

Aus der Klinik für Anästhesiologie mit Schwerpunkt operative  
Intensivmedizin CCM & CVK  
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Evidenzbasierte medizinische Ausbildung: Notwendigkeit,  
Validität und Wege der Implementation von Ausbildungszielen

zur Erlangung des akademischen Grades  
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von

Wolf Hautz, geb. Blaum

aus Frankfurt am Main

**Datum der Promotion: 30.05.2015**

# Inhalt

ZUSAMMENFASSUNG .....	3
ABSTRACT .....	3
EINLEITUNG .....	4
GRÜNDE FÜR OBE .....	4
STAND DER UMSETZUNG .....	5
PROBLEMDEFINITION .....	7
ZIELSTELLUNG .....	8
METHODEN .....	9
ZIELDEFINITION WEITERBILDUNG .....	9
<u>Teilnehmer</u> .....	9
<u>Inhalte</u> .....	9
<u>Konsensverfahren</u> .....	10
VALIDIERUNG „KONSENSUSSTATEMENT PRAKTISCHE FERTIGKEITEN“ .....	10
IDENTIFIKATION VON TAXONOMIEN UND ONTOLOGIEN .....	11
<u>Suche</u> .....	11
<u>Kriterien zur Evaluation der Suchergebnisse</u> .....	12
ERGEBNISSE .....	13
ZIELDEFINITION WEITERBILDUNG .....	13
VALIDIERUNG „KONSENSUSSTATEMENT PRAKTISCHE FERTIGKEITEN“ .....	14
IDENTIFIKATION VON TAXONOMIEN UND ONTOLOGIEN .....	15
DISKUSSION .....	17
LIMITATIONEN .....	20
QUELLEN .....	22
EIDESSTATTLICHE VERSICHERUNG .....	24
ANTEILSERKLÄRUNG .....	25
DRUCKEXEMPLARE DER AUSGEWÄHLTEN PUBLIKATIONEN .....	25
LEBENS LAUF .....	64
PUBLIKATIONEN .....	65
DANKSAGUNG .....	66

## **Zusammenfassung**

Ergebnisorientierte Ausbildung (outcome based education OBE) wird international als Standard medizinischer Ausbildung angesehen. Während international OBE in der medizinischen Aus- und Weiterbildung fest etabliert ist, fehlen bisher geeignete outcome frameworks (OF) für die Aus- und Weiterbildung in Deutschland. Zudem ist die Implementation von OBE wegen der Vielzahl der miteinander zu verknüpfenden Elemente aufwendig und wenig fortgeschritten. Am Beispiel der Facharztweiterbildung Anästhesiologie werden in dieser Arbeit mittels eines Konsensusverfahrens erstens Möglichkeiten einer Zieldefinition als Grundlage der Definition eines OF für die Weiterbildung exemplarisch erprobt. Zweitens wird das „Konsensusstatement praktische Fertigkeiten im Medizinstudium“ als ein bestehendes OF mit einem existierenden Curriculum vergleichen und so die Verwendbarkeit lernzielorientierter OF für die Curriculumsentwicklung untersucht. Anhand einer systematischen Literaturrecherche werden drittens Taxonomien und Ontologien identifiziert, die zur Verschlagwortung curricularer Elemente der medizinischen Aus- und Weiterbildung geeignet sind.

## **Abstract**

Outcome-based education OBE is regarded as the gold standard of medical education internationally. While internationally OBE is firmly established in medical education, Germany lacks usable outcome frameworks (OF) for under- and postgraduate education. The implementation of OBE is further complicated by the large number of curricular elements to be linked among each other. Using residency training in anaesthesiology as one example, this work aims to first evaluate possibilities of OF definition using a consensus method. Second, the established consensus statement on practical skills in medical school as one OF is compared to an existing curriculum to thus evaluate the usability of learning objective based OF. Third a systematic literature search then leads to the identification of taxonomies and ontologies that allow for the annotation of curricular elements in under- and postgraduate medical education.

## **Einleitung**

Ergebnisorientierte Ausbildung („outcome based education“ (OBE))<sup>1</sup> wird international zurzeit als Standard medizinischer Ausbildung angesehen [2].

Dabei wird die Ausbildung ausgehend von den Zielen der Ausbildung gestaltet, welche üblicherweise im Rahmen meist nationaler Konsensverfahren definiert und in Ergebnisrahmenwerken („outcome frameworks“ (OF)) festgehalten werden [3].

OBE unterscheidet sich von anderen edukativen Konzepten vor allem durch diese deduktive Ableitung von Inhalten und Formaten für Unterricht und Prüfung. Im Gegensatz dazu gehen z.B. problemorientierte Konzepte davon aus, dass durch die Befassung mit sorgfältig ausgewählten Problemen des späteren Aufgabengebiets eine zielgerichtete Lernmotivation und ein inhaltlich relevanter Lernzuwachs induziert werden [4].

### ***Gründe für OBE***

Die Gründe für die weltweit zunehmende Bedeutung von OBE sind vielfältig.

Durch die global zunehmende Mobilität medizinischen Personals steigt der Bedarf an gegenseitiger Anerkennung nationaler Abschlüsse. Auch die fortschreitende europäische Integration setzt die Vergleichbarkeit nationaler Abschlüsse voraus, die über den Vergleich verschiedener OF, die Zusammenführung mehrerer unterschiedlicher OF oder die Entwicklung eines gemeinsamen OF realisiert werden kann [5].

Weiterhin sind Bildungssysteme und –Einrichtungen zunehmend rechenschaftspflichtig gegenüber der sie finanzierenden Gesellschaft [6]. In Deutschland sind die Universitäten in der Ausbildung von Ärztinnen und Ärzten gesetzlich verpflichtet, „Kenntnisse, Fähigkeiten und Fertigkeiten in allen Fächern [zu] vermitteln, die für eine umfassende Gesundheitsversorgung der Bevölkerung erforderlich sind“ [7]. OBE erleichtert es Bildungseinrichtungen, die Ziele, Formate und Inhalte der von ihnen angebotenen Ausbildung zu rechtfertigen [6].

---

<sup>1</sup> In dieser Arbeit werden „outcome based education (OBE)“ und „competency based education (CBE)“ synonym verwendet. Nach Albanese wird der Unterschied zwischen beiden Konzepten im wesentlichen durch die Perspektive der Lernenden bestimmt. Während „outcome frameworks (OF)“ in OBE aus Sicht der ausbildenden Einrichtung beschreiben, welche ziele Absolventen erreichen (sollen), beschreiben OF in CBE, welche Kompetenzen Absolventen besitzen [1].

Daneben soll OBE die Planung und Umsetzung von Curricula vereinfachen [8]. Der Abgleich aller Ausbildungsinhalte und –Formate mit einer konsentierten Zieldefinition soll Redundanzen ebenso wie Lücken im Curriculum verhindern. Zudem werden Prüfungsformate und –Inhalte durch ihre Ableitung aus einem OF transparent und selbst überprüfbar [3].

Neben diesen eher praktischen Gründen sprechen auch lerntheoretische Überlegungen für OBE. Biggs und Tang haben diese Überlegungen in ihrem „constructive alignment concept“ zusammen gefasst [9]: Dabei gehen sie, entsprechend der konstruktivistischen Lerntheorie davon aus, das Lernende Zusammenhänge zwischen neuen Erkenntnissen und bereits vorhandenem Wissen durch ihr Tun konstruieren und setzen dieses Konstruieren dem Lernen gleich. Die Aufgabe Lehrender besteht demnach darin, alle Aktivitäten des Lernenden auf die gewünschten Ergebnisse des Lernens auszurichten. Empirische Befunde unterstützen dieses Konzept insofern, als dass Zielklarheit und Zielverbundenheit die Zufriedenheit und Leistung Studierender fördern [10].

## ***Stand der Umsetzung***

Als frühes nationales OF wurde 1996 das CanMEDs framework for specialist physicians vom Royal College of Physicians and Surgeons of Canada verabschiedet [11]. Dabei handelt es sich um eine viel zitierte, inzwischen in der dritten Auflage erschienene Beschreibung von Wissen, Fähigkeiten und Fertigkeiten, über die Fachärzte<sup>2</sup> verfügen sollen. Zahlreiche, auch für Studierende entwickelte nationale OF referenzieren die CanMEDs, oft ohne auf die teilweise erheblichen Unterschiede in den Zielgruppen einzugehen [12].

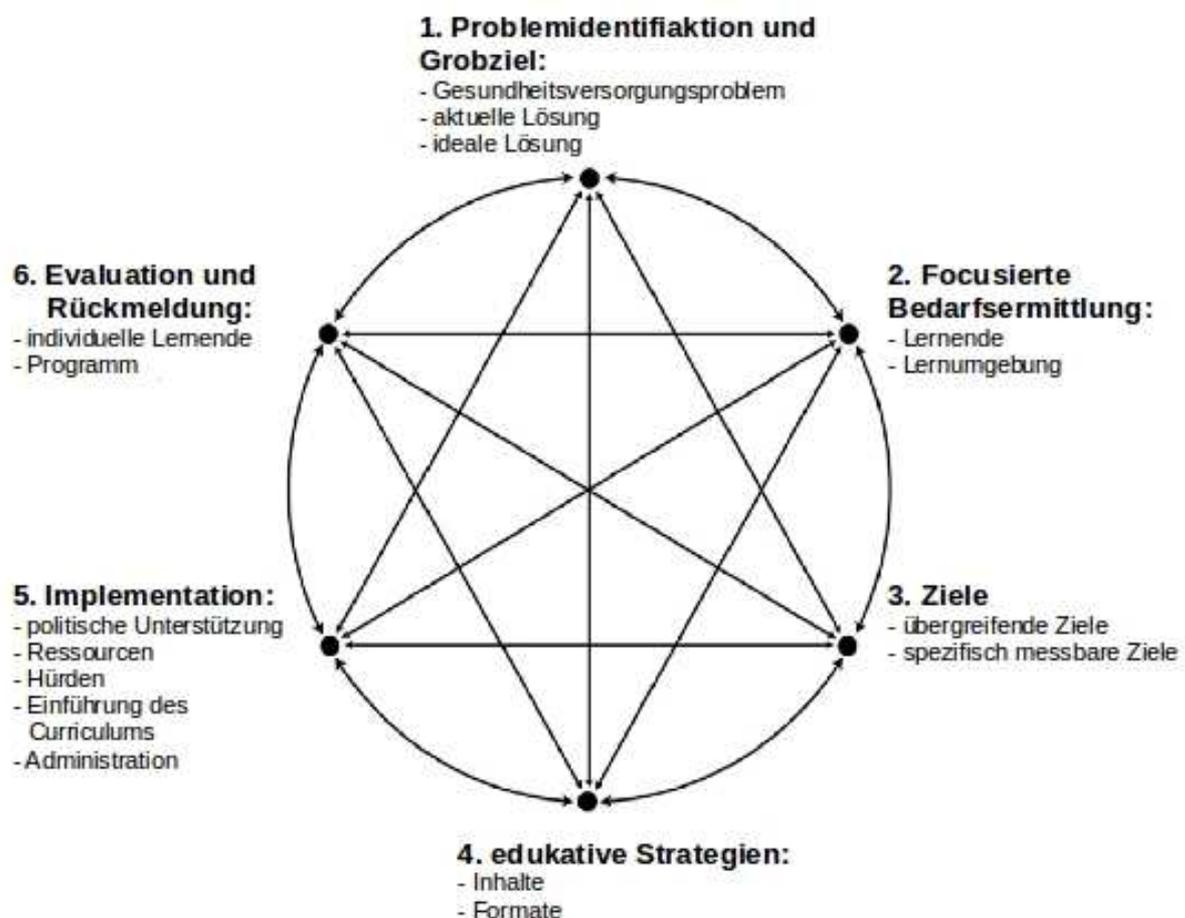
In Deutschland hat die Kultusministerkonferenz den Medizinischen Fakultätentag (MFT) und die Gesellschaft für Medizinische Ausbildung (GMA) mit der Entwicklung eines Nationalen Kompetenzbasierten Lernzielkataloges Medizin (NKLM) für das Studium beauftragt [13]. Derzeit befindet sich der NKLM in einer Entwurfsfassung in der Abstimmung mit den medizinischen Fakultäten Deutschlands sowie den Fachgesellschaften der Arbeitsgemeinschaft medizinisch-wissenschaftlicher Fachgesellschaften (AWMF). Für den Bereich der Weiterbildung existieren die (Muster-)

---

<sup>2</sup> Um die Lesbarkeit der Arbeit zu steigern, wird in dieser Arbeit auf die explizite Nennung beider Geschlechter verzichtet. Es sind stets beide Geschlechter gleichermaßen gemeint.

Weiterbildungsordnungen der Ärztekammern [14]. Die Entwicklung eines Nationalen Kompetenzbasierten Lernzielkatalogs Weiterbildung Medizin (NKLWM) wurde kürzlich angeregt [15]. Auf europäischer Ebene existiert mit dem TUNING-Project ein Versuch, eine übergreifende Zieldefinition aus den bestehenden nationalen OF zu entwickeln [5].

Als Standard der Planung, Implementation und Evaluation outcomeorientierter medizinischer Curricula haben sich zyklische Curriculumsmodelle etabliert. Ein weit verbreitetes Modell der Curriculumsentwicklung ist der Zyklus nach Kern [16]. Dieser sieht nach der Definition zunächst globaler Ziele die Erstellung (meist operationalisierter) Lernziele vor. Orientiert an diesen Zielen werden geeignete Unterrichtsformate ausgewählt, deren Umsetzung organisiert sowie deren Wirkung evaluiert (wobei auch klassische Prüfungen aus curricularer Sicht als Evaluation der Passung von Formaten und Implementation mit den Zielen verstanden werden können [16]) – vgl. Abbildung 1.



**Abbildung 1:** Curriculumsentwicklung nach Kern. Übersetzt aus [16]

## ***Problemdefinition***

Während international OBE in der medizinischen Aus- und Weiterbildung fest etabliert ist [2,3] fehlen bisher OF für die Aus- und Weiterbildung in Deutschland. Bestehende ausländische OF sind nicht einfach auf den Kontext anderer Nationen übertragbar [17]. Dies liegt einerseits an den teils erheblichen Unterschieden in Zielgruppen und Strukturen der OF [12]. Andererseits unterschieden sich auch die Aus- und Weiterbildungssysteme verschiedener Nationen in ihrer Dauer sowie ihre zeitlichen und strukturellen Gliederung [18].

