 10.1046/j.1365-2923.2000.00632.x
- Kamin CS, O'Sullivan PS, Younger M, Deterding R. Measuring Critical Thinking in Problem-Based Learning Discourse. *Teach Learn Med.* 2001;13(1):27-35. DOI: 10.1207/S15328015TLM1301_6
- Simpson E, Courtney M. Critical thinking in nursing education: Literature review. *Int J Nurs Pract.* 2002;8(2):89-98. DOI: 10.1046/j.1440-172x.2002.00340.x
- Eshach H, Bittermann H. From case-based reasoning to problem-based learning. *Acad Med.* 2003;78(5):491-496. DOI: 10.1097/00001888-200305000-00011
- Hmelo CE. Cognitive Consequences of Problem-Based Learning for the Early Development of Medical Expertise. *Teach Learn Med.* 1998;10(2):92-100. DOI: 10.1207/S15328015TLM1002_7
- Rochmawati E, Wiechula R. Education strategies to foster health professional students' clinical reasoning skills. *Nurs Health Sci.* 2010;12(2):244-250. DOI: 10.1111/j.1442-2018.2009.00512.x
- Tiwari A, Leighton-Beck L, So M, Yuen K. A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Med Educ.* 2006;40(6):547-554. DOI: 10.1111/j.1365-2929.2006.02481.x
- Yuan H, Kunaviktikul W, Klunkklin A, Williams BA. Improvement of nursing students' critical thinking skills through problem-based learning in the People's Republic of China: A quasi-experimental study. *Nurs Health Sci.* 2008; 10(1):70-76. DOI: 10.1111/j.1442-2018.2007.00373.x
- Windish DM. An Innovative Curriculum Teaching the Integration of Communication and Clinical Reasoning Skills to Medical Students. Liverpool: MPH Capstone Project; 2004.
- Bowen JL. Educational strategies to promote diagnostic clinical reasoning. *N Engl J Med.* 2006;355(21):2217-2225. DOI: 10.1056/NEJMra054782
- Schmidmaier R, Eiber S, Ebersbach R, Schiller M, Hege I, Holzer M, Fischer MR. Learning the facts in medical school is not enough: which factors predict successful application of procedural knowledge in a laboratory setting? *BMC Med Educ.* 2013;13:28. DOI: 10.1186/1472-6920-13-28
- Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. *Med Educ.* 1990;24(5):413-425. DOI: 10.1111/j.1365-2923.1990.tb02650.x

30. Jones UF. The reliability and validity of the Bordage, Grant & Marsden diagnostic thinking inventory for use with physiotherapists. *Med Teach.* 1997;19(2):133-140. DOI: 10.3109/01421599709019366
31. Sobral DT. Medical Students' Mindset for Reflective Learning: A Revalidation Study of the Reflection-In-Learning Scale. *Adv Health Sci Educ.* 2005;10(4):303-314. DOI: 10.1007/s10459-005-8239-0
32. Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Med Teach.* 2002;24(5):507-515. DOI: 10.1080/01421590220145743
33. Round AP. Teaching clinical reasoning – a preliminary controlled study. *Med Educ.* 1999;33(7):480-483. DOI: 10.1046/j.1365-2923.1999.00352.x
34. Beullens J, Struyf E, van Damme B. Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Med Educ.* 2006;40(12):1173-1179. DOI: 10.1111/j.1365-2929.2006.02627.x
35. Jerant AF, Azari R. Validity of Scores Generated by a Web-Based Multimedia Simulated Patient Case Software: A Pilot Study. *Acad Med.* 2004;79(8):805-811. DOI: 10.1097/00001888-200408000-00017
36. Stieger S, Praschinger A, Kletter K, Kainberger F. Diagnostic grand rounds: A new teaching concept to train diagnostic reasoning. *Eur J Radiol.* 2009;78(3):349-352. DOI: 10.1016/j.ejrad.2009.05.015
37. Lee A, Joynt GM, Lee AKT, Ho A, Groves M, Vlantis AC, Ma RC, Fung CS, Aun CS. Using illness scripts to teach clinical reasoning skills to medical students. *Fam Med.* 2010;42(4):255-261.
38. Jünger J, Schellberg D, Nikendei C. Subjektive Kompetenzeinschätzung von Studierenden und ihre Leistung im OSCE. *GMS Z Med Ausbild.* 2006;23(3):Doc51. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2006-23/zma000270.shtml>
39. Ilgen JS, Humbert AJ, Kuhn G, Hansen ML, Norman GR, Eva KW, Charlin B, Sherbino J. Assessing diagnostic reasoning: a consensus statement summarizing theory, practice, and future needs. *Acad Emerg Med.* 2012;19(12):1454-1461. DOI: 10.1111/acem.12034
40. Page G, Bordage G. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Acad Med.* 1995;70(2):104-110. DOI: 10.1097/00001888-199502000-00012
41. Lubarsky S, Charlin B, Coe DA, Chalk C, van der Vleuten CP. Script concordance testing: a review of published validity evidence. *Med Educ.* 2011;45(4):329-338. DOI: 10.1111/j.1365-2923.2010.03863.x
42. Rahman M. Tackling situational judgment tests. *BMJ Careers.* 2007;334:gp189-gp190.

Korrespondenzadresse:

Dr. Kirsten Gehlhar, MME

Carl von Ossietzky Universität Oldenburg, Fakultät für Medizin und Gesundheitswissenschaften, Carl von Ossietzky Str. 9-11, 26129 Oldenburg, Deutschland
kirsten.gehlhar@uni-oldenburg.de

Bitte zitieren als

Gehlhar K, Klimke-Jung K, Stosch C, Fischer MR. Wirken sich unterschiedliche medizinische Curricula auf das selbst eingeschätzte klinische Denken von Studierenden aus? *GMS Z Med Ausbild.* 2014;31(2):Doc23.
 DOI: 10.3205/zma000915, URN: urn:nbn:de:0183-zma0009152

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2014-31/zma000915.shtml>

Eingereicht: 02.07.2013**Überarbeitet:** 13.03.2014**Angenommen:** 02.04.2014**Veröffentlicht:** 15.05.2014**Copyright**

©2014 Gehlhar 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.

Do different medical curricula influence self-assessed clinical thinking of students?

Abstract

Objectives: As a fundamental element of medical practice, clinical reasoning should be cultivated in courses of study in human medicine. To date, however, no conclusive evidence has been offered as to what forms of teaching and learning are most effective in achieving this goal. The Diagnostic Thinking Inventory (DTI) was developed as a means of measuring knowledge-unrelated components of clinical reasoning. The present pilot study examines the adequacy of this instrument in measuring differences in the clinical reasoning of students in varying stages of education in three curricula of medical studies.

Methods: The Diagnostic Thinking Inventory (DTI) comprises 41 items in two subscales (“Flexibility in Thinking” and “Structure of Knowledge in Memory”). Each item contains a statement or finding concerning clinical reasoning in the form of a stem under which a 6-point scale presents opposing conclusions. The subjects are asked to assess their clinical thinking within this range. The German-language version of the DTI was completed by 247 student volunteers from three schools and varying clinical semesters. In a quasi-experimental design, 219 subjects from traditional and model courses of study in the German state of North Rhine-Westphalia took part. Specifically, these were 5th, 6th and 8th semester students from the model course of study at Witten/Herdecke University (W/HU), from the model (7th and 9th semester) and traditional (7th semester) courses of study at the Ruhr University Bochum (RUB) and from the model course of study (9th semester) at the University of Cologne (UoC). The data retrieved were quantitatively assessed.

Results: The reliability of the questionnaire in its entirety was good (Cronbach’s alpha between 0.71 and 0.83); the reliability of the subscales ranged between 0.49 and 0.75. The different groups were compared using the Mann-Whitney test, revealing significant differences among semester cohorts within a school as well as between students from similar academic years in different schools. Among the participants from the model course of study at the W/HU, scores increased from the 5th to the 6th semester and from the 5th to the 9th semester. Among individual cohorts at RUB, no differences could be established between model and traditional courses of study or between 7th and 9th semester students in model courses of study. Comparing all participating highest semester students, the 8th semester participants from the W/HU achieved the highest scores – significantly higher than those of 9th semester RUB students or 9th semester UoC students. Scores from the RUB 9th semester participants were significantly higher than those of the 9th semester UoC participants.

Discussion: The German-language version of the DTI measures self-assessed differences in diagnostic reasoning among students from various semesters and different model and traditional courses of study with satisfactory reliability. The results can be used for discussion in the context of diverse curricula. The DTI is therefore appropriate for further research that can then be correlated with the different teaching method characteristics and outcomes of various curricula.

Keywords: clinical thinking, clinical reasoning, PBL, diagnostic thinking inventory

Kirsten Gehlhar¹

Kathrin Klimke-Jung²

Christoph Stosch³

Martin R. Fischer⁴

1 Carl von Ossietzky University of Oldenburg, School of Medicine and Health Sciences, Oldenburg, Germany

2 Ruhr University Bochum, Bochum, Germany

3 University of Cologne, Cologne, Germany

4 Clinic of the Ludwig Maximilian University of Munich, Institute for Medical Education, Munich, Germany

Introduction

Clinical thinking is a central component of physician competence. Optimal patient care depends on thorough analysis of the information provided by the patient and on the risk-benefit assessment of diagnostic tests and therapies.

It follows that every university-level programme of medical education must have the objective of forming good clinical reasoning skills in its students. Conclusive evidence of the particular advantages of specific types of teaching in fostering these skills, however, has yet to be provided.

Clinical problems present essential differences in comparison to well-structured tasks. Necessary information is not available in its entirety at the beginning of the process, and the nature of a problem can change its dynamics during the diagnostic-therapeutic process (case history and examination). The task is a matter of complex problem solving because there are no standardised procedures for arriving at a solution; instead, each problem is unique, and the physician can never be fully certain that the solution found is actually correct [1], [2].

Models of clinical reasoning

This subject has been intensively worked on within the field of educational research since the 1970s. The objective is to better qualify which processes are at work during clinical reasoning and to identify differences between experts and novices. From this information, it should be possible to draw conclusions that will aid better instruction in clinical reasoning.

Over the course of this time, many concepts explaining the acquisition of expertise in clinical reasoning alternated and were further developed [3]. One concept, for example, is that of so-called "illness scripts" [4]. These are representations of problems (diseases, for example), syndromes or groups of diseases along with the conditions in which they occur, their manifestations, diagnoses and therapy concepts, as well as their pathophysiological bases.

Other models emphasise the particular benefit of practical experience in achieving medical expertise [5], [6]. Empirical evidence shows that skills associated with clinical problem solving were more developed among experts than among novices [7]. It was shown, for example, that the expert's knowledge in memory is primarily linked together by probabilities [4], or that it is characterised by reference models which are spontaneously recognised and then confirmed with hypothetico-deductive methods [2]. Furthermore, experts are better at using so-called "semantic qualifiers" that classify a symptom on a bipolar scale or in the quality of key features recognised in a patient history [8].

Clinical reasoning, however, is also dependent on content and context [9], and expertise in one area does not necessarily mean that comparable skills are available in a different specialised discipline or medical case. In fact, it is not even possible to generalise expertise within a field [10], [11].

None of these approaches are seen as mutually exclusive today [12]. As a result, experts make parallel use of analytical (deductive, controlled) and non-analytical (unconscious, spontaneous) processes in resolving patient cases. This is because effective clinical problem solving is based on clear procedures of information gathering, hypothesis formation and hypothesis testing on the one hand, while a diagnosis depends on knowledge of the underlying mechanisms and the problem-related integration and mental organisation of this information on the other. Recognition of patterns accelerates recollection [13]. The two processes complement each other and a bidirectional exchange of information takes place.