Die 2009 erstmals und 2011 zum zweiten Mal von der Bundesärztekammer beauftragte bundesweite Evaluation der Weiterbildungen in Deutschland hat unter anderem eine erhebliche Notwendigkeit der Entwicklung strukturierter Weiterbildungsprogramme aufgezeigt [19,20]. Notwendige Voraussetzung dafür sind klare Zieldefinitionen [16], die in Deutschland bisher fehlen [15]. Die (Muster-)Weiterbildungsordnungen der Ärztekammern beschreiben bisher regelhaft kaum Ziele einer Facharztweiterbildung als vielmehr strukturelle Merkmale wie etwa die Dauer der Weiterbildung, die Anzahl durchgeführter Prozeduren oder den Mindestumfang verschiedener Rotationen [14].

Für das Studium befindet sich in Deutschland mit dem NKLM ein OF in der Entwicklung [13]. International existieren verschiedene Ansätze zur Definition von OF. Während beispielsweise die CanMEDs auf wenigen Seiten die Rollen und dafür erforderlichen Kompetenzen eines Arztes beschreiben [11], umfasst der Schweizer Lernzielkatalog (SCLO) mehrere hundert teils sehr spezifisch formulierte Lernziele [21]. Einzelne in den NKLM eingeflossene Arbeitspakete, wie das „Konsensusstatement praktische Fertigkeiten im Medizinstudium“ [22], sind bereits veröffentlicht und im Detailniveau ihres Inhalts eher dem SCLO als den CanMEDs vergleichbar. Erkenntnisse zur Validität einzelner OF existieren bisher ebenso wenig wie Informationen zu ihrer Wirkung auf die Implementation medizinischer Curricula.

Ein weiteres Problem outcomeorientierter Studiengänge ist deren praktische Umsetzung. Die Zuordnung jeder einzelnen Unterrichtsveranstaltung und Prüfungsaufgabe zu den Zielen des Curriculums, wie sie etwa das Curriculumsmodell nach Kern vorsieht [16], ist logistisch anspruchsvoll und wird durch die zyklischen

Änderungen des Curriculums zudem kompliziert. Allein die grundständigen Studiengänge der Charité umfassen rund 40.000 Lehrveranstaltungen pro Semester [23], denen curriculare Elemente wie Lernziele, Formate, Prüfungsaufgaben, geeignete Ressourcen und Evaluationsdaten zugeordnet werden sollen. Eine technische Möglichkeit dieser Zuordnungen ist die Software-gestützte Kartierung des Curriculums [24]. Bisher allerdings existieren kaum Lösungen, curriculare Elemente einander automatisiert zuzuordnen, wie dies durch die Verschlagwortung und semantische Zuordnung aller Elemente möglich wäre [25]. Insbesondere fehlt in Deutschland eine Übersicht über geeignete Schlagwortkataloge, wie sie strukturierte Taxonomien und Ontologien bieten.

## **Zielstellung**

In der vorliegenden Arbeit werden entlang des in Abbildung 1 dargestellten Modells drei Aufgaben bearbeitet:

- 1) Am Beispiel der Facharztweiterbildung Anästhesiologie sollen Möglichkeiten einer Zieldefinition als Grundlage der Definition eines OF für die Weiterbildung exemplarisch erprobt werden.
- 2) Am Beispiel des „Konsensusstatement praktische Fertigkeiten im Medizinstudium“ [22] soll die Verwendbarkeit lernzielorientierter OF für die Curriculumsentwicklung untersucht werden.
- 3) Anhand einer systematischen Literaturrecherche sollen Taxonomien und Ontologien identifiziert werden, die zur Verschlagwortung curricularer Elemente der medizinischen Aus- und Weiterbildung geeignet sind.

## **Methoden**

### ***Zieldefinition Weiterbildung***

Kern empfiehlt zur Definition von Zielen eines Curriculums einen „Stakeholder-approach“ [16]. Während die Ansprüche der Weiterbildner sowie der ärztlichen Fachgesellschaften an die Weiterbildung durch deren Einfluss auf die existierenden Weiterbildungsordnungen gesichert ist, fehlen bisher klare Ziele der Weiterzubildenden. Zur Identifikation der Ziele dieser Gruppe haben wir einen eintägigen Konsensusworkshop mit Weiterbildungsassistenten durchgeführt, der im Folgenden aus [26] zusammengefasst ist:

### Teilnehmer

Es wurden die Assistentensprecher der Anästhesiologie aller 39 deutschen Universitätsklinika eingeladen: Zunächst wurden alle universitären Anästhesiologien per Email um Entsendung eines Assistentensprecher gebeten. Für alle Einrichtungen ohne Rückantwort wurde die Einladung wiederholt. Die Assistentensprecher der verbliebenen Einrichtungen ohne Antwort wurden persönlich per Email und Telefon kontaktiert. 31 Assistentensprecher (79,5%), davon 14 (45,2%) weiblich, haben am Workshop teilgenommen. Deren mittlere Weiterbildungsdauer lag bei 3,16 Jahren (Median 4, Minimum 2, Maximum 7). Die Datenerhebung wurde dem Datenschutzbeauftragten der Charité angezeigt, auf eine umfängliche Prüfung des Vorhabens hat die Ethikkommission bei fehlender Patientenbeteiligung verzichtet. Alle Teilnehmer gaben ihre schriftliche Zustimmung zur Auswertung und Publikation der Ergebnisse.

### Inhalte

Zur Vorbereitung auf ihre Workshopteilnahme erhielten alle Teilnehmer zwei Wochen zuvor eine Liste von drei Workshopthemen: Charakteristika von 1.) guten Lehrenden, 2.) guten Bedingungen und 3.) guten Curricula. Die Themen entsprechen denen durch die bundesweite Evaluation der Weiterbildung identifizierten Handlungsschwerpunkten der Bundesärztekammer [20]. Jeder Teilnehmer wurde gebeten, sich für die Mitarbeit an einem der drei Themen zu entscheiden. Plätze wurden nach Eingang der Entscheidungen vergeben, jeder Teilnehmer konnte dem Thema seiner Wahl zugeordnet werden. Zur inhaltlichen Vorbereitung erhielten alle Teilnehmer themenspezifische Literatur.

## Konsensverfahren

Jede der drei Workshopgruppen wurde durch je einen Anästhesisten und einen weiteren Moderator geleitet. Aufgabe jeder Gruppe war die Zusammenstellung einer priorisierten Liste von themenspezifischen Zielen der Weiterbildung aus Sicht der Assistenten. Alle Gruppen setzten dazu die nominale Konsenstechnik (nominal group technique NGT) ein [27,28]. Die Ergebnisse aller drei Gruppen wurden in einer abschließenden Plenarveranstaltung konsentiert.

## **Validierung „Konsensusstatement praktische Fertigkeiten“**

Das „Konsensusstatement praktische Fertigkeiten im Medizinstudium“ [22] umfasst 289 Lernziele, von denen 232 als Kernziele gekennzeichnet sind. Jedes Lernziel ist einem, das Lernziel Nummer 275 zwei von 16 Organsystemen zugeordnet. Für drei definierte Zeitpunkte des Studiums ist weiter für jedes Lernziel eine zu diesem Zeitpunkt zu erreichende Tiefe definiert, die drei Stufen umfasst (siehe Tabelle 1).

Zur Validierung des Konsensusstatements und zur Einschätzung der Möglichkeit, Unterrichtsangebote anhand des Statements weiter zu entwickeln, wurde das Konsensusstatement mit dem etablierten und hervorragend evaluierten Tutorienangebot des Lernzentrums der Charité [29] verglichen. Das freiwillige und kostenfreie Tutorienprogramm umfasst jährlich rund 500 Termine zu einem von 48 Themen, wird jährlich von rund 4500 Studierenden wahrgenommen und wurde 2012 mit dem Preis für Lehrende Studierende der GMA ausgezeichnet [29]. Vier langjährige studentische Tutoren bewerteten die Tiefe, in der jedes der 289 Lernziele des Konsensusstatements in jedem der 48 Tutorien behandelt wird (13920 Bewertungen) in einem ersten Schritt unabhängig von einander anhand der in Tabelle 1 dargestellten Klassen. Aus diesen Bewertungen wurde die Interrater-Reliabilität als Korrelation nach Spearman errechnet. In einem zweiten Schritt wurden ungleiche Klassifikationen durch Diskussion und Konsens in eine endgültige Klassifikationen überführt, aus der die prozentuale Abdeckung der Lernziele des Konsensusstatements durch die Tutorien ermittelt wurde. Dabei wurden Gründe für den initialen Dissens protokolliert und abschließend gemeinsamen thematisch gruppiert, um Anhaltspunkte für die Validität des Konsensusstatements zu entwickeln. Eine tiefgehender Darstellung der Methodik findet sich in [30].

Tiefe	Bedeutung zur Klassifikation	Bedeutung im Konsensusstatement
0	Das Lernziel wird im Tutorium nicht behandelt.	keine
1	Die im Lernziel beschriebene Fertigkeit wird im Tutorium inklusive der theoretischen Voraussetzungen demonstriert.	demonstriert bekommen haben, inklusive der theoretischen Voraussetzungen
2	Die im Lernziel beschriebene Fertigkeit wird im Tutorium unter Aufsicht durchgeführt, ggf. am Modell.	unter Aufsicht durchgeführt haben, wenigstens einige Male
3	Die im Lernziel beschriebene Fertigkeit wird im Tutorium routiniert handwerklich eingesetzt, ggf. am Modell. Tutoriumsteilnehmer können Indikation und Konsequenzen benennen.	routiniert handwerklich können, situationsadäquat einsetzen können und die Konsequenzen kennen

**Tabelle 1:** Tiefenskala zur Klassifikation der Tiefe, in der ein Tutorium ein Lernziel behandelt. Adaptiert nach [22], aus [30]

## ***Identifikation von Taxonomien und Ontologien***

Zur Identifikation strukturierter Vokabulare einerseits und der Bewertung ihrer Eignung zur Verschlagwortung curricularer Elemente der medizinischen Ausbildung andererseits führten wir eine zweistufige Suche durch, die in [31] im Detail beschreiben ist:

### Suche

Anhand einer systematischen Suche in den Datenbanken der German Medical Science Publikationen identifizierten wir Publikationen, die mindestens eine Taxonomie oder Ontologie für die medizinische Ausbildung oder deren dortigen Einsatz beschreiben. Suchworte sind in Tabelle 2 dargestellt. Aus den so identifizierten Publikationen extrahierten wir im Konsens Kriterien zur Beurteilung der Eignung strukturierter Vokabulare für die Verschlagwortung curricularer Elemente der medizinischen Ausbildung.

Deutsch	5/37	English	15/564
medizinische Taxonomie	2/9	Medical education ontology	6/37
medizinische Ontologie	3/12	Taxonomy medical education german	0/93
deutschsprachige Taxonomie	0/2	Taxonomy medical education elearning	5/238
Taxonomie medizinische Lehre	0/2	taxonomy learning resources	1/127
Ontologie medizinische Lehre	0/6	medical education semantic network	2/15
Taxonomie medizinische Lernressourcen	0/3	medical education curriculum mapping taxonomy	0/18
Taxonomie medizinische Lerninhalte	0/3	taxonomy medical education learning objectives	1/36

**Tabelle 2:** Suchbegriffe der systematischen Literaturrecherche (Zahlenangaben: eingeschlossene Publikationen / Gesamtzahl der Treffer). Aus [31]

Anschließend dehnten wir die Suche unter Verwendung englischsprachiger Suchbegriffe auf die Medline-Datenbank aus (siehe Tabelle 2). Anhand der Abstracts wurden alle Publikationen zur weiteren Analyse eingeschlossen, die ein strukturiertes Vokabular mit Bezug zur medizinischen Ausbildung beschreiben und die in der Publikation referenzierte Taxonomie oder Ontologie aus dem Volltext identifiziert. Tabelle 2 stellt die verwendeten Suchbegriffe sowie die Anzahl der gefundenen und eingeschlossenen Publikationen zusammen.

### Kriterien zur Evaluation der Suchergebnisse

Aus den in Tabelle 2 (Deutsch) identifizierten Publikationen extrahierten wir im Konsens Kriterien zur Beurteilung der Eignung strukturierter Vokabulare für die Verschlagwortung curricularer Elemente der medizinischen Ausbildung. Diese Kriterien umfassen [31]:

Thema des Vokabulars: Es sollen Ausbildungsformate und -Methoden, sowie biomedizinische Inhalte beschreibbar sein.

Struktur: Das Vokabular soll Synonym- und Äquivalenzbeziehungen zwischen Termen unterstützen.

Sprache: Das Vokabular soll eine Verschlagwortung mit deutschsprachigen Termen unterstützen.

Umfang: Das Vokabular soll fein granular genug sein, um Inhalte und Methoden des Medizinstudiums in Deutschland differenziert beschreiben zu können, aber den annotierenden Nutzer nicht durch zu hohe Granularität zu irrelevanten Unterscheidungen zwingen.

Wartung: Das Vokabular sollte durch eine definierte Organisation oder Institution wenigstens jährlich gewartet werden.

Technik: Datenformat, Verfügbarkeit und Urheberrecht sollen einen kostengünstigen Einsatz zur webbasierten Beschreibung von Elementen und Ressourcen medizinischer Curricula erlauben.

# Ergebnisse

## **Zieldefinition Weiterbildung**

Unter den Teilnehmern der den Workshop abschließenden Plenarveranstaltung bestand Einvernehmen, dass Patientensicherheit Voraussetzung und Ziel jeder Weiterbildung sein muss und daher nicht separat unter den priorisierten Merkmalen guter Weiterbildung aufgenommen werden soll. Die von allen Workshopteilnehmern konsentierten Ergebnisse der einzelnen Gruppen sind in den Tabellen 3 - 5 dargestellt.

Ziel	Rang
Wille und Wunsch, weiterzubilden	1
Fachkompetenz auf neustem Stand	2
kann erklären	3
geduldig Anleiten/ Erklären	4
hat ein Konzept	4
Schaffen eines vertrauensvollen, offenen Klimas	6
kann beurteilen	7
Fähigkeit, sich Zeit zu nehmen	8
Tutorenausbildung	9
Kritikfähigkeit	10
Sicherheit für Weiterzubildende	11

**Tabelle 3:** Ergebnisse der Gruppe "Charakteristika guter Lehrender". Aus [26]

Ziel	Rang
Verbindlichkeit	1
Mitarbeiter Wertschätzung	2
Klare WB-Struktur	3
Supervision	4
Arbeitsklima (interprofessionell, interdisziplinär)	5
„Produktionsdruck“ überwinden	6
Fehlerkultur	7
Einhaltung des Tarifvertrags	8
Einhaltung des Arbeitszeitgesetzes	9
OP-Organisation WB geeignet	10
Dienstplansicherheit vs Flexibilität	11
Familienfreundlichkeit	12
Tägl. Feedback	13
SOPs	14
Transparente Personalpolitik	15

**Tabelle 4:** Ergebnisse der Gruppe "Charakteristika guter Bedingungen". Aus [26]

Ziel	Rang
Weiterbildungskultur (Zeit, Geld, Kultur des Lehrens und Lernens, Evaluation, ausgebildete Tutoren)	1
Abschnittsweise Definition von theoretischen und praktischen Lernzielen und deren Vermittlung	2
Strukturierte Einarbeitungen	3
Planbare strukturierte und transparente Rotationsplanung	4
Kommunikations- und Fehlerkultur	5
Simulatortraining regelmäßig im Team	6
Abschnittsweise summative und formative Überprüfung des theoretischen und praktischen Lernerfolges	7
Fortbildung intern und extern ermöglichen	8

**Tabelle 5:** Ergebnisse der Gruppe "Charakteristika guter Curricula" . Aus [26]

### ***Validierung „Konsensusstatement praktische Fertigkeiten“***

Die Übereinstimmung der Bewerter in der Frage, in welcher Tiefe welches Tutorium welches Lernziel des Konsensusstatements praktische Fertigkeiten abdeckt, schwankt je nach Organsystem, dem das Lernziel zugeordnet ist, erheblich. Während die Bewerter für die 192 Klassifikationen des Organsystems „Psyche“ eine maximale Übereinstimmung erzielten ( $r=1,0$ ), konnte für das Organsystem „Harn-/Geschlechtsorgane“ lediglich eine geringe Übereinstimmung erzielt werden ( $r=0,2$ ). Der Grad der Übereinstimmung der Bewerter ist dabei unabhängig von der Anzahl der Lernziele pro Organsystem.

Die protokollierten Gründe für Dissens in der Klassifikation lassen sich einem von drei Gründen zuordnen [30]: am häufigsten hatten die Bewerter Schwierigkeiten mit unscharf formulierten Lernzielen; beispielsweise war ihnen unklar, was unter Lernziel 115 „Indikationsstellung und Anordnung technischer Untersuchungen“ genau zu verstehen ist.

Eine zweite wesentliche Ursache für Dissens sahen die Bewerter in den teils erheblich unterschiedlich formulierten Lernzielen. Während etwa das Lernziel „Aktive und passive Untersuchung der oberen Extremität inkl. Schulter-, Ellenbogen- und Handgelenken, sowie der Langfinger und Daumen (insbesondere Inspektion, Auffinden anatomischer Landmarken, Durchführung der Neutral-Null-Methode und der Funktion für die Gelenke)“ sehr umfangsspezifisch formuliert ist, lässt beispielsweise das Lernziel 182 „Neuroradiologische Untersuchungsmethoden“ mehr Interpretation hinsichtlich Art und Umfang der zu erwerbenden Fähigkeiten zu.

Die drittens identifizierten lernzielunabhängigen Ursachen beziehen sich auf das Konsensusstatement als Ganzes. Beispielsweise war den Bewertern unklar, ob die Tatsache, dass einzelne Lernziele explizit Untersuchungen am Modell erwähnen, impliziert, dass dies für alle anderen Lernziele nicht gilt.

Das mit dem Konsensusstatement verglichene Tutorienangebot des Lernzentrums deckt 65,9% aller Lernziele und 73,7% der Kernziele des Konsensusstatements ab. Die Tiefe der im Konsensusstatement geforderten Abdeckung eines Lernziels nimmt über die Zeit zu, die Abdeckung durch das Tutorienangebot bis zu dieser Tiefe nimmt für die im Statement definierten Zeitpunkte ab: die Abdeckung der Lernziele in bis zur Famulaturreife geforderten Tiefe durch die Tutorien beträgt 63% für alle Lernziele (70,3% für die Kernziele bis zur Famulaturreife). Die bis zur PJ-Reife geforderte Tiefe wird zu 48,6% abgedeckt (Kernziele 53%), bis zur Weiterbildungsreife zu 39,3% (Kernziele 41,8%). Auch die Abdeckung der Lernziele unterscheidet sich nach Organsystem teils erheblich. So deckt das Tutorienangebot 90% der Lernziele des Organsystems Nervensystem (Kernziele 100%) aber nur 30% des Systems Atmung (Kernziele 50%) ab. Eine detaillierte Darstellung der oben zusammen gefassten Ergebnisse findet sich in [30].

### ***Identifikation von Taxonomien und Ontologien***

Es konnten, wie aus Tabelle 2 ersichtlich, 601 Publikationen identifiziert werden, von denen 20 eines von insgesamt 14 identifizierten strukturierten Vokabularen beschreiben. Mit dem „Unified Medical Language System“ (UMLS) existiert darüber hinaus eine Metaontologie, die ein semantisches Netz von über 130 einzelnen Terminologien umfasst. Die 14 identifizierten Vokabulare konnten einem der drei primären Anwendungsgebiete „Ausbildungsformate und -methoden“, „biomedizinische Inhalte“ sowie „Dokumentation und Verwaltung“ zugeordnet werden. Im Wesentlichen unterscheiden sich die Vokabulare jedoch in ihrem Aufbau durch die Art der Verknüpfungen zwischen ihren Termini: Die polyhierarchischen Ontologien einerseits umfassen in aller Regel nicht nur Eltern-Kind-Beziehungen zwischen den Termini, sondern ermöglichen auch Äquivalenz- und Assoziationsverweise. Die hier beschriebenen monohierarchischen Taxonomien andererseits lassen regelhaft nur eine Form von Verknüpfung ihrer Termini zu und sind allenfalls teilweise um eine

Verknüpfung vom Typ „Synonym“ ergänzt. Die einem Vokabular zugrunde liegende Struktur richtet sich bei den hier betrachteten Vokabularen wesentlich nach dem Zweck und den durch das Vokabular zu beschreibenden Inhalten. Die eher zur Verwaltung und Dokumentation entwickelten Vokabulare folgen einer monohierarchischen Struktur, die Ontologien zur Beschreibung von Ausbildungsformaten und -methoden sind, ebenso wie jene zur Beschreibung biomedizinischer Inhalte, polyhierarchisch strukturiert.

Drei der identifizierten Vokabulare zur Beschreibung von Ausbildungsformaten und -methoden liegen primär in Englisch, „Topics for Indexing Medical Education“ (TIME) darüber hinaus auch in Französisch vor. Die TIME-Ontologie ist zur inhaltlichen Beschreibung medizinischer Lehr- und Lerninhalte konzipiert, mit der Taxonomie der „Medical Education Taxonomy Research Organization“ (METRO) können vor allem Methoden und Prozesse der medizinischen Ausbildung annotiert werden. Der „British Education Thesaurus“ (BET) ist ein allgemeiner Ausbildungsthesaurus ohne spezifischen Bezug zur Medizin. Mit der „Ontology of Bio-Medical Educational Objectives“ (OBEO) steht darüber hinaus eine deutschsprachige Ontologie zur semantischen Annotation vor allem von Lernzielen in der medizinischen Ausbildung zur Verfügung. Eine ausführliche Darstellung der hier zusammen gefassten Ergebnisse und ein tabellarischer Vergleich der identifizierten Vokabulare findet sich in [31].

## **Diskussion**

OBE setzt die Definition von Zielen der Aus- oder Weiterbildung voraus, um auf deren Grundlage Unterrichts- und Prüfungsformate sowie deren Implementation zu entwickeln [3]. Derartige Zieldefinitionen fehlen in Deutschland bisher für die Weiterbildung [15] und sind nicht ohne weiteres aus dem Ausland übertragbar [17]. Eine gemeinsame Zieldefinition nach Kern [16] sollte die bestehenden Sichtweisen der Weiterbildungsbefugten und Kammern in Form der existierenden Weiterbildungsordnungen [14] unter anderem um die Perspektive der Weiterzubildenden ergänzen. Die in dieser Arbeit identifizierten Bedürfnisse der Weiterzubildenden der Anästhesie spiegeln dabei die in der Literatur beschriebene Notwendigkeit transparenter Ziele und Weiterbildungsstrukturen [15] wieder, die nach den Ergebnissen der bundesweiten Evaluation der Weiterbildung bisher nur unzureichend erreicht sind [19,20]. Die Ergebnisse des Konsensusworkshops bestätigen und erweitern zudem frühere Erkenntnisse zur anästhesiologischen Weiterbildung [32–34]. In einer Umfrage unter 770 Weiterzubildenden in deutschen Krankenhäusern konnte nur ein Drittel eine Struktur in ihrer Weiterbildung erkennen [32]. Diese Situation ist insbesondere unbefriedigend, weil davon auszugehen ist, dass strukturierte Programme zu effektiverem Lernen und damit verbesserter klinischer Praxis und ultimativ erhöhter Patientensicherheit führen [35].

Allerdings sind nicht alle in dieser Arbeit identifizierten Anforderungen der Weiterzubildenden an ihre anästhesiologische Weiterbildung ohne Weiteres umsetzbar. Der Wunsch nach formativen und summativen Prüfungen im Laufe der Weiterbildung erfordert die Implementation von Formaten der arbeitsplatzbasierten Prüfung wie der „Direct Observation of Procedural skills DOPS“, der „Mini-Clinical Examinations Mini CEX“ [36] oder „Multi Source Feedback MSF“ [37], die einen erheblichen Aufwand in Vorbereitung, Umsetzung und Auswertung erfordern. Bisher allerdings wird der Weiterbildungsaufwand den weiterbildenden Einrichtungen in Deutschland nicht über die Kosten der Krankenversorgung hinaus vergütet. Eine darüber hinaus gehende Finanzierung erscheint notwendig, auch weil Ertmer und Kollegen abnehmende Supervision in der anästhesiologischen Weiterbildung in Deutschland beobachten [34].

Dass sorgfältig definierte OF grundsätzlich geeignet sind, bestehende, auch gut bewertete Programme, weiter zu entwickeln, zeigt der Abgleich des Tutorienangebots des Lernzentrums der Charité mit dem „Konsensusstatement praktische Fertigkeiten“ [22]. Mit den gewonnenen Erkenntnissen können Lücken und Überlappungen im Tutorienangebot identifiziert werden. Durch den Vergleich der Behandlungstiefe eines Lernziels im Tutorium mit der zu unterschiedlichen Zeitpunkten im Konsensusstatement geforderten Tiefe ist es darüber hinaus möglich, Studierenden mögliche sinnvolle Reihenfolgen für ihre Tutoriumsteilnahme zu empfehlen. Das Konsensusstatement ist also geeignet, „bestehende Curricula neu zu strukturieren“ [22]. Gleichzeitig bietet der fortwährende Vergleich eines etablierten Programms mit dem es steuernden OF im Rahmen der zyklischen Curriculumentwicklung [16] (vgl. Abb. 1) Möglichkeiten zur kontinuierlichen Weiterentwicklung des OF. Das aktuelle Tutorienangebot deckt sowohl insgesamt als auch pro Organsystem einen deutlich höheren Anteil an Kernzielen ab, als es über die Gesamtheit der Lernziele des Konsensusstatements abdeckt. Die Priorisierung der Inhalte der Tutorien gleicht und unterstützt also die im Konsensusstatement getroffene Priorisierung. Weiter nimmt die geforderte Komplexität der Lernziele des Konsensusstatements über die drei Reifezeitpunkte zu. Die Abdeckung dieser Lernziele in der geforderten Tiefe durch die Tutorien nimmt sowohl insgesamt als auch getrennt nach Organsystemen ab. Die Komplexitätsbewertungen des Konsensusstatements gleichen also denen des Tutorienangebots, was die im Statement vorgenommene Tiefenforderungen unterstützt. Die sehr unterschiedliche Übereinstimmung der Bewerter in Abhängigkeit vom Organsystem lässt vermuten, dass sich die im Konsensusstatement formulierten Lernziele hinsichtlich der Möglichkeit ihrer Zuordnung zu konkreten Lehrveranstaltungen unterscheiden. Mögliche Ursachen könnten die aus den Gründen für Dissens extrahierten Themen „unscharfe Formulierung der Lernziele“, „Unterschiede im Umfang der Lernziele“ und „lernzielunabhängige Ursachen“ sein. Eine umfassende Diskussion der hier zusammen gefassten Aspekte findet sich in [30]. Eine systematische Aufarbeitung dieser Ansatzpunkte sollte zu einer Weiterentwicklung des Konsensusstatements beitragen [16].

In der Umsetzung von OBE allerdings stehen Lehrende und Planende vor der Herausforderung, die inhaltlichen Zusammenhänge in den von ihnen verantworteten Curricula zu kartieren [24], was bisher eine manuelle Verknüpfung von curricularen

Elementen voraussetzt. Diese Voraussetzung könnte einer der Gründe für die zwar zahlreich definierten aber wenig konsequent implementierten OF sein [38].

Besonders die Entwicklung des Internets hin zum semantischen Web (Web 3.0) [39] birgt ein erhebliches Potential zur automatisierten inhaltlichen Verknüpfung von Lernressourcen ebenso wie zum Einsatz in der Curriculumskartierung. Der Begriff „Web 3.0“ (auch „semantic Web“ genannt) bezeichnet ein Konzept zur Speicherung von Daten im Netz, welches durch Hinzufügen von maschinenlesbaren Metainformationen (z. B. mittels strukturierter Vokabulare) die Möglichkeit bietet, inhaltliche Verknüpfungen zwischen Datenobjekten von Maschinen automatisch generieren bzw. aufdecken zu lassen [39]. Voraussetzung dafür ist eine maschinell lesbare, inhaltliche Beschreibung der in den Daten enthaltenen Informationen [39]. Taxonomien und Ontologien stellen formalisierte Vokabulare zur inhaltlichen Beschreibung (semantische Annotation) von Objekten, wie zum Beispiel Lernressourcen, dar [25]. Durch die darüber hinaus in Taxonomien hinterlegten Beziehungen der Vokabeln untereinander kann eine (manuell angestoßene) automatische Suche curriculare Objekte mit inhaltlich ähnlichen Beschreibungen auch an voneinander unabhängigen Orten identifizieren [40]. Eine ausführliche Diskussion von Nutzen und Risiken strukturierter Vokabulare in der Curriculumsentwicklung bietet [31].