Teaching clinical reasoning

The question of how to optimally teach students such complex abilities has, for good reason, yet to be definitively answered. Students who had been instructed in the heuristic method of concurrent hypotheses (Bayes' Theorem), for example, were able to implement the method satisfactorily following the teaching unit but later failed in the clinic when it came to transferring this abstract model into practice [14]. Other studies showed that work on didactically chosen patient cases with errors and elaborated feedback or on case studies with instruction had a positive effect on clinical reasoning in the respective trial setting [15], [16], [17].

To date, two teaching and learning environments have almost unanimously proved effective in facilitating clinical reasoning: problem-based learning (PBL) and clinical praxis/clinical experience [18], [19], [20].

Because it is particularly well-suited to learning and practicing the problem solving process [21], PBL should contribute to students' mastering of clinical reasoning [18] [20]. PBL combines content with context, and the application of memorised knowledge to clinical problems fosters the development of coherent, pathophysiological concepts [21], [22]. Meta-analysis [23] of the question of which types of instruction promote clinical reasoning revealed two studies showing that students improved their critical thinking with PBL [24], [25] and that they were able to deliver more accurate, more coherent and more comprehensive explanations for medical problems than other students [22]. Critical thinking and clinical reasoning are not necessarily comparable, however.

Concrete clinical experience with patients is a further indispensable factor in the development of clinical reasoning. Firstly, students must also practice their acquired abilities [26], and secondly, clinical reasoning is strongly influenced by experience [12] and is a consequence of resulting multidimensional knowledge. It was shown that the students' stages of development often corresponded to their clinical experience, for example, rather than to their academic year [27], [28].

On the basis of these findings, traditional curricula, in which instruction is limited to the subject at hand, little PBL is implemented and clinical experience comes later

and often to a lesser extent, presumably offers students less opportunity to learn and practice clinical reasoning.

Measuring clinical reasoning

In 1990, in order to measure clinical reasoning by means of self-assessment, Bordage and Marsden developed a corresponding inventory [29] - the Diagnostic Thinking Inventory (DTI). Until then, inventories were used that measured critical instead of clinical reasoning, for instance. Subjects were asked to outline their thinking processes either orally or in written form, or the solving of a concrete clinical problem was tested. The DTI, on the other hand, is an instrument that measures both self-evaluated flexibility in thinking and structure of knowledge in memory independently of context and can also differentiate between medical experts of varying degrees of training. It is comprised of 41 items in which subjects use a scale to evaluate themselves in predefined situations. The answers, in turn, represent a specific degree of clinical reasoning. The original research was conducted among 270 test subjects with different degrees of training - from first-semester students to experienced physicians. Significant differences between students and physicians were revealed, whereas the differences between physicians of varying degrees of experience were insignificant. The DTI inventory was implemented and validated in various studies in the following years [30] and was also analysed in conjunction with other cognitive or psychometric tests [31].

In total, the present study results reached with the DTI are not consistent. In the majority of studies, the students' DTI results improved significantly with instruction in clinical problem solving, instruction in diagnostic procedure errors, through the handling of patient cases, through participation in diagnostic case discussions or with increasing semesters of study [32], [33], [34], [35], [36]. Two other studies, however, showed that the conducted interventions did not lead to any improvements in the DTI [26], [37].

The inventory measures the self-assessment of the participants on the type and structure of their clinical reasoning, but it does not measure their actual diagnostic ability. For this reason, the correlation between DTI results and resolved cases or determined diagnoses is often weak. 6,32 In addition to clinical reasoning skills, a corresponding knowledge base in the respective specialist field and clinical experience are required to arrive at the correct diagnosis.

Objectives

The medical curricula of different schools vary significantly in their objectives and focal points. Despite their differing paths, their common goal is an education that will produce good physicians. In the long term therefore it is necessary to provide evidence as to the effect that different curricula have on acquiring clinical reasoning as a core competence of physicians. As a first step, this pilot study,

conducted in three different medical schools in North Rhine-Westphalia, examined the measurability of differences between students of varying curricula, on the one hand, and between differing stages of study, on the other. Answers to the following questions were sought:

1. Can an increase in the students' self-assessed clinical reasoning competence over the course of their studies be evidenced using DTI?
2. Are there differences regarding clinical-reasoning-related self-assessment between students from varying courses of study and curricula?

Presumably, clinical reasoning competence would increase with the advance of study (i.e., advancing academic semesters). Differences between students of varying curricula would be apparent when, for example, the amount of elements such as PBL or clinical training differed. Students at W/HU and in the model course of study at RUB undergo model curricula that are characterised in part by problem-oriented learning (PBL) from their start and - particularly at W/HU - by long and numerous clinical traineeships. PBL is the central element of the first four semesters at W/HU. In addition, the amount of clinical experience prescribed by the curricula of W/HU is the highest among all of the courses of study. Starting from the second half of the fourth semester until the practical year at the end of their studies, 46 weeks are devoted to blocks of practical traineeships, while six weeks are spent in on-site observation in general medicine. The two courses of study at RUB differ in regard to their curriculum. The traditional course of study has a subject-related traditional curriculum with a six-week PBL block in the fourth semester. The model course of study is theme-based and praxis-related. TBL is a structuring element in the first four semesters and is offered concurrently with 5th, 8th and 9th semester studies. The two courses of study also differ in numbers of students. While 42 candidates are accepted to the model course of study programme per academic year, the traditional course of study admits 200 students per year. The survey included all students present in the 7th and 9th semesters of the model course of study as well as two seminar groups from the general medicine course in the 7th semester of the traditional course of study.