In dieser Arbeit konnte für den deutschsprachigen Raum keine ad hoc nutzbare Taxonomie oder Ontologie zur Annotation von Ressourcen und Inhalten der medizinischen Ausbildung identifiziert werden [31]: Die in deutscher Sprache verfügbaren strukturierten Vokabulare sind entweder - wie etwa OBEO - bisher nicht fein granular genug, um eine sinnvolle inhaltliche Unterscheidung zwischen Lernressourcen zu ermöglichen oder zwar sehr fein granular, aber eher für die inhaltliche Annotation von Patientendaten (SNOMED-CT) oder biomedizinischen Publikationen (MeSH) geeignet.

Die in englischer Sprache verfügbaren, speziell für die Domäne der medizinischen Ausbildung entwickelten Ontologien, wie TIME oder METRO, müssten zunächst ins Deutsche überführt werden. Dabei entstehen Probleme in der Auswahl inhaltlicher Entsprechungen. So findet der englischsprachige Begriff „medical education“ eine deutsche Entsprechung nur in der Trias aus medizinischer Aus-, Fort- und Weiterbildung, die im Englischen eher den Begriffen „graduate-“, „continous-“ und „postgraduate-“ medical education entsprechen.

Obgleich also in seiner praktischen Umsetzung OBE zahlreiche Schwierigkeiten bereitet, stellt OBE derzeit den Standard medizinischer Ausbildung dar [2]. Gleichzeitig ist die Frage, ob OBE kompetentere Ärzte hervor bringt, bisher empirisch nicht abschließend beantwortet [3]. Zudem ist zwar die steuernde Wirkung von OF auf Prüfungsformate und –Inhalte gut belegt, gleichzeitig aber bleibt ihre Wirkung auf Unterrichtsformate und –Inhalte ebenso wie die Struktur von Curricula wenig untersucht [3]. Zukünftige Arbeiten sollten neben der Entwicklung von Möglichkeiten zur tatsächlichen Umsetzung von OBE deren Wirkung auf die Ergebnisse der Aus- und Weiterbildung zum Gegenstand haben.

### ***Limitationen***

Neben der nur exemplarischen Bearbeitung dreier Aspekte von OBE – Definition von Zielen, Validität von Zieldefinitionen und Methoden zur Implementation – weist diese Arbeit einige Einschränkungen auf, die hier aus [26,30,31] zusammen gefasst sind:

Die am Konsensusworkshop zur Definition von Zielen der anästhesiologischen Ausbildung beteiligten Assistentensprecher kamen ausschließlich aus universitären Häusern. Ihre Ansichten könnten sich von denen der Assistenten in kleineren Häusern unterscheiden. Aktuell arbeiten etwa je ein Drittel der Weiterbildungsassistenten der Anästhesie in Deutschland in Häusern der Grund-, Regel – oder Maximalversorgung [34].

Zudem wurden ausschließlich Weiterzubildende, nicht jedoch Weiterbildende nach ihren Ansprüchen an die Weiterbildung befragt. Andere Studien konnten zeigen, dass beide Gruppen ähnliche Vorstellungen vom idealen klinischen Lehre – wenigstens in der Allgemeinmedizin – haben [41]. Der nächste Schritt zur Entwicklung eines OF für die anästhesiologische Weiterbildung in Deutschland könnte ein nationaler Delphiprozess zur Herstellung eines breiteren Konsenses sein, wie er in zahlreichen Publikationen zur Curriculumsentwicklung empfohlen wird [16,28].

Im Vergleich der Lernziele des Konsensusstatements „praktische Fertigkeiten“ [22] mit denen im Lernzentrum angebotenen Tutorien wird deren Behandlungstiefe durch unabhängige Bewerter/-innen klassifiziert. Die Interrater-Reliabilität wird als Korrelation nach Spearman berechnet, weil die Autoren die in Tabelle 2 dargestellte Skala als ordinal skaliert betrachten. Es existiert eine umfangreiche Diskussion in der Literatur zur Frage, ob Skalen wie die hier eingesetzte mit parametrischen statistischen Verfahren behandelt werden können (vgl. [42,43]). Einige Autoren berechnen für solche und selbst

für nominal skalierte Merkmale Intraklassen-Korrelationen (IKK). Die Berechnung von Korrelationen nach Spearman reduziert die Power der Untersuchung (und damit die Stärke der Korrelation), die Berechnung von IKKs würde stärkere Korrelationen ergeben, könnte aber den  $\alpha$ -Fehler erhöhen [42,43]. Insofern stellen die hier berechneten Korrelationen eine konservative Schätzung der tatsächlichen Übereinstimmung der Bewerter dar.

Zudem wurde nicht untersucht, inwiefern die studentische Bewertung der in ihren Tutorien behandelten Lernziele der Einschätzung durch approbierte Ärzte gleicht, die sich dazu allerdings zunächst detailliert praktische Erfahrungen mit dem Tutorienprogramm aneignen müssten.

Die Literaturrecherche nach strukturierten Vokabularen ist limitiert durch Suche und Auswahl der verwendeten Publikationen. Darüber hinaus besteht die Möglichkeit, dass strukturierte Vokabulare existieren, zu denen keine in den verwendeten Datenbanken indizierten Publikationen vorliegen.

In der Vorbereitung auf die hier dargestellte Arbeit konnten die Autoren zudem keine allgemein akzeptierten Kriterien zur Beschreibung und zum Vergleich strukturierter Vokabulare identifizieren. In einem ersten Schritt wurden deshalb zunächst in einer Teilmenge der durchsuchten Datenbanken Taxonomien identifiziert und aus diesen Publikationen die von den jeweiligen Autoren selbst verwendeten Kriterien zur Beschreibung ihres strukturierten Vokabulars ermittelt. Aus diesen Kriterien wurden für diese Arbeit im Konsens solche extrahiert, die den Autoren geeignet erscheinen, Taxonomien für die medizinische Ausbildung miteinander zu vergleichen. In der Auswahl dieser Kriterien haben die Autoren sich um die Generalisierbarkeit der Ergebnisse auf die Situation an den deutschsprachigen medizinischen Fakultäten bemüht. Lokal könnten andere oder weitere, hier nicht behandelte Kriterien von Bedeutung sein.

## Quellen

1. Albanese MA, Mejicano G, Mullan P, Kokotailo P, Gruppen L. Defining characteristics of educational competencies. *Medical Education*. 2008;42:248–55.
2. Cooke M, Irby DM, O'Brien BC. *Educating physicians: A call for reform of medical school and residency*. 1st ed. The Jossey-Bass higher and adult education series. San Francisco, CA: Jossey-Bass; 2010.
3. Morcke AM, Dornan T, Eika B. Outcome (competency) based education: an exploration of its origins, theoretical basis, and empirical evidence. *Adv in Health Sci Educ*. 2013;18:851–63.
4. Norman GR, Schmidt HG. The psychological basis of problem-based learning: a review of the evidence. *Acad Med*. 1992;67:557–65.
5. Cumming A, Ross M. The Tuning Project for Medicine – learning outcomes for undergraduate medical education in Europe. *Med Teach*. 2007;29:636–41.
6. David Hodges B. A Tea-Steeping or i-Doc Model for Medical Education? *Academic Medicine*. 2010;85:S34.
7. [http://www.bgbl.de/banzxaver/bgbl/start.xav?startbk=Bundesanzeiger\\_BGBI&bk=Bundesanzeiger\\_BGBI&start=/\\*%255B@attr\\_id=%2527bgbl112s1539.pdf%2527%255D#\\_\\_bgbl\\_\\_%2F%2F\\*%5B%40attr\\_id%3D%27bgbl112s1539.pdf%27%5D\\_\\_1403704248901](http://www.bgbl.de/banzxaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBI&bk=Bundesanzeiger_BGBI&start=/*%255B@attr_id=%2527bgbl112s1539.pdf%2527%255D#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl112s1539.pdf%27%5D__1403704248901) (accessed June 25, 2014).
8. Harden JCMDM. AMEE Guide No. 14: Outcome-based education: Part 5-From competency to meta-competency: a model for the specification of learning outcomes. *Med Teach*. 1999;21:546–52.
9. Biggs JB, Tang CS. *Teaching for quality learning at university*. 4th ed. Maidenhead: Open University Press; 2011.
10. Locke EA, Latham GP. *A theory of goal setting & task performance*. Englewood Cliffs, N.J: Prentice Hall; 1990.
11. [http://www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/framework/the\\_7\\_canmeds\\_roles\\_e.pdf](http://www.royalcollege.ca/portal/page/portal/rc/common/documents/canmeds/framework/the_7_canmeds_roles_e.pdf) (accessed June 25, 2014).
12. Hautz S, Blaum WE, Feufel MA, Spies CD. Comparability of Outcome Frameworks in Medical Education: Implications for framework development. *Med Teach*. 2014;under review.
13. Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT). *GMS Z Med Ausbild*. 2009;Doc35.
14. [http://www.bundesaerztekammer.de/downloads/20130628-MWBO\\_V6.pdf](http://www.bundesaerztekammer.de/downloads/20130628-MWBO_V6.pdf) (accessed June 25, 2014).
15. David DM, Euteneier A, Fischer MR, Hahn EG, Johannink J, Kulike K, et al. The future of graduate medical education in Germany - position paper of the Committee on Graduate Medical Education of the Society for Medical Education (GMA). *GMS Z Med Ausbild*. 2013;30:Doc26.
16. Kern DE, Thomas PA, Hughes MT. *Curriculum development for medical education: A six-step approach*. 2nd ed. Baltimore, Md: Johns Hopkins University Press; 2009.
17. Ringsted C, Hansen TL, Davis D, Scherpbier A. Are some of the challenging aspects of the CanMEDS roles valid outside Canada? *Med Educ*. 2006;40:807–15.
18. Wijnen-Meijer M, Burdick W, Alofs L, Burgers C, Cate O ten. Stages and transitions in medical education around the world: clarifying structures and terminology. *Med Teach*. 2013;35:301–07.
19. Hibbeler B, Korzilius H. Evaluation der Weiterbildung: Ein erster Schritt. *Dtsch Arztebl International*. 2010;107:A-417.
20. [http://www.bundesaerztekammer.de/downloads/EVA\\_Bundesrapport\\_final\\_16042010.pdf](http://www.bundesaerztekammer.de/downloads/EVA_Bundesrapport_final_16042010.pdf) (accessed June 25, 2014).
21. Bloch R, Bürgi H. The Swiss catalogue of learning objectives. *Med Teach*. 2002;24:144–50.

22. Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, et al. A consensus statement on practical skills in medical school - a position paper by the GMA Committee on Practical Skills. *GMS Z Med Ausbild.* 2011;28:Doc58.
23. [http://aco.charite.de/ueber\\_uns/stunden\\_raumplanung](http://aco.charite.de/ueber_uns/stunden_raumplanung) (accessed June 25, 2014).
24. Harden R. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach.* 2001;23:123–37.
25. Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. *Med Teach.* 2007;29:655–59.
26. Ortwein H, Blaum WE, Spies CD. Anesthesiology residents' perspective about good teaching--a qualitative needs assessment. *Ger Med Sci.* 2014;12:Doc05.
27. Delbecq AL. A Group Process Model for Problem Identification and Program Planning. *The Journal of Applied Behavioral Science.* 1971;7:466–92.
28. Lloyd-Jones, Fowell, Bligh. The use of the nominal group technique as an evaluative tool in medical undergraduate education. *Med Educ.* 1999;33:8–13.
29. [http://aco.charite.de/studierende/lernzentrum/tutorien/evaluation\\_der\\_tutorien/](http://aco.charite.de/studierende/lernzentrum/tutorien/evaluation_der_tutorien/) (accessed June 25, 2014).
30. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch A, Ahlers O. The practical use of the consensus statement on practical skills in medical school--a validation study. *GMS Z Med Ausbild.* 2012;29:Doc58.
31. Blaum WE, Jarczewski A, Balzer F, Stötzner P, Ahlers O. Towards Web 3.0: taxonomies and ontologies for medical education -- a systematic review. *GMS Z Med Ausbild.* 2013;30:Doc13.
32. Radtke RM, Hahnenkamp K. Weiterbildung im klinischen Alltag: Bestandsaufnahme und Strategien. *Anästh Intensivmed.* 2007;5:240-250.
33. Lehmann KA, Schultz J. Zur Lage der anästhesiologischen Weiter- und Fortbildung in Deutschland. *Der Anaesthesist.* 2001;50:248–61.
34. Ertmer C, van Aken H, Skorning M, Hahnenkamp K. Praxis der ärztlichen Weiterbildung und Rahmenbedingungen im Wandel – Evaluation der Weiterbildung im Fachgebiet Anästhesiologie über einen Zeitraum von fünf Jahren (2006–2011). *Anästh Intensivmed.* 2012;9:452-469.
35. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach.* 2007;29:642–47.
36. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach.* 2007;29:855–71.
37. Violato C. Multisource feedback: a method of assessing surgical practice. *BMJ.* 2003;326:546–48.
38. Harden RM. Outcome-based education – the ostrich, the peacock and the beaver. *Med Teach.* 2007;29:666–71.
39. Berners-Lee T, Fischetti M. Weaving the web: The original design and ultimate destiny of the World Wide Web by its inventor. New York, NY: HarperBusiness; 2010.
40. Willett TG, Marshall KC, Broudo M, Clarke M. It's about TIME: a general-purpose taxonomy of subjects in medical education. *Med Educ.* 2008;42:432–38.
41. Masunaga H, Hitchcock MA. Residents' and faculty's beliefs about the ideal clinical teacher. *Fam Med.* 2010;42:116–20.
42. Norman G. Likert scales, levels of measurement and the “laws” of statistics. *Adv in Health Sci Educ.* 2010;15:625–32.
43. Jamieson S. Likert scales: how to (ab)use them. *Med Educ.* 2004;38:1217–18.

## Eidesstattliche Versicherung

„Ich, Wolf Hautz, geb. Blaum, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: *Evidenzbasierte medizinische Ausbildung: Notwendigkeit, Validität und Wege der Implementation von Ausbildungszielen* selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -[www.icmje.org](http://www.icmje.org)) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.

Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

01.07.2014

---

Unterschrift

## Anteilerklärung

Wolf Hautz, geb. Blaum hatte folgenden Anteil an den folgenden Publikationen:

**Publikation 1:** Ortwein H, **Blaum WE**, Spies CD. Anesthesiology residents' perspective about good teaching--a qualitative needs assessment. Ger Med Sci. 2014;12:Doc05.

Beitrag im Einzelnen:

- Konzeption der Studie zu gleichen Teilen mit den Co-Autoren: 33%
- Datenerhebung und -Analyse gemeinsam mit der Erstautorin: 50%
- Entwurf und Korrektur des Manuskripts gemeinsam mit der Erstautorin: 40%

**Publikation 2:** **Blaum WE**, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch A, Ahlers O. The practical use of the consensus statement on practical skills in medical school--a validation study. GMS Z Med Ausbild. 2012;29:Doc58.

Beitrag im Einzelnen:

- Konzeption der Studie: 80%
- Datenerhebung: 20%
- Datenauswertung: 90%
- Entwurf und Korrektur des Manuskripts: 60%

**Publikation 3:** **Blaum WE**, Jarczewski A, Balzer F, Stötzner P, Ahlers O. Towards Web 3.0: taxonomies and ontologies for medical education -- a systematic review. GMS Z Med Ausbild. 2013;30:Doc13.

Beitrag im Einzelnen:

- Konzeption der Studie: 60%
- Datenerhebung: 25%
- Datenauswertung: 80%
- Entwurf und Korrektur des Manuskripts: 60%

Unterschrift, Datum und Stempel des betreuenden Hochschullehrers/der betreuenden Hochschullehrerin

---

Unterschrift des Doktoranden/der Doktorandin

---

# Anesthesiology residents' perspective about good teaching – a qualitative needs assessment

## Gute Facharztweiterbildung im Fach Anästhesie aus Sicht der Weiterbildungsassistenten – eine qualitative Bedarfsanalyse

### Abstract

**Background:** Germany, like many other countries, will soon have a shortage of qualified doctors. One reason for the dissatisfaction amongst medical residents are the relatively unstructured residency training programs despite increasing importance of outcome-based education. The aim of our study was to identify characteristics and requirements for good teaching during anesthesiology residency training from the resident's point of view.

**Methods:** A consensus workshop with residents from all medical universities in Germany was held. Participants were allocated to one of the three topics, chosen based on a 2009 nationwide evaluation of residency. The three topics were (A) characteristics of helpful/good teachers, (B) characteristics of helpful/good conditions and (C) characteristics of helpful/good curricular structure. Each group followed a nominal group technique consensus process to define and rank characteristics for a good residency.

**Results:** 31 (79.5%) resident representatives were present. The consented results put emphasis on the importance of structured curricula including transparent goals and objectives, in training formative assessments and quality assurance measures for the program. Residents further long for trained trainers with formal teaching qualifications and protected teaching time.

**Conclusions:** Good residency training requires careful consideration of all stakeholders' needs. Results reflect and extend previous findings and are at least to some degree easily implemented. These findings are an important step to establish a broader consensus within the discipline.

**Keywords:** residents, curriculum development, needs assessment, anesthesiology, consensus

### Zusammenfassung

**Hintergrund:** In Deutschland wird es, wie in vielen Ländern, in absehbarer Zukunft einen Mangel an qualifizierten Ärzten geben. Einer der Gründe für Unzufriedenheit unter den Weiterbildungsassistenten sind die relativ unstrukturierten Weiterbildungsprogramme trotz der zunehmenden Wichtigkeit von Ergebnis-orientierter Weiterbildung. Ziel dieser Studie ist es, Charakteristika von und Anforderungen an gute Lehre während der anästhesiologischen Facharztweiterbildung aus Sicht der Weiterzubildenden zu identifizieren.

**Methoden:** Es wurde eine Konsensus-Konferenz mit Weiterzubildenden von allen 39 anästhesiologischen Universitätskliniken Deutschlands durchgeführt. Die Teilnehmer wurden Gruppen zu einem von drei Themen zugeordnet, die anhand der nationalen Evaluation der Weiterbildung 2009 identifiziert wurden. Diese Themen waren (A) Charakteristika hilfreicher/guter Weiterbildner, (B) Charakteristika hilfreicher/guter Bedingungen und (C) Charakteristika hilfreicher/guter Curriculastruktur.

Heiderose Ortwein<sup>1</sup>

Wolf E. Blaum<sup>1,2</sup>

Claudia D. Spies<sup>1</sup>

1 Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Charité – Universitätsmedizin Berlin, Campus Virchow Klinikum und Campus Mitte, Berlin, Germany

2 Lernzentrum, Abteilung für Curriculaorganisation, Charité – Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany

tur. Jede Gruppe folgte dem Nominal-Group-Technik-Konsensprozess, um Charakteristika guter Weiterbildung zu definieren und zu priorisieren.

**Ergebnisse:** 31 (79,5%) Assistentensprecher waren anwesend. Die Ergebnisse betonen die Bedeutung strukturierter Curricula inklusive transparenter Ziele und Inhalte, formativer Zwischenprüfungen und Qualitätssicherungsmaßnahmen der Programme. Weiterbildungsassistenten wünschen sich zudem ausgebildete Weiterbildende mit formaler Lehrqualifikation und geschützte Zeit zum Lehren.

**Schlussfolgerung:** Gute Facharztweiterbildung setzt eine sorgfältige Berücksichtigung der Bedürfnisse aller Beteiligten voraus. Die Ergebnisse bestätigen und erweitern bisherige Publikationen und können zum Teil einfach umgesetzt werden. Die Ergebnisse sind ein wichtiger Schritt hin zu einem breiteren Konsens innerhalb einer Fachdisziplin.

**Schlüsselwörter:** Weiterbildungsassistenten, Curriculumsentwicklung, Bedarfsanalyse, Anästhesie, Konsensus

## Introduction

Germany, like many other countries, will soon have a shortage of qualified doctors. Many possible reasons for this shortage are discussed in the literature: The current generation of medical students (generation Y) are unlikely to accept poor training conditions. They may well leave the hospital or indeed the country to get better training [1], [2], [3]. Hospitals currently react to these new challenges with random isolated measures [3].

New planning tools are necessary to cope with the challenges medicine and especially anesthesiology will face in the future. Outcome-based education has become the gold standard in curriculum development for residency training [4], curricula focus on outcomes and their description [5]. Many countries have already developed a structured outcome-based curriculum for various levels of training (e.g., for specialties and subspecialties) in order to assure high quality of future health care [4], [6], [7], [8]. In Germany the German Association for Medical Education (Gesellschaft fuer Medizinische Ausbildung (GMA)) and The Council of Medical Faculties (Medizinischer Fakultäetentag (MFT)) have begun to develop a competency-based catalogue of learning objectives to inform undergraduate medical education [9]. A catalogue for the purpose of residency training is planned by the same Association [10].

In 2009 the German Medical Council (Bundesärztekammer) conducted the first nationwide evaluation of residency training in all medical subspecialties including residency trainers and residents. The evaluation was conducted analogous to German school marks (1 = excellent, 6 = poor). Global rating of residency training was fairly good with 2.54. There were eight question domains (global rating, medical expert teaching, learning culture, leadership, culture of error prevention, culture of decision making, workplace culture, application of evidence-based medicine). Compared to all specialties, anaesthesiology performed below average in seven of the eight question domains. Only "management of medical errors" received higher-than-average ratings. The nationwide evaluation

identified different fields where need for change was deemed necessary:

1. Trainers: need to develop a tutor system for residency, train the trainer programs for clinical teachers.
2. Conditions: need to develop family friendly hourly working schedule, optimization of pathways in hospitals and within a specialty department.
3. Curriculum: need to develop well-structured residency training curricula, publication of best practice models, reform of residency training regulations, possibility of residency training programs or rotations to different sites as part of a network [11], [12].

The follow-up survey in 2011 showed only slightly better results with the same pattern [13].

Given this situation there is a need for more structured approaches to curriculum development in Germany including development of appropriate learning goals, implementation and assessment as well as quality assurance of graduate medical education on a national level [10]. According to Kern different stakeholders have to be involved in this task. Societal and patient needs should be taken into account as well as the needs of clinical teachers, residents and employers [14]. While the needs of teachers, medical societies and employers at least partly manifest through council regulations of residency training programs, patient as well as residents needs are generally less frequently sought. We approached one of these groups to contribute to the necessary discussion. The aim of our study was to identify characteristics and requirements of good teaching in anaesthesiology residency training from the resident's point of view. We conducted a focused general needs assessment in the specialty of anaesthesiology as defined by Kerns six-step approach to curriculum development [14].

## Methods

In order to generate a generic list of characteristics and requirements for all three main topics of the above mentioned nationwide evaluation we conducted a con-

**Table 1: Topics and literature for preparation**

Topics	Literature for preparation
A. Characteristics of helpful/good teachers	<ul style="list-style-type: none"> <li>– AMEE Guide No. 27: Effective educational and clinical supervision [33]</li> <li>– Workplace-based assessment as an educational tool: AMEE Guide No. 31 [28]</li> </ul>
B. Characteristics of helpful/good conditions	<ul style="list-style-type: none"> <li>– Continuing medical education: AMEE Education Guide No 35 [34]</li> <li>– Learning in inter-professional teams: AMEE Guide No 38 [35]</li> <li>– Being a young and inexperienced trainee anesthetist: a phenomenological study on tough working conditions [36]</li> </ul>
C. Characteristics of helpful/good curriculum structure and rotation culture	<ul style="list-style-type: none"> <li>– ABC of learning and teaching in medicine. Curriculum design [37]</li> <li>– Competency-based education as a steering instrument for residency training in anesthesiology [38]</li> </ul>

sensus workshop for residents. It took place during one of the main conferences of the German Association of Anesthesiologists (Deutsche Gesellschaft fuer Anaesthesiologie und Intensivmedizin) in Berlin in September 2011 (13. Hauptstadtkongresses der Deutschen Gesellschaft fuer Anaesthesiologie und Intensivmedizin).

### Sample (participants)

To maximize legitimacy we invited one elected resident representative (Assistentensprecher) from all 39 German departments of anesthesiology at university hospitals [15]. The invitation was carried out in three steps: All anesthesiology departments at university hospitals were asked by e-mail to send one of their resident representatives to the workshop. In a second step all institutions that hadn't replied were reminded by a second e-mail and asked to send one of their resident representatives or a substitute. In a third step all resident representatives of non-responding institutions were contacted personally by WB by e-mail or phone and asked to attend or send a substitute. All workshop participants gave their written consent to publish the data gained in the consensus process as well as aggregated participant information. Ethical approval was deemed unnecessary for this workshop in our institution.

### Workshop preparation and topics

Participants were informed two weeks prior to the workshop about the three topics of the workshop (see Table 1), were asked to subscribe to one of the three workshop groups and to provide a preference list of the given main topics in case the preferred group could not be allocated to everyone. Workshop places were allocated on a first come first served basis, everyone was allocated to the group of their first choice (see Table 1). To familiarize participants with concepts and current debates in their topic all participants received two to three articles to prepare for the workshop five days before the workshop took place (see Table 1). Selected articles were reviews and related to the group topics as well as articles focused on residency training in anesthesiology related to group topics.

### Instruments: Workshop and Nominal Group Technique

The workshop was divided into three parts from noon to 5.30 pm. After a brief welcome and introduction we stated the workshop goal: to identify characteristics and requirements for good residency training in anaesthesiology from the resident's point of view. We split the group into the three working groups covering the topics (A) characteristics of helpful/good teachers, (B) characteristics of helpful/good conditions and (C) characteristics of helpful/good curricular structure and rotation culture. Each group was facilitated by at least one anesthesiologist and a second facilitator. At least one of the facilitators per group was familiar with the Nominal Group Technique (NGT) method and had applied it previously in other workshops. Groups were obliged to use the NGT to steer the workshop as recommended for clinical guideline development [16], [17], [18], [19]. The NGT consists of five steps: (1) brainstorming, (2) prioritizing, (3) collection of ideas, (4) discussion and clustering, (5) ranking (see Table 2). The result of the NGT is a prioritized list of idea-clusters and concepts. At the top of the list is the idea-cluster with the highest ranking priority followed by all idea-clusters in decreasing priority. In part three of the workshop, group results were presented to all participants and discussed in a plenary session. Group results could only be changed from the plenary if there was consensus that there is strong concern of misunderstanding within the group results. All participants consented to the results.

### Statistics

Data were collected by photography and manually transcribed. Descriptive statistics were carried out using SPSS version 20.

## Results

### Sample (participants)

31 (79.5%) resident representatives (or their substitute) out of 39 invited participants were present during the consensus process. 14 (45.2%) participants were female and 17 (54.8%) were male. Five (16.1%) of the parti-

**Table 2: The five stages of nominal group technique used during the workshop [16], [17], [19], duration and aim in the described consensus process**

Step	Duration	Description	Aim
Brain-storming (silent generation of ideas)	10 min	Every participant writes down all his/her ideas to the given topic. Every idea is put down on an individual card. Participants don't consult or discuss.	Open collection of ideas independent of group process to allow as many ideas to rise as possible.
Individual prioritization	5 min	Every participant performs individual ranking of relevance regarding his/her ideas.	Preparation of structured sharing
Sharing ideas	30 min	The first priority idea-cards are shared one at a time by all participants and collected on a pin-board. The ideas of lower priority subsequently follow. Every participant can explain the content briefly with no debate about the items. New ideas that arise from the sharing process should be written down also. The round robin process continues until all ideas have been presented or until 30 min have passed. Ideas that weren't mentioned within 30 min are discarded.	This process ensures all participants get an opportunity to make an equal contribution and provides a written record of all ideas generated by the group.
Discussion and clustering/ categorizing	60 min	Participants can seek verbal explanation or further details about any of the ideas that may not be clear to them. The process is kept as neutral as possible to avoid judgment and criticism. The group may suggest new items for discussion and combine items into categories.	Individual ideas are put together to concepts.
Ranking	15 min	Prioritizing the recorded categories or clusters by ranking them. All participants are to rank individually at the same time. The workshop question must be kept in mind. Rankings of all participants are summed up.	The concepts are ranked by importance. The result is a list of categories in order of highest priority first and less important categories last.

Participants had already passed the anesthesiology specialty examination.

Mean time of residency duration (excluding the candidates that had already passed their specialty examinations) was 3.16 years (median 4, minimum 2, maximum 7) out of a 5-year program. Only one participant had children.

## Consensus process

During the plenary presentation of the group results and the consensus discussion, the topic "patient safety" was heavily debated. Some groups had listed "patient safety" amongst their results. The plenary group discussed whether it was suitable to allocate "patient safety" to a prioritizing list even if this was not listed as the first priority. Patient safety was stressed as an overarching theme of residency training which was not explicitly matching any other specific category and therefore removed from the prioritizing lists to avoid misunderstanding. This was the only topic on which the plenary made a change to the results of the groups.