Methods

Participants, survey instrument

In order to measure the self-assessment of memory-unrelated components of clinical reasoning, the DTI (Diagnostic Thinking Inventory) was implemented. To this end, the German-language translation of the questionnaire (courtesy of Dr. Stieger [36]) was administered in the model courses of study at three schools in North Rhine-Westphalia (Bochum, Cologne and Witten/Herdecke) and in one traditional course of study (Bochum) in machine-readable form.

In order to answer the question of whether clinical reasoning improved with the duration of study, different semesters within a course were included. This was the case for the model course at W/HU and the model course at RUB. As a measurement of differences in self-assessed clinical reasoning between students of varying courses and curricula, advanced students from all three locations were surveyed prior to their practical (final) year. At the time of our study, the 8th semester was the most advanced pre-practical-year semester available at W/HU because these students had enrolled in a period when registration was only possible for the summer semester (beginning in spring at German universities). It was possible to include 9th semester students from RUB and UoC, as their studies had commenced in the winter semester (beginning in autumn at German universities).

The questionnaires were distributed for completion and immediately collected again at the participating universities in a period between October and December 2010 at a specific time, in each case, when face-to-face instruction was taking place, at a point of greatest possible attendance of that semester's students.

The students were briefed on the purpose of the project directly preceding the distribution of questionnaires. Participation was voluntary and the survey was anonymous; separate declarations of consent were not given.

Following evaluation, students' scores and respective commentary were made available to them upon request and presentation of the correlating questionnaire number. The results were then made available under that number's reference.

Instrument

The DTI questionnaire measures the self-assessment of non-knowledge-related components of clinical reasoning and comprises 41 questions (see sample questions, table 1). Twenty of these can be allocated to the subscale "flexibility in thinking" and 21 to the subscale "structure of memory" (see examples below). "Flexibility in thinking" illustrates the participants' ability during the diagnostic process to arrive at the correct diagnosis and to flexibly incorporate new information. "Structure in memory" reflects the degree of organisation and accessibility of the memorised knowledge from which the participants' draw. Every question is composed of a stem (usually a statement) and a response scale. At each end, the response scale offers two opposing answers/statements in response to the initial statement, with six boxes between to select from. Participants are asked to check the box in the scale that best reflects their position between the response options. The options are randomly placed either at the left or the right end of the scale which means that the response reflecting more advanced clinical reasoning could be at either end.

Sample question from the flexibility subscale (see figure 1).

When I am interviewing a patient,

I often seem to have a fixed idea in my mind about what might be wrong.

I usually find it easy to explore various possible diagnoses.

--	--	--	--	--	--	--

Figure 1

Sample question from the structure subscale (see figure 2).

When I am collecting information about a patient,

The individual units of information usually seem to group themselves in my mind.

I am often unable to see how the units of information relate to each other.

--	--	--	--	--	--	--

Figure 2

Evaluation

All of the completed questionnaires were scanned and were then read using analysis software (FormPro 2.5). Only questionnaires in which all questions had been clearly answered were included. Score results were calculated by assigning values of 1 to 6 to the selection boxes. The response showing the most pronounced clinical reasoning was given the highest score. The total score achieved (a maximum of 126 points possible) as well as the points in the subscales "flexibility" (max. 120 points) and "structure" (max. 126 points) were calculated for each participant. Cronbach's alpha was used in determining the consistency of the DTI. Effect size was calculated as Cohen's *d*.

All data were statistically analysed using SPSS 19.0. Since the data were not normally distributed (checked with the Kolmogorov-Smirnov goodness-of-fit test), the individual groups were checked for differences using the Mann-Whitney test. A significance level of 5% was chosen. Considering the number of comparisons performed, a Bonferroni correction of alpha errors was not necessary.

Results

Questionnaires

Between 48% and 78% of the students solicited took part in the survey. Between 71% and 95% of the returned questionnaires were complete and evaluable.

The test cohorts' evaluable questionnaires were divided as follows (see table 1 and 2).

Intra-faculty comparison

An initial evaluation of the three represented academic years from Witten/Herdecke University and the three co-

Table 1: number and allocation of questionnaires

University	Type of studies	Semester	Number of questionnaires distributed	Number of completed questionnaires	Number of evaluated questionnaires	Distribution female/male [%]
Witten/Herdecke	Model	5	42	33	31	58.6 / 41.4
Witten/Herdecke	Model	6	42	25	24	66.6 / 33.4
Witten/Herdecke	Model	8	42	25	23	56.5 / 43.3
Ruhr University	Model	7	30	28	20	50.0 / 50.0
Ruhr University	Traditional	7	40	31	20	53.5 / 46.5
Ruhr University	Model	9	37	36	29	68.0 / 32.0
University of Cologne	Model	9	150	141	101	58.5 / 41.5

Table 2: Cronbach's alpha for all schools, listed by total score as well as by subscale scores.
Questionnaire reliability

Group	Sum total	Flexibility scale	Structure scale
UWH	0.82	0.58	0.75
UoC	0.71	0.49	0.60
RUB	0.83	0.74	0.73

horts from Ruhr University Bochum was targeted at revealing the value of DTI in measuring significant differences in the self-assessment of clinical reasoning between varying academic years of one curriculum or, respectively, differences between academic years of different curricula.

Ideally, abilities in clinical reasoning would increase with the number of semesters attended. Among students at W/HU (see figure 3), a distinct corresponding semester-related increase in self-assessed clinical reasoning was evidenced using DTI; the increase from the 5th to the 8th semester is significant and has a large effect size. An increase from the 5th to the 6th semester, of medium effect size, is also evidenced, while no significant difference could be detected between the 6th and 8th semesters.

If model courses of study compared are more conducive to the facilitation of clinical reasoning in comparison with traditional courses, then differences between cohorts with differing curricula but the same length of study should be detectable. In a comparison of 7th semester RUB students, however, no difference could be evidenced between those from traditional studies and those from model studies (see figure 4).

Inter-faculty comparison

In a further study, students from the three participating model courses were compared; namely 9th semester students from both the University of Cologne and Ruhr University Bochum, and 8th semester students from Witten/Herdecke University. The three cohorts' results from the self-assessment of clinical reasoning using DTI differed significantly. The students from W/HU achieved higher scores than the students from both of the other schools, and the students from RUB achieved higher overall scores than the students from UoC (lesser effect size in this comparison) – see figure 5.

Discussion

The objective of the project was to examine the effectiveness of measuring differences in self-assessment with the aid of the Diagnostic Thinking Inventory not only between students from different faculties but also between students in different semesters within one faculty.

The internal consistency of the questionnaire was acceptable to good; that of the subscales, however, was meagre.

Overall score: University of Witten/Herdecke

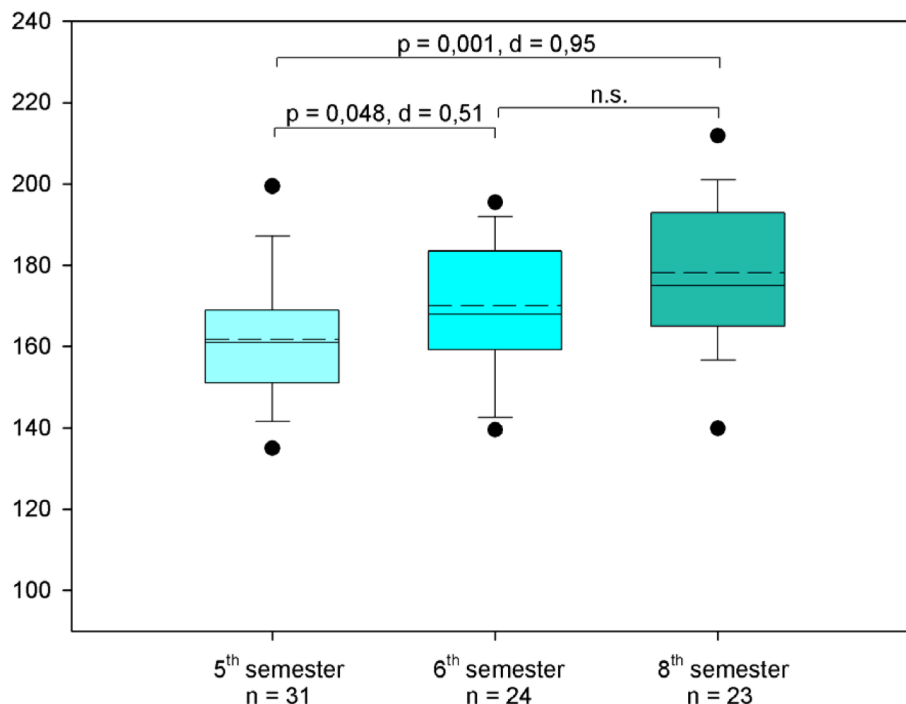


Figure 3: Overall Score W/HU - Comparison of 5th, 6th and 8th semesters at Witten/Herdecke University (German abbr. UW/H). The respective maximum scores achieved in the questionnaires are depicted in a boxplot with median (solid line), mean score (dashed line), upper and lower quartile (box), the 5%/95% percentiles (whiskers) and minimum or maximum score (black circle). The respective p-value (p) and effect size (d) are listed above the comparisons.

As a result, only the respective total scores were compared in all groups.

Significant differences could be measured both between individual academic years at a university (increasing scores in the self-assessed clinical reasoning of students from Witten/Herdecke University) as well as between comparable semesters at different universities. The results from students at W/HU corresponded to the initial hypothesis that clinical reasoning skills increase with the advance of course studies. Both of the elements that, according to references, can aid this development (PBL [24], [25] and clinical experience in the form of blocks of practical training or clinical traineeships [27], [28]), are soundly represented in the curriculum of W/HU. If PBL alone had an effect on clinical reasoning, then higher scores would be expected from students in the model courses of study at RUB than from their fellow students from traditional courses. This, however, was not evidenced in this study.

A comparison of all students from the most advanced academic years of the schools surveyed also shows that courses of study with the greatest amount of PBL and practical clinical experience achieve the highest scores, whereas those students with curricula displaying lesser amounts of these elements attain lower results.

In order not to limit the resulting data to use within the survey groups and to screen the attained scores for plausibility, they were compared with the results measured by Bordage [29]. The lowest measured total score

of a group in the present study (students from the UoC, 9th semester with 150.5 points) are below the results recorded by Bordage for students in their 3rd academic year (158.3 points), while results from the group with the highest achieved scores in our study (W/HU, 8th semester with 178.2 points) are above the results measured for senior house officers (168.4 points) and general practitioners (172.3 points). This suggests that DTI is particularly well-suited to intra-faculty comparisons due to the difficulty of calibration for varying systems. Whereas the Bordage study showed evidence of a continual score increase from first-year students (153.9 points) to registrars (180.2 points), the differences found between the surveyed groups in our study was greater.

The questionnaire was implemented in a non-standardised setting among three different medical schools in the German State of North Rhine-Westphalia. The results are therefore not necessarily representative for the respective cohorts. The variance of results in the data collected for the traditional course of study (7th semester, RUB) is particularly great, and the participants (40) represent only a small, randomly chosen group from the whole of the cohort. Considering this, it is questionable if the data are representative for students of this academic year on the whole.