Group A worked on the topic "characteristics of helpful/good teachers" and developed eleven categories. Fifteen categories were developed by Group B on the topic "characteristics of helpful/good conditions". Group C worked on the topic "characteristics of helpful/good curricular structure and rotation culture" and developed eight categories. The consented categories of group A, B and C

with a ranking sum and in ranking order are summarized in Table 3, Table 4 and Table 5.

## Discussion

We were able to ensure a broad consensus within the group of workshop participants on the key characteristics for a good residency. The results give a broad overview of possible changes suggested by residents and a lot of the prioritized characteristics for a good residency would require little effort to put into action. Although the ranking orders are not equidistant, they provide an order of more and less importance.

Our results elaborate on the needs and expectations of generation Y. Most of the results from group A on the characteristics of helpful/good teacher like "Can patiently instruct and explain", "Creates a trusting and open learning climate" as well as the results B2 "Appreciates staff" and B5 "Good interprofessional and interdisciplinary working climate" of group B on the characteristics of helpful/good conditions strengthen the introductory claim, that medical students and junior residents especially from generation Y are not willing to accept poor training conditions.

The reported need for development of appropriate and transparent learning goals and structure [10] is reflected in our findings. Namely in the results A 4.2 "Has a concept", B3 "Clear structure to residency training" and C1 to C3 "Training culture", "Step-by step definition of

**Table 3: Results of group A “characteristics of helpful/good teachers”**

Number	Ranking place	Categories	Ranking sum
A 1	1	Will and wish to teach	12
A 2	2	Medical expertise is up to date	31
A 3	3	Can explain	48
A 4.1	4	Can patiently instruct and explain	50
A 4.2	4	Has a concept	50
A 6	6	Creates a trusting and open learning climate	62
A 7	7	Can assess	64
A 8	8	Able to take the time to teach	77
A 9	9	Has completed teacher training	85
A 10	10	Is open to criticism	89
A 10	11	Provides a safe environment for the resident	92

**Table 4: Results of group B “characteristics of helpful/good conditions”**

Number	Ranking place	Categories	Ranking sum
B 1	1	Accountability	43
B 2	2	Appreciates staff	44
B 3	3	Clear structure to residency training	60
B 4	4	Sufficient supervision	65
B 5	5	Good interprofessional and interdisciplinary working climate	70
B 6	6	Overcome pressure of production	82
B 7	7	Culture for managing of medical errors	89
B 8	8	Adherence to labor agreements	102
B 9	9	Adherence to laws on working hours	103
B 10	10	Operating room organization suitable for residency training	107
B 11	11	Reliable on duty-schedule versus flexibility	113
B 12	12	Family friendly workplace	120
B 13	13	Everyday feedback	122
B 14	14	Standard operating procedures	134
B 15	15	Transparent staff policy	138

**Table 5: Results of group C “characteristics of helpful/good curricular structure and rotation culture”**

Number	Ranking place	Categories	Ranking sum
C 1	1	Training culture (time, money, culture of teaching and learning, evaluation, trained teachers)	25
C 2	2	Step-by step definition of theoretical and practical learning objectives and corresponding teaching	33
C 3	3	Structured familiarization period	36
C 4	4	Predictable, structured and transparent rotation planning	38
C 5	5	Communication and culture for management of medical errors	52
C 6	6	Regular simulator training as a team	59
C 7	7	Step-by step summative and formative assessment of theoretical and practical skills/learning success	66
C 8	8	Education internally and externally possible	71

theoretical and practical learning objectives and their teaching”, “Structured familiarization period”. These requirements are not met in all programs as the nationwide surveys by the Bundesärztekammer demonstrates [11], [12], [13]. The results also confirm and extend previous

findings on residency training in the specialty of anesthesiology [20], [21], [22], [23]. One survey of 770 residents at German hospitals revealed that only about one third of residents are able to find a structure in their training program [23]. In the most recent survey, only 42.2% of

residents rated their training program as structured [24]. This situation is unsatisfactory, as one would assume that well-structured residency curricula would facilitate a more effective learning environment, followed by enhanced clinical practice and improved patient safety in future generations of physicians [25], [26]. Necessary changes in this domain are relatively easy to achieve and are inexpensive. They might be implemented by political will in single hospitals.

The topic of implementation and assessment of competency-based graduate medical education [10] is reflected in A7 "Can assess" and C7 "Step-by-step summative and formative assessment of theoretical and practical skills/learning success". Changes in assessment methods like performance testing, namely OSCE and portfolio [27], workplace-based assessment e.g. using DOPS or Mini-CEX [28] or even multi source feedback [29], [30] and others need content experts, are labor intensive and therefore need nationwide initiatives and financial support.

The important aspect of quality assurance of graduate medical education [10] is reflected in the following findings: "Predictable, structured and transparent rotation planning", B1 "Accountability", as well as A9 "Has completed teacher training". Taking into account that hospitals are not reimbursed for residency training in the German medical system, departments can rarely afford staff specifically dedicated to clinical teaching and supervision.

The proposed changes may only be feasible if they are developed and implemented nationally to save costs. Financial support to develop more structured curricula and provide more supervision seems necessary as Ertmer reports a decrease in clinical supervision of anaesthesiology residents in German hospitals [24].

Further research and regular program evaluation is needed to reveal which measures are more educationally and financially effective than others [10].

There are several limitations to the interpretation of the presented data. We invited only one elected resident representative (Assistentensprecher) per anesthesiology department from all German university hospitals. Their point of view may vary from residents trained in smaller hospitals. In the most recent anesthesiology survey 32% of residents were trained in small hospitals (Regelversorgung), 39% at large hospitals with all subspecialties, 29% at university hospitals [24].

In addition, only one of the participants had children, which may explain why family friendliness at the workplace does not seem to be of great importance in our results. Mean time of residency duration in our study population was slightly above three years out of a five-year program. 45.2% of our participants were female. These aspects might play a role in preferences. A Swiss survey reported that career related issues became less important for specialty choice with advancing training status. The same survey reported work- and time-related aspects are more important for women than for men.

With an increasing number of female residents, training programs should adapt to this particular aspect [31].

Furthermore, we only asked residents and not teachers, representatives of other health care professions or patients to participate in our workshop. We wanted to give residents the opportunity to exclusively give their point of view without being influenced by others. Other studies have shown that clinical faculty and residents at least in family medicine have a shared view of the ideal clinical teacher. The believed ideal clinical teacher for residents and faculty was stimulating, encouraging, competent and communicative; and not conventional, cautious or controlling, a finding well reflected in the results of our workshop group A. They only differed in the descriptors "Probing" and "Innovative", both more important for faculty than for residents, thus differences in views might be minimal [32].

The next step to develop quality residency in German anesthesiology could be a nationwide Delphi process involving at least residents and resident teachers to accomplish a broader consensus and in order to involve and motivate people that are involved in everyday teaching in the hospitals. Delphi processes are shown to be effective and recommended in other curriculum development processes [14], [18], [19].

The results of our workshop were presented in a scientific session to senior faculty staff during the congress in which the workshop took place: the reception of results and the following discussion showed a general agreement with the presented findings among the auditorium and conveyed their encouragement to proceed in seeking consensus from a larger group of stakeholders. It would be beneficial to involve other healthcare professionals and patients in such a process if deemed feasible.

## Conclusions

Good residency training curricula require careful consideration of all stakeholders' needs. Characteristics of good residency as defined by residents are at least to some degree easy to implement and reflect previous findings from the literature. These preliminary findings are an important stepping stone in establishing a broader consensus within the discipline.

## Notes

### Authorship

HO and WEB contributed equally to this article.

### Acknowledgements

The authors thank the German Association of Anesthesiologists (Deutsche Gesellschaft fuer Anaesthesiologie und Intensivmedizin) and the organizers of the congress (Hauptstadtkongress für Anaesthesie und Intensivmedi-

zin), especially Dr. Dominik Mittler, for the possibility to conduct the consensus process in the context of the conference. We thank Dr. Eva Kornemann, Dr. Jochen Öltjenbruns and Dr. Andreas Kopf of the Department of Anesthesiology at Charité Berlin for their support during the preparation of the workshop. We thank Dr. Max Skorning, Department of Anesthesiology, University Hospital Aachen for his presentation on the work of the group "young anesthesia" of the German Association of Anesthesiologists.

We especially thank the participants of the consensus process for their dedicated and creative work and their support during the presentation of the workshop results at the congress meeting.

We would further like to thank Stephanie Matthews of the skills lab at Charité Berlin, Ara Tekian, Department of Medical Education, University of Illinois at Chicago and two anonymous reviewers for their constructive critique of the manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Disclosure of funding

German Association of Anesthesiologists (Deutsche Gesellschaft fuer Anaesthesiologie und Intensivmedizin), Roritzer Straße 27, 90419 Nuremberg funded travel costs and conference fees for participants.

## References

- Schlitzkus LL, Schenarts KD, Schenarts PJ. Is your residency program ready for Generation Y? *J Surg Educ.* 2010 Mar-Apr;67(2):108-11. DOI: 10.1016/j.jsurg.2010.03.004
- Bürkle H. X, Y, ... Personalmangel [X, Y, ... lack of personnel?]. *Anaesthesist.* 2011 Jun;60(6):505-6. DOI: 10.1007/s00101-011-1904-1
- Schmidt GN, Fiege M, Goetz AE. Weiterbildung in der Anästhesiologie. Umsetzung am Universitätsklinikum Hamburg-Eppendorf [Further education in anesthesiology. Implementation at the University Hospital Hamburg-Eppendorf]. *Anaesthesist.* 2011 Apr;60(4):366-74. DOI: 10.1007/s00101-010-1836-1
- The Royal College of Physicians and Surgeons in Canada. CanMEDS 2005 Framework. 2005 [cited 2005 Oct 30]. Available from: <http://www.royalcollege.ca/portal/page/portal/rc/canmeds/framework>
- Leung WC. Competency based medical training: review. *BMJ.* 2002 Sep;325(7366):693-6. DOI: 10.1136/bmj.325.7366.693
- Paterson Davenport LA, Hesketh EA, Macpherson SG, Harden RM. Exit learning outcomes for the PRHO year: an evidence base for informed decisions. *Med Educ.* 2004 Jan;38(1):67-80. DOI: 10.1111/j.1365-2923.2004.01736.x
- Ringsted C, Henriksen AH, Skaarup AM, Van der Vleuten CP. Educational impact of in-training assessment (ITA) in postgraduate medical education: a qualitative study of an ITA programme in actual practice. *Med Educ.* 2004 Jul;38(7):767-77. DOI: 10.1111/j.1365-2929.2004.01841.x
- Brown AK, Roberts TE, O'connor PJ, Wakefield RJ, Karim Z, Emery P. The development of an evidence-based educational framework to facilitate the training of competent rheumatologist ultrasonographers. *Rheumatology (Oxford).* 2007 Mar;46(3):391-7. DOI: 10.1093/rheumatology/kel415
- Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT) [National Competence-Based Learning Objectives for Undergraduate Medical Education (NKLM) in Germany: Cooperation of the Association for Medical Education (GMA) and the Association of Medical Faculties in Germany (MFT)]. *GMS Z Med Ausbild.* 2009;26(3):Doc35. DOI: 10.3205/zma000627
- David DM, Euteneier A, Fischer MR, Hahn EG, Johannink J, Kulike K, Lauch R, Lindhorst E, Noll-Hussong M, Pinilla S, Weih M, Wennekes V. The future of graduate medical education in Germany – position paper of the Committee on Graduate Medical Education of the Society for Medical Education (GMA). *GMS Z Med Ausbild.* 2013;30(2):Doc26. DOI: 10.3205/zma000869
- Eidgenössische Technische Hochschule Zürich (ETH Zürich) – Institute for Environmental Decisions (IED) Consumer Behavior; Bundesärztekammer. Ergebnisse der Evaluation der Weiterbildung – 1. Befragungsrunde 2009. Bundesrapport. 2010 [cited 2010 5 Nov]. Available from: [http://www.bundesaeztekammer.de/downloads/EVA\\_Bundesrapport\\_final\\_16042010.pdf](http://www.bundesaeztekammer.de/downloads/EVA_Bundesrapport_final_16042010.pdf)
- Hibbeler B, Korzilius H. Evaluation der Weiterbildung: Ein erster Schritt. *Dtsch Arztebl.* 2010;107(10):A417–20.
- Korzilius H. Evaluation der Weiterbildung: Im Ergebnis eine gute Zwei minus. *Dtsch Arztebl.* 2011;108(50):B2250-1.
- Kern DE, Thomas PA, Hughes MT. Curriculum Development for Medical Education. A Six-Step Approach. 2nd ed. Baltimore: Johns Hopkins University Press; 2009.
- Calder J. Survey research methods. *Med Educ.* 1998 Nov;32(6):638-52. DOI: 10.1046/j.1365-2923.1998.00227.x
- Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften – Ärztliches Zentrum für Qualität in der Medizin. Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. 2005 [cited 2006 Jul 11]. Available from: <http://www.leitlinien.de/leitlinienqualitaet/leitlinie/delbi/pdf/delbi05.pdf>
- Delbecq AL, Van de Ven AH. A Group Process Model for Problem Identification and Program Planning. *J Appl Behav Sci.* 1975;7(4):466-92. DOI: 10.1177/002188637100700404
- Lloyd-Jones G, Fowell S, Bligh JG. The use of the nominal group technique as an evaluative tool in medical undergraduate education. *Med Educ.* 1999 Jan;33(1):8-13. DOI: 10.1046/j.1365-2923.1999.00288.x
- Kiessling C, Dieterich A, Fabry G, Hölzer H, Langewitz W, Mühlhans I, Pruskil S, Scheffer S, Schubert S; Committee Communication and Social Competencies of the Association for Medical Education Gesellschaft für Medizinische Ausbildung; Basel Workshop Participants. Communication and social competencies in medical education in German-speaking countries: the Basel consensus statement. Results of a Delphi survey. *Patient Educ Couns.* 2010 Nov;81(2):259-66. DOI: 10.1016/j.pec.2010.01.017
- Goldmann K, Steinfeldt T, Wulf H. Die Weiterbildung für Anästhesiologie an deutschen Universitätskliniken aus Sicht der Ausbilder – Ergebnisse einer bundesweiten Umfrage [Anaesthesia education at german university hospitals: the teachers' perspective – results of a nationwide survey]. *Anesthesiol Intensivmed Notfallmed Schmerzther.* 2006 Apr;41(4):204-12. DOI: 10.1055/s-2006-925367