In Cologne and at WH/U, the questionnaires were handed out following a progress test, which could explain the low response rate, the poor reliability and the lower overall results. The reliability and overall scores were higher at

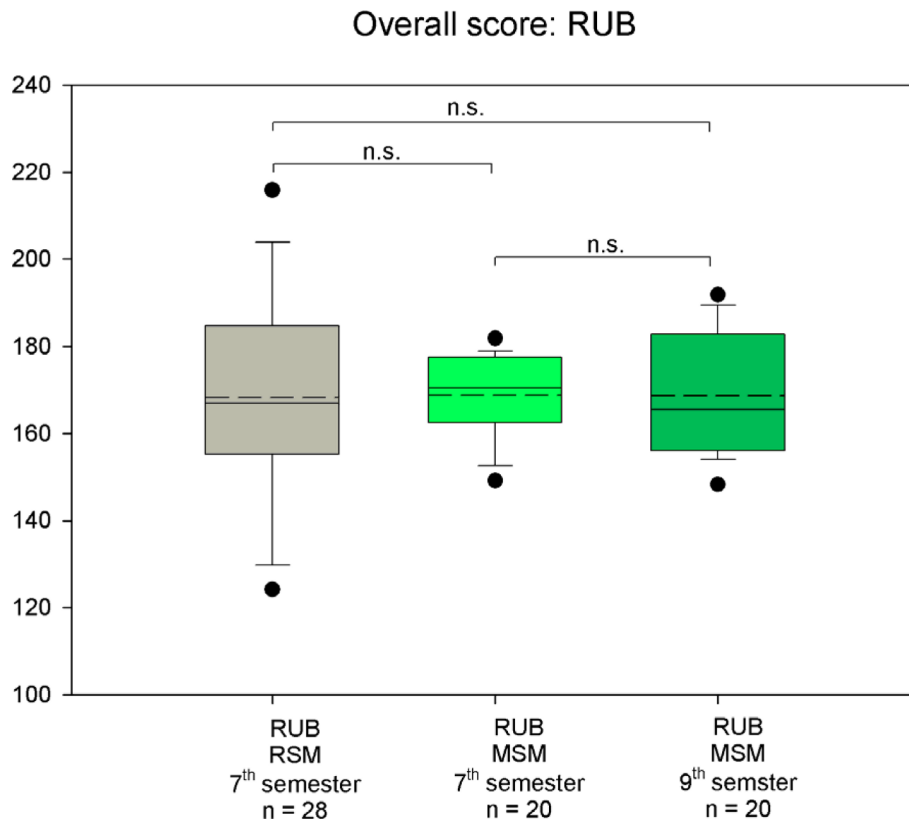


Figure 4: Overall Score RUB - Comparison of 7th semester traditional studies in medicine (German abbr. RSM) and 7th and 8th semester model studies in medicine (MSM) at Ruhr University Bochum. The respective maximum scores achieved in the questionnaires are depicted in a boxplot with median (solid line), mean score (dashed line), upper and lower quartile (box), the 5%/95% percentiles (whiskers) and minimum or maximum score (black circle). The respective p-value (p) and effect size (d) are listed above the comparisons.

W/HU. Greater peer pressure may have had a motivating effect in the smaller cohort in this case. The percentage of nonevaluable questionnaires from W/HU was also lower. Being that approximately one third of the questionnaires from the UoC were not evaluable, comparisons using their results should be viewed critically. Questionnaires without clear responses to all of the questions were not factored in. Despite the instruction to check one of the boxes between the proposed statements, many students chose to mark the dividing line between two boxes, evidently unable to decide on a value offered in the continuum. This was also observed in the questionnaires filled out by students from RUB and is the reason for the low percentage of evaluable returns from this location. Furthermore, we must bear in mind when interpreting the data that although the DTI is an established instrument, it does not objectively gauge clinical reasoning as such but serves as a measure of the self-assessment of students in this discipline. An (indeed desirable) overall critical stance in respect to their own competence could have a negative influence on the students' self-assessment concerning clinical reasoning. Undoubtedly, the surveyed cohorts can also differ in the extent of their self-criticism. No further instrument was implemented that could have served to standardise this aspect. One can assume, for instance, that male students tend to over-assess their competence [38]. The male/female distribution did not differ significantly in the surveyed groups,

however. For this reason, the differences discovered in this study should not be attributable to a gender-related overassessment among the students at W/HU and an underassessment among those in Cologne.

Conclusions

In summary, the present pilot study indicates that the DTI is an appropriate instrument for a comparative survey of self-assessed clinical reasoning among students with various curricula.

The study did, however, also show that further steps must be taken if inferences are to be drawn from the measured differences about the effect of apportionment of different curricular elements such as PBL or clinical-practical experiences.

Principally, it must be ensured that the survey is conducted under comparable circumstances, whereby a higher rate of return must be reached. Further, it would seem necessary to implement an additional inventory in order to investigate possible self-overassessments or self-underassessments by the students. The DTI itself should be further validated with other specialised, objective tests for clinical reasoning and conduct [39], for example key feature [40], script concordance [41] or situational judgement [42] tests.

Comparison of Overall Scores W/HU, UoC, RUB

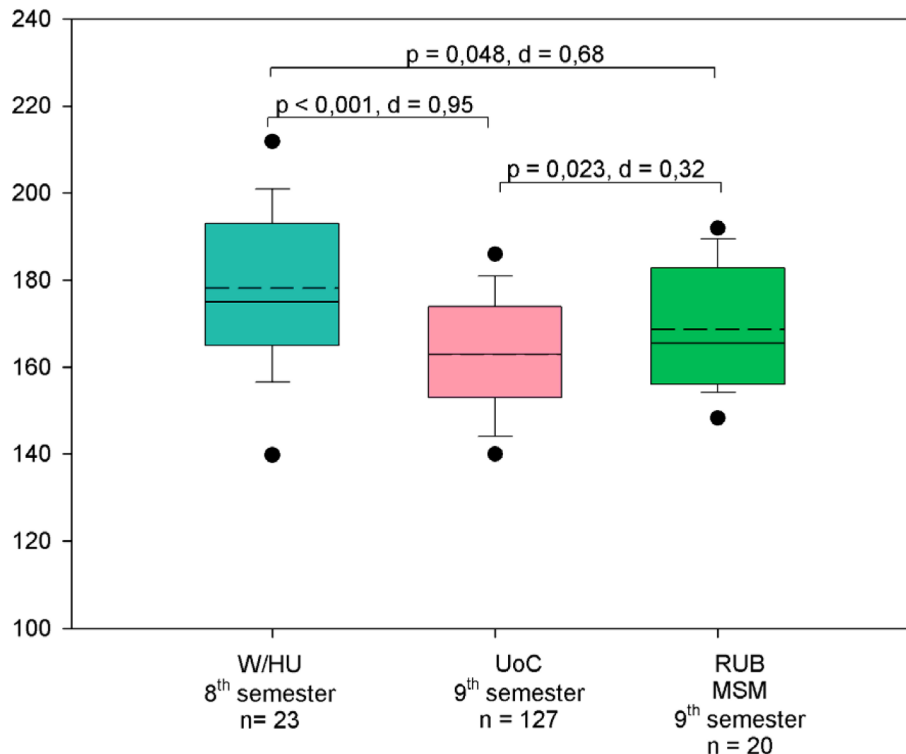


Figure 5: Comparison of Overall Scores W/HU, UoC, RUB - Comparison of the 8th and 9th semesters respectively from Witten/Herdecke University (German abbr. UW/H), the University of Cologne (German abbr. UzK) and Ruhr University Bochum (RUB). The respective maximum scores achieved in the questionnaires are depicted in a boxplot with median (solid line), mean score (dashed line), upper and lower quartile (box), the 5%/95% percentiles (whiskers) and minimum or maximum score (black circle).

The respective p-value (p) and effect size (d) are listed above the comparisons.

Competing interests

The authors declare that they have no competing interests.

References

- Barrows HS, Feltovich PJ. The clinical reasoning process. *Med Educ.* 1987;21(2):86-91. DOI: 1365-2923.1987.tb00671.x
- Pelaccia T, Tardif J, Tribby E, Charlin B. An analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. *Med Educ Online.* 2011;14:16. DOI: 10.3402/meo.v16i0.5890
- Norman G. Research in clinical reasoning: past history and current trends. *Med Educ.* 2005;39(4):418-427. DOI: 10.1111/j.1365-2929.2005.02127.x
- Charlin B, Tardif J, Boshuizen HPA. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Acad Med.* 2000;75(2):182-190. DOI: 10.1097/00001888-200002000-00020
- Schmidt HG, Norman GR, Boshuizen HP. A Cognitive Perspective on Medical Expertise: Theory and Implications. *Acad Med.* 1990;65(10):611-621. DOI: 10.1097/00001888-199010000-00001
- Schmidt HG, Boshuizen HP. On acquiring expertise in medicine. *Educ Psychol Rev.* 1993;5(3):205-221. DOI: 10.1007/BF01323044
- Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. *Med Teach.* 2003; 25(3):308-313. DOI: 10.1080/0142159031000100427
- Grant J, Marsden P. The structure of memorized knowledge in students and clinicians: an explanation for diagnostic expertise. *Med Educ.* 1987;21(2):92-98. DOI: 10.1111/j.1365-2923.1987.tb00672.x
- Eva KW, Neville AJ, Norman GR. Exploring the Etiology of Content Specificity: Factors Influencing Analogic Transfer and Problem Solving. *Acad Med.* 1998;73(10):S1-S5. DOI: 10.1097/00001888-199810000-00028
- Dory C, Gagnon R, Charlin B. Is case-specificity content-specificity? An analysis of data from extended-matching questions. *Adv Health Sci Educ.* 2010;15(1):55-63. DOI: 10.1007/s10459-009-9169-z
- Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ.* 2006;40(7):618-623. DOI: 10.1111/j.1365-2929.2006.02511.x
- Kassirer JP. Teaching clinical reasoning: Case-based and coached. *Acad Med.* 2010;85(7):1118-1124. DOI: 10.1097/ACM.0b013e3181d5dd0d
- Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ.* 2004;39(1):98-106. DOI: 10.1111/j.1365-2929.2004.01972.x