21. Lehmann KA, Schultz JH. Zur Lage der anästhesiologischen Weiter- und Fortbildung in Deutschland. Ergebnisse einer Repräsentativumfrage [Anesthesiology education and training in Germany. Results from a representative questionnaire]. *Anaesthesist*. 2001 Apr;50(4):248-61. DOI: 10.1007/s001010170028
22. Prien T, Siebolds M. Beurteilung der Facharztweiterbildung durch Ärzte in Weiterbildung anhand eines validierten Fragebogens [Evaluation of medical residency training programmes by residents using a validated questionnaire]. *Anaesth Intensivmed*. 2004;45(1):25-31.
23. Radtke RM, Hahnenkamp K. Weiterbildung im klinischen Alltag: Bestandsaufnahme und Strategien [Further education during the daily clinical routine: Stocktaking and strategies]. *Anaesth Intensivmed*. 2007;48(5):240-50.
24. Hahnenkamp K, Ertmer C, Van Aken H, Skorning M. Praxis der ärztlichen Weiterbildung und Rahmenbedingen im Wandel – Evaluation der Weiterbildung im Fachgebiet Anästhesiologie über einen Zeitraum von fünf Jahren (2006–2011) [Postgraduate education in a changing occupational environment – Evaluation of postgraduate education in anaesthesiology over a period of five years (2006–2011)]. *Anaesth Intensivmed*. 2012;53(9):452-69.
25. Frank JR, Danoff D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach*. 2007 Sep;29(7):642-7. DOI: 10.1080/01421590701746983
26. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach*. 2007 Sep;29(7):648-54. DOI: 10.1080/01421590701392903
27. Shumway JM, Harden RM; Association for Medical Education in Europe. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003 Nov;25(6):569-84. DOI: 10.1080/0142159032000151907
28. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*. 2007 Nov;29(9):855-71. DOI: 10.1080/01421590701775453
29. Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. *BMJ*. 2003 Mar;326(7388):546-8. DOI: 10.1136/bmj.326.7388.546
30. Wilkinson TJ, Wade WB, Knock LD. A blueprint to assess professionalism: results of a systematic review. *Acad Med*. 2009 May;84(5):551-8. DOI: 10.1097/ACM.0b013e31819fbba2
31. van der Horst K, Siegrist M, Orlov P, Giger M. Residents' reasons for specialty choice: influence of gender, time, patient and career. *Med Educ*. 2010 Jun;44(6):595-602. DOI: 10.1111/j.1365-2923.2010.03631.x
32. Masunaga H, Hitchcock MA. Residents' and faculty's beliefs about the ideal clinical teacher. *Fam Med*. 2010 Feb;42(2):116-20.
33. Kilminster S, Cottrell D, Grant J, Jolly B. AMEE Guide No. 27: Effective educational and clinical supervision. *Med Teach*. 2007 Feb;29(1):2-19. DOI: 10.1080/01421590701210907
34. Davis N, Davis D, Bloch R. Continuing medical education: AMEE Education Guide No 35. *Med Teach*. 2008;30(7):652-66. DOI: 10.1080/01421590802108323
35. Hammick M, Olckers L, Campion-Smith C. Learning in interprofessional teams: AMEE Guide no 38. *Med Teach*. 2009 Jan;31(1):1-12. DOI: 10.1080/01421590802585561
36. Larsson J, Rosenqvist U, Holmström I. Being a young and inexperienced trainee anesthetist: a phenomenological study on tough working conditions. *Acta Anaesthesiol Scand*. 2006 Jul;50(6):653-8. DOI: 10.1111/j.1399-6576.2006.01035.x
37. Prideaux D. ABC of learning and teaching in medicine. Curriculum design. *BMJ*. 2003 Feb;326(7383):268-70.
38. Ortwein H, Dirkmorfeld L, Haase U, Herold KF, Marz S, Rehberg-Klug B, Scheid A, VargasHein O, Spies C. Zielorientierte Ausbildung als Steuerungsinstrument für die Facharztweiterbildung in der Anästhesiologie [Competency-based education as a steering instrument for residency training in anaesthesiology]. *Anaesth Intensivmed*. 2007;48(7):420-9.

#### Corresponding author:

Dr. Heiderose Ortwein

Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Charité – Universitätsmedizin Berlin, Campus Charité Mitte, Charité Platz 1, 13353 Berlin, Phone: 030-450-631197, Fax: 030-450-531911  
heiderose.ortwein@charite.de

#### Please cite as

Ortwein H, Blaum WE, Spies CD. Anesthesiology residents' perspective about good teaching – a qualitative needs assessment. *GMS Ger Med Sci*. 2014;12:Doc05.  
DOI: 10.3205/000190, URN: urn:nbn:de:0183-0001909

#### This article is freely available from

<http://www.egms.de/en/journals/gms/2014-12/000190.shtml>

Received: 2013-08-03

Revised: 2014-01-10

Published: 2014-02-19

#### Copyright

©2014 Ortwein 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.

# Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie

## Zusammenfassung

**Zielsetzung:** Die Bedeutung des Erwerbs praktischer Fertigkeiten im Medizinstudium nimmt zu. Mit dem Konsensusstatement „praktische Fertigkeiten“ hat die GMA im Rahmen der Entwicklung des Nationalen Kompetenzbasierten Lernzielkatalogs Medizin (NKLM) eine Referenz für deren Vermittlung entwickelt, die 290 Lernziele nach Organsystem, Typ (Kern- oder Wahlziel), Reifezeitpunkt der Studierenden und Lernzieltiefe gliedert. Durch den Abgleich eines umfangreichen und gut evaluierten studentischen Tutorienangebots mit dem Konsensusstatement soll einerseits der praktische Nutzen des Konsensusstatements analysiert und andererseits das Tutorienangebot selbst auf Vollständigkeit hin untersucht werden.

**Methodik:** Vier Bewerter/-innen haben in einem ersten Schritt in zwei Gruppen alle Lernziele des Konsensusstatements durch eines der 48 angebotenen Tutorien unabhängig voneinander klassifiziert. Zwischen den Bewertern/-innen einer Gruppe wurde die Interrater-Reliabilität insgesamt und in Abhängigkeit vom Organsystem berechnet. In einem zweiten Schritt wurde Dissens in der Klassifikation durch Diskussion und Konsensfindung gelöst. Anschließend wurde die Abdeckung der Lernziele in der geforderten Tiefe durch das Tutorienangebot getrennt nach Lernzieltyp und Organsystem analysiert. Gründe für den initialen Dissens wurden protokolliert und thematisch gruppiert.

**Ergebnisse:** Die Klassifikationen der beiden Bewerter/-innen korrelieren signifikant in moderater Stärke, die Stärke der Korrelation variiert in Abhängigkeit vom Organsystem und damit auch der Formulierung der einzelnen Lernziele. Nach Konsentierung ergab sich folgendes Bild: 66% aller Lernziele und 74% der Kernziele wurden durch das Tutorienangebot abdeckt. Der Grad der Abdeckung unterschied sich abhängig von Organsystem und Reifezeitpunkten.

**Schlussfolgerung:** Das Konsensusstatement ist geeignet, ein Unterrichtsangebot systematisch zu analysieren und weiterzuentwickeln. Der Abgleich mit etablierten Curricula eröffnet darüber hinaus Möglichkeiten für die Weiterentwicklung des Konsensusstatements und damit des NKLMs.

**Schlüsselwörter:** Fertigkeiten, Praktische Fertigkeiten, Klinische Fertigkeiten, medizinische Ausbildung, peer-teaching, curriculum, curricular mapping, Lernziele

## Einleitung

Die Bedeutung des Erwerbs praktischer (ärztlicher) Fertigkeiten während des Medizinstudiums hat in den letzten Jahren zugenommen [1], [2]. Gleichzeitig orientieren sich medizinische Curricula zunehmend an Outcomedefinitionen und/oder (nationalen) Lernzielkatalogen [3], [4]. In Deutschland arbeitet die Gesellschaft für Medizinische Ausbildung (GMA) gemeinsam mit dem Medizinischen Fakultätentag (MFT) seit 2009 an der Entwicklung eines Nationalen Kompetenzbasierten Lernzielkatalogs Medizin

(NKLM) [5]. Im Rahmen dieser Entwicklung des NKLM ist kürzlich das Konsensusstatement „praktische Fertigkeiten im Medizinstudium“ [6] publiziert worden, das 289 Lernziele, gegliedert in Kern- und Wahlziele sowie in 16 Organsysteme umfasst. Das Konsensusstatement „soll einen formativen Effekt auf die Fakultäten haben, ihre praktischen Unterrichtsinhalte entsprechend der Leitlinien auszurichten“ [6].

Die vorliegende Validierung untersucht das Konsensusstatement als nationale Referenz auf seine Eignung, ein

Wolf E. Blaum<sup>1,2</sup>

Katja A. Dannenberg<sup>2</sup>

Torsten Friedrich<sup>2</sup>

Anne Jarczewski<sup>2</sup>

Anne-Katrin Reinsch<sup>2</sup>

Olaf Ahlers<sup>1,2</sup>

1 Charité - Universitätsmedizin  
Berlin, Klinik für  
Anästhesiologie mit  
Schwerpunkt operative  
Intensivmedizin, Berlin,  
Deutschland

2 Charité - Universitätsmedizin  
Berlin, Abteilung für  
Curriculumsorganisation,  
Lernzentrum, Berlin,  
Deutschland

umfangreiches Curriculum – hier das Tutorienangebot des Lernzentrums der Charité – zu „kartieren“. „Curriculumskartierung“ („curriculum mapping“) beinhaltet u.a. die transparente, feingranulare Abbildung von Ausbildungsinhalten und -zielen sowie ihrer Zusammenhänge und thematischen Zuordnungen untereinander. Sie ist geeignet, Aspekte wie Gliederung und Vollständigkeit, Relevanz, Komplexität, Stimmigkeit und Organisation des Curriculums darstellbar und nachvollziehbar zu machen und wird von der Association for Medical Education in Europe AMEE empfohlen (vgl. [7]). Eine Kartierung jedes (medizinischen) Curriculums ist dabei aus drei Gründen essentiell:

1. Die medizinischen Fakultäten sind der Gesellschaft und dem Gesetzgeber Rechenschaft über die Eignung der Absolventen zum Arztberuf schuldig (§41 ÄAPPO) [8].
2. Zielklarheit und Zielverbundenheit fördern Zufriedenheit und Leistung der Studierenden [9], [10], [11], [12].
3. Die Orientierung von Lehrenden und Studierenden über die inhaltliche Einbettung ihres Unterrichts in das Curriculum steigert die Unterrichtsqualität und die Prüfungsleistung der Studierenden [13].

Das Lernzentrum der Charité, das aus dem ehemaligen „Trainingszentrum für ärztliche Fertigkeiten“ hervorgegangen ist, betreibt ein umfangreiches Programm an peer-teaching Tutorien zur Vermittlung praktischer Fertigkeiten [14]. Dazu beschäftigt das Lernzentrum zurzeit 19 studentische Tutoren/-innen, die regelmäßig medizinisch-fachlich und didaktisch geschult werden. Die Tutoren/-innen ermitteln den Bedarf an Tutorien eigenständig, ebenso konzipieren sie neue Tutorien - unter Einbindung von Experten/-innen. Jährlich werden rund 500 Termine zu einem von aktuell 48 Themen angeboten. Das Angebot wird jährlich von etwa 4500 Studierenden wahrgenommen, die Teilnahme ist freiwillig und kostenlos. Alle Tutorien werden systematisch durch den Evaluationsbereich der Charité evaluiert. Die Studierenden sind sehr zufrieden mit Inhalt und Umfang der Tutorien und ihrem Lernerfolg durch dieses Training (Median 1 auf einer siebenstufigen Likert Skala) [15]. Die Tutoren/-innen entwickeln die einzelnen Tutorien und das gesamte Angebot anhand der Evaluationsergebnisse selbstständig weiter.

## Fragestellung

Ziel der vorliegenden Arbeit ist es, das Angebot an peer-teaching Tutorien im Lernzentrum durch eine Kartierung mit dem Konsensusstatement „praktische Fertigkeiten im Medizinstudium“ abzugleichen. Der Abgleich soll drei Fragen beantworten:

1. Ist das Konsensusstatement geeignet, ein umfangreiches und sehr gut evaluiertes Tutorienangebot neu zu strukturieren und damit die „praktischen Unterrichtsinhalte entsprechend der Leitlinien auszurichten“ [6], wie es das Statement anstrebt?

2. Ist das Tutorienangebot des Lernzentrums im Sinne des Konsensusstatements vollständig/ bestehen Lücken oder ungewollte Redundanzen?
3. Ergibt sich durch den Abgleich mit dem Konsensusstatement eine sinnvolle Reihenfolge, um Tutorien zu besuchen?

## Methoden

### Abgleich von Tutorienangebot und Konsensusstatement

Das Konsensusstatement beschreibt 289 Lernziele, von denen 232 als Kernziele gekennzeichnet sind. Jedes Lernziel ist mindestens einem von 16 Organsystemen zugeordnet. Das Lernziel 275 ist dem Organsystem Wachstum/Altern, Notfall zugeordnet. Für diese Arbeit wurde dieses Lernziel doppelt aufgenommen und den beiden Organsystemen „Wachstum/Altern“ und „Kompetenzbereich Notfall“ je einmal zugeordnet, so dass sich insgesamt 290 Lernziele ergeben.

Für jedes Lernziel ist weiterhin für einen von drei Zeitpunkten die zu erreichende Tiefe auf einer von drei Stufen definiert. Die Zeitpunkte umfassen Famulaturreife (bis zum Beginn der ersten Famulatur), PJ-Reife (bis zum Beginn des praktischen Jahres) und Weiterbildungsreife (bis zum Beginn der Weiterbildung), die Tiefen sind beschrieben als

1. „demonstriert bekommen haben“,
2. „unter Aufsicht durchgeführt haben“ und
3. „routiniert handwerklich können“ (vgl. [6]).

Um die Lernziele des Konsensusstatements mit den Tutorien abgleichen zu können, wurde klassifiziert, in welcher Tiefe ein Tutorium ein Lernziel behandelt. Die Klassifikation erfolgt auf der im Konsensusstatement verwendeten Tiefenskala, die für diesen Zweck leicht modifiziert wurde und in Tabelle 1 dargestellt ist.

Um für 48 Tutorien abzubilden, in welcher Tiefe eines von 290 Lernzielen darin behandelt wird, sind 13920 (48\*290) Klassifikationen nötig.