14. Wolf FM, Gruppen LD, Billi JE. Use of the competing-hypotheses heuristic to reduce 'pseudodiagnosticity'. *Acad Med*. 1988;63(7):548-554. DOI: 10.1097/00001888-198807000-00006
15. Kopp V, Möltner A, Fischer MR. Key-feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. *GMS Z Med Ausbild*. 2006;23(3):Doc50. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2006-23/zma000269.shtml>
16. Kopp V, Stark R, Fischer MR. Förderung von Diagnose-Kompetenz in der medizinischen Ausbildung durch Implementation eines Ansatzes zum fallbasierten Lernen aus Lösungsbeispielen. *GMS Z Med Ausbild*. 2007;24(2):Doc107. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2007-24/zma000401.shtml>
17. Gräsel C, Mandl H. Förderung des Erwerbs diagnostischer Strategien in fallbasierten Lernumgebungen. *Unterrichtswissenschaft*. 1993;21:355-370.
18. Maudsley G, Strivers J. Promoting professional knowledge, experiential learning and critical thinking for medical students. *Med Educ*. 2000;34(7):535-544. DOI: 10.1046/j.1365-2923.2000.00632.x
19. Kamin CS, O'Sullivan PS, Younger M, Deterding R. Measuring Critical Thinking in Problem-Based Learning Discourse. *Teach Learn Med*. 2001;13(1):27-35. DOI: 10.1207/S15328015TLM1301_6
20. Simpson E, Courtney M. Critical thinking in nursing education: Literature review. *Int J Nurs Pract*. 2002;8(2):89-98. DOI: 10.1046/j.1440-172x.2002.00340.x
21. Eshach H, Bittermann H. From case-based reasoning to problem-based learning. *Acad Med*. 2003;78(5):491-496. DOI: 10.1097/00001888-200305000-00011
22. Hmelo CE. Cognitive Consequences of Problem-Based Learning for the Early Development of Medical Expertise. *Teach Learn Med*. 1998;10(2):92-100. DOI: 10.1207/S15328015TLM1002_7
23. Rochmawati E, Wiechula R. Education strategies to foster health professional students' clinical reasoning skills. *Nurs Health Sci*. 2010;12(2):244-250. DOI: 10.1111/j.1442-2018.2009.00512.x
24. Tiwari A, Leighton-Beck L, So M, Yuen K. A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Med Educ*. 2006;40(6):547-554. DOI: 10.1111/j.1365-2929.2006.02481.x
25. Yuan H, Kunaviktikul W, Klunklin A, Williams BA. Improvement of nursing students' critical thinking skills through problem-based learning in the People's Republic of China: A quasi-experimental study. *Nurs Health Sci*. 2008; 10(1):70-76. DOI: 10.1111/j.1442-2018.2007.00373.x
26. Windish DM. An Innovative Curriculum Teaching the Integration of Communication and Clinical Reasoning Skills to Medical Students. Liverpool: MPH Capstone Project; 2004.
27. Bowen JL. Educational strategies to promote diagnostic clinical reasoning. *N Engl J Med*. 2006;355(21):2217-2225. DOI: 10.1056/NEJMr054782
28. Schmidmaier R, Eiber S, Ebersbach R, Schiller M, Hege I, Holzer M, Fischer MR. Learning the facts in medical school is not enough: which factors predict successful application of procedural knowledge in a laboratory setting? *BMC Med Educ*. 2013;13:28. DOI: 10.1186/1472-6920-13-28
29. Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. *Med Educ*. 1990;24(5):413-425. DOI: 10.1111/j.1365-2923.1990.tb02650.x
30. Jones UF. The reliability and validity of the Bordage, Grant & Marsden diagnostic thinking inventory for use with physiotherapists. *Med Teach*. 1997;19(2):133-140. DOI: 10.3109/01421599709019366
31. Sobral DT. Medical Students' Mindset for Reflective Learning: A Revalidation Study of the Reflection-In-Learning Scale. *Adv Health Sci Educ*. 2005;10(4):303-314. DOI: 10.1007/s10459-005-8239-0
32. Groves M, Scott I, Alexander H. Assessing clinical reasoning: a method to monitor its development in a PBL curriculum. *Med Teach*. 2002;24(5):507-515. DOI: 10.1080/01421590220145743
33. Round AP. Teaching clinical reasoning – a preliminary controlled study. *Med Educ*. 1999;33(7):480-483. DOI: 10.1046/j.1365-2923.1999.00352.x
34. Beullens J, Struyf E, van Damme B. Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Med Educ*. 2006;40(12):1173-1179. DOI: 10.1111/j.1365-2929.2006.02627.x
35. Jerant AF, Azari R. Validity of Scores Generated by a Web-Based Multimedia Simulated Patient Case Software: A Pilot Study. *Acad Med*. 2004;79(8):805-811. DOI: 10.1097/00001888-200408000-00017
36. Stieger S, Praschinger A, Kletter K, Kainberger F. Diagnostic grand rounds: A new teaching concept to train diagnostic reasoning. *Eur J Radiol*. 2009;78(3):349-352. DOI: 10.1016/j.ejrad.2009.05.015
37. Lee A, Joynt GM, Lee AKT, Ho A, Groves M, Vlantis AC, Ma RC, Fung CS, Aun CS. Using illness scripts to teach clinical reasoning skills to medical students. *Fam Med*. 2010;42(4):255-261.
38. Jünger J, Schellberg D, Nikendei C. Subjektive Kompetenzeinschätzung von Studierenden und ihre Leistung im OSCE. *GMS Z Med Ausbild*. 2006;23(3):Doc51. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/zma/2006-23/zma000270.shtml>
39. Ilgen JS, Humbert AJ, Kuhn G, Hansen ML, Norman GR, Eva KW, Charlin B, Sherbino J. Assessing diagnostic reasoning: a consensus statement summarizing theory, practice, and future needs. *Acad Emerg Med*. 2012;19(12):1454-1461. DOI: 10.1111/acem.12034
40. Page G, Bordage G. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Acad Med*. 1995;70(2):104-110. DOI: 10.1097/00001888-199502000-00012
41. Lubarsky S, Charlin B, Coe DA, Chalk C, van der Vleuten CP. Script concordance testing: a review of published validity evidence. *Med Educ*. 2011;45(4):329-338. DOI: 10.1111/j.1365-2923.2010.03863.x
42. Rahman M. Tackling situational judgment tests. *BMJ Careers*. 2007;334:gp189-gp190.

Corresponding author:

Dr. Kirsten Gehlhar, MME

Carl von Ossietzky University of Oldenburg, School of Medicine and Health Sciences, Carl von Ossietzky Str. 9-11, 26129 Oldenburg, Germany
 kirsten.gehlhar@uni-oldenburg.de

Please cite as

Gehlhar K, Klimke-Jung K, Stosch C, Fischer MR. Wirken sich unterschiedliche medizinische Curricula auf das selbst eingeschätzte klinische Denken von Studierenden aus? GMS Z Med Ausbild. 2014;31(2):Doc23.
DOI: 10.3205/zma000915, URN: urn:nbn:de:0183-zma0009152

This article is freely available from

<http://www.egms.de/en/journals/zma/2014-31/zma000915.shtml>

Received: 2013-07-02

Revised: 2014-03-13

Accepted: 2014-04-02

Published: 2014-05-15

Copyright

©2014 Gehlhar 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.