### Klassifikation durch Bewerter

Vier studentische Tutoren/-innen mit langjähriger Erfahrung als peer teacher des Lernzentrums wurden in zwei Bewertergruppen geteilt:

Die Autorinnen AKR und AJ haben die Behandlung der 129 Lernziele zu den Organsystemen Atmung, Blut/Abwehr, GI-Trakt, Herz-Kreislauf, Nervensystem, Psyche, Sinnesorgane und Wachstum/Altern in den 48 angebotenen Tutorien klassifiziert, KAD und TF die Behandlung der 161 Lernziele zu den Organsystemen Bewegungsapparat, endokrines System, Grenzbereich Kommunikation, Grenzbereich Notfall, Grenzbereich soft skills, Harn-/Geschlechtsorgane, Haut und Organsystem übergreifende Fertigkeiten in ebenfalls allen 48 Tutorien.

**Tabelle 1: Tiefenskala zur Klassifikation der Tiefe, in der ein Tutorium ein Lernziel behandelt. Adaptiert nach [6].**

Tiefe	Bedeutung zur Klassifikation	Bedeutung im Konsensusstatement
0	Das Lernziel wird im Tutorium nicht behandelt.	keine
1	Die im Lernziel beschriebene Fertigkeit wird im Tutorium inklusive der theoretischen Voraussetzungen demonstriert.	demonstriert bekommen haben, inklusive der theoretischen Voraussetzungen
2	Die im Lernziel beschriebene Fertigkeit wird im Tutorium unter Aufsicht durchgeführt, ggf. am Modell.	unter Aufsicht durchgeführt haben, wenigstens einige Male
3	Die im Lernziel beschriebene Fertigkeit wird im Tutorium routiniert handwerklich eingesetzt, ggf. am Modell. Tutoriumsteilnehmer können Indikation und Konsequenzen benennen.	routiniert handwerklich können, situationsadäquat einsetzen können und die Konsequenzen kennen

Die Klassifikation haben die Bewerter/-innen in einem ersten Schritt unabhängig voneinander vorgenommen. Anschließend wurde die Interrater-Reliabilität ermittelt. In einem zweiten Schritt wurden ungleiche Klassifikationen durch Diskussion und Konsens in eine endgültige Klassifikation überführt, aus denen die prozentuale Abdeckung der nationalen Lernziele durch die aktuell angebotenen Tutorien ermittelt wurde. Dabei wurden die Gründe für den initialen Dissens protokolliert.

## Datenauswertung

In Libre Office 3 wurden die Klassifikationen gesammelt und die Abdeckung der Lernziele durch die Tutorien getrennt nach Organsystem, Kern- oder Wahlziel sowie Reifezeitpunkt der Studierenden berechnet.

Zur Berechnung der Interrater-Reliabilität wurde die Korrelation zwischen den Ratern nach Spearman mit SPSS 19 getrennt für jede Bewertergruppe berechnet. Zudem wurde die Korrelation getrennt nach Organsystem berechnet, um die Frage zu prüfen, welchen Interpretationsspielraum die Lernzielformulierungen des Konsensusstatements zulassen.

Die protokollierten Gründe für Dissens in der Klassifikation wurden von den Bewertergruppen gemeinsam thematisch gruppiert, um Ansatzpunkte für die weitere Entwicklung des Konsensusstatements zu benennen.

## Ergebnisse

### Qualität der Lernzielformulierung im Konsensusstatement

Bewertergruppe 1 hat die Behandlung von 129 Lernzielen in 48 Tutorien durch zusammen 6192 Klassifikationen bewertet und dabei eine signifikante Übereinstimmung moderater Stärke erzielt ( $r=0,66$ ).

Bewertergruppe 2 hat die Behandlung von 161 Lernzielen in 48 Tutorien durch zusammen 7728 Klassifikationen bewertet und dabei eine signifikante Übereinstimmung mäßiger Stärke erzielt ( $r=0,31$ ).

Die Übereinstimmung der Bewerter/-innen ist abhängig vom Organsystem, die Höhe der Korrelation unterscheidet sich zwischen den Organsystemen teilweise erheblich: Die höchste Übereinstimmung wurde für die 192 Klassifikationen des Organsystems Psyche erzielt ( $r=1,0$ ), die geringste Übereinstimmung für die 624 Lernziele des Organsystems Harn-/Geschlechtsorgane ( $r=0,2$ ). In Tabelle 2 sind die Anzahl der Klassifikationen und die Übereinstimmung der Bewertungen nach Organsystemen gegliedert dargestellt. Der Grad der Übereinstimmung ist, wie aus Tabelle 2 ersichtlich, unabhängig von der Anzahl der Lernziele im Konsensusstatement und damit der Zahl der Klassifikationen pro Organsystem.

### Abdeckung der Lernziele

Das aktuelle Tutorienangebot des Lernzentrums deckt 65,9% aller Lernziele und 73,7% der Kernziele des Konsensusstatements ab. Einige Lernziele werden in nahezu allen Tutorien behandelt: 42 Tutorien behandeln das Ziel „Berücksichtigung des Arbeits- und Gesundheitsschutzes in der eigenen ärztlichen Tätigkeit und Verantwortung, z.B. Nadelstichverletzung, rücken-schonendes Arbeiten, Recapping, Handschuhtragen, ...“ (Grenzbereich soft skills), 39 Tutorien behandeln das Ziel „Vorgehen bei Untersuchung an Patienten vermitteln [zu] können“ (Grenzbereich Kommunikation). Die Tiefe der im Konsensusstatement geforderten Abdeckung eines Lernziels nimmt über die Zeit zu, die Abdeckung durch das Tutorienangebot bis zu dieser Tiefe nimmt für die im Statement definierten Zeitpunkte ab: die Abdeckung der Lernziele in bis zur Famulaturreife geforderten Tiefe durch die Tutorien beträgt 63% für alle Lernziele (und 70,3% für die Kernziele bis zur Famulaturreife). Die bis zur PJ-Reife geforderte Tiefe wird zu 48,6%

**Tabelle 2: Anzahl der Klassifikationen und Übereinstimmung der Bewertungen nach Organsystem. \* Korrelation nach Spearman, \*\*signifikant auf einem Niveau von  $p < 0,01$**

Organsystem	Anzahl Klassifikationen	r*
Psyche	192	1,0**
Nervensystem	480	0,85**
Herz-Kreislauf	1104	0,72**
Sinnesorgane	1920	0,69**
Atmung	480	0,62**
Wachstum/Altern	1008	0,55**
Endokrines System	96	0,5**
GI-Trakt	624	0,48**
Blut Abwehr	480	0,41**
Grenzbereich Notfall	1680	0,37**
Organsystemübergreifende Fertigkeiten	1728	0,31**
Grenzbereich Kommunikation	672	0,29**
Bewegungsapparat	528	0,27**
Grenzbereich soft skills	1152	0,27**
Haut	624	0,23**
Harn-/Geschlechtsorgane	1104	0,2**

abgedeckt (Kernziele 53%), bis zur Weiterbildungsreife zu 39,3% (Kernziele 41,8%).

Die Abdeckung der Lernziele unterscheidet sich zudem nach Organsystem teils erheblich. So deckt das Tutorienangebot 90% der Lernziele des Organsystems Nervensystem (Kernziele 100%) aber nur 30% des Systems Atmung (Kernziele 50%) ab.

In Tabelle 3 ist die Abdeckung der Lernziele durch das Tutorienangebot getrennt nach Organsystem und Zeitpunkt dargestellt. Für alle Organsysteme nimmt die Abdeckung in der geforderten Tiefe mit Zunahme über die Reifestadien ab.

## Ansatzpunkte für die weitere Entwicklung des Konsensusstatements

Die protokollierten Gründe für einen Dissens in der Klassifikation lassen sich einer von drei Ursachen zuordnen: unscharfe Formulierung der Lernziele, Unterschiede im Umfang der Lernziele und lernzielunabhängige Ursachen. Am häufigsten sahen die Bewerber/-innen den Grund für Unterschiede in ihrer Klassifikation in unscharf formulierten Lernzielen des Konsensusstatements. Beispielsweise war den Bewertern/-innen unklar, welche Art von Zugängen im Lernziel 90 „Zugänge anlegen“ eingeschlossen sind oder was genau unter dem Lernziel 115 „Indikationsstellung und Anordnung technischer Untersuchungen“ zu verstehen ist.

Eine zweite wesentliche Ursache für Unterschiede in ihrer Klassifikation sahen die Bewerber/-innen im teils erheblichen Unterschied im Umfang der formulierten Lernziele. Während etwa das Lernziel „Aktive und passive Untersuchung der oberen Extremität inkl. Schulter-, Ellenbogen- und Handgelenken, sowie der Langfinger und Daumen (insbesondere Inspektion, Auffinden anatomischer Landmarken, Durchführung der Neutral-Null-Methode

und der Funktion für die Gelenke)“ sehr umfangsspezifisch formuliert ist, lässt beispielsweise das Lernziel 182 „Neuroradiologische Untersuchungsmethoden“ mehr Interpretation hinsichtlich Art und Umfang der zu erwerbenden Fähigkeiten zu.

Die lernzielunabhängigen Ursachen beziehen sich auf das Konsensusstatement als Ganzes. Beispielsweise war den Bewertern/-innen unklar, ob die Tatsache, dass einzelne Lernziele explizit Untersuchungen am Modell erwähnen, impliziert, dass dies für alle anderen Lernziele nicht gilt.

## Diskussion

### Konsensusstatement

Das Tutorienangebot des Lernzentrums der Charité ist seit 1999 etabliert und wird kontinuierlich weiterentwickelt. Seit seiner Entwicklung wird das Programm systematisch evaluiert, die Ergebnisse fließen ebenso wie die Erfahrungen und Eindrücke der nun siebten Tutorengeneration in das Programm ein. Der Abgleich dieses umfangreichen und etablierten Programms mit dem Konsensusstatement kann also als ein Argument für die Bewertung der Validität des Konsensusstatements dienen. Das aktuelle Tutorienangebot deckt sowohl insgesamt als auch pro Organsystem ausnahmslos einen deutlich höheren Anteil an Kernzielen ab, als es über die Gesamtheit der Lernziele des Konsensusstatements abdeckt. Die Priorisierung der Inhalte der Tutorien gleicht und unterstützt also die im Konsensusstatement getroffene Priorisierung.

Weiter nimmt die geforderte Komplexität der Lernziele des Konsensusstatements über die drei Reifezeitpunkte zu. Die Abdeckung dieser Lernziele in der geforderten

**Tabelle 3: Abdeckung der Lernziele durch das Tutorienangebot getrennt nach Organsystem und Zeitpunkt.**

Organsystem	Kernziele				Alle Ziele			
	Anzahl Lernziele	Abdeckung bis Famulaturreife	Abdeckung bis PJ-Reife	Abdeckung bis Weiterbildungsreife	Anzahl Lernziele	Abdeckung bis Famulaturreife	Abdeckung bis PJ-Reife	Abdeckung bis Weiterbildungsreife
Atmung	6	50 %	33 %	33 %	10	30 %	20 %	20 %
Bewegungsapparat	8	75 %	75 %	62 %	11	73 %	73 %	64 %
Blut Abwehr	9	78 %	44 %	33 %	10	70 %	40 %	30 %
Endokrines System	2	50 %	50 %	50 %	2	50 %	50 %	50 %
GI-Trakt	8	63 %	25 %	0 %	13	54 %	31 %	15 %
Grenzbereich Kommunikation	14	100 %	100 %	93 %	14	100 %	100 %	93 %
Grenzbereich Notfall	35	77 %	63 %	57 %	35	77 %	63 %	57 %
Grenzbereich soft skills	23	74 %	65 %	61 %	24	71 %	63 %	58 %
Harn- / Geschlechtsorgane	12	75 %	75 %	67 %	24	67 %	63 %	54 %
Haut	9	44 %	22 %	11 %	13	31 %	15 %	8 %
Herz-Kreislauf	18	50 %	22 %	11 %	23	48 %	26 %	17 %
Nervensystem	8	100 %	50 %	13 %	10	90 %	50 %	20 %
Organsystemübergreifende Fertigkeiten	31	77 %	50 %	55 %	36	72 %	64 %	53 %
Psyche	4	25 %	25 %	25 %	4	25 %	25 %	25 %
Sinnesorgane	28	79 %	36 %	14 %	40	63 %	30 %	15 %
Wachstum, Altern	17	35 %	35 %	29 %	21	33 %	33 %	29 %

Tiefe durch die Tutorien nimmt sowohl insgesamt als auch getrennt nach Organsystemen ab. Die Komplexitätsbewertungen des Konsensusstatements gleichen also denen des Tutorienangebots, was die im Statement vorgenommene Tiefenforderungen unterstützt.

Die in Tabelle 2 dargestellte, sehr unterschiedliche Übereinstimmung der Bewerter/-innen in Abhängigkeit vom Organsystem lässt vermuten, dass sich die im Konsensusstatement formulierten Lernziele hinsichtlich der Möglichkeit ihrer Zuordnung zu konkreten Lehrveranstaltungen unterscheiden. Mögliche Ursachen könnten die aus den Gründen für Dissens extrahierten Themen „unscharfe Formulierung der Lernziele“, „Unterschiede im Umfang der Lernziele“ und „lernzielunabhängige Ursachen“ sein. Eine systematische Aufarbeitung dieser Ansatzpunkte kann zu einer Weiterentwicklung des Konsensusstatements beitragen. Es wäre darüber hinaus interessant, zu untersuchen, ob die Vermittlung praktischer Fertigkeiten durch die curriculare Pflichtlehre der Fakultäten im Sinne des Konsensusstatements vollständig und frei von Redundanz ist.

## Tutorienprogramm

Mit den gewonnenen Erkenntnissen können Lücken und Überlappungen im Tutorienangebot identifiziert werden. Diese Daten werden zur Restrukturierung des Angebots eingesetzt.

Durch den Vergleich der Behandlungstiefe eines Lernziels im Tutorium mit der zu unterschiedlichen Zeitpunkten im Konsensusstatement geforderten Tiefe ist es darüber hinaus möglich, Studierenden mögliche sinnvolle Reihenfolgen für ihre Tutoriumsteilnahme zu empfehlen. Das Konsensusstatement ist also geeignet, bestehende Curricula neu zu strukturieren. Wir haben begonnen, das Curriculum der studentischen Tutorien in die bereits vorhandene Lehrveranstaltungs- und Lernzielplattform der

Charite [13] zu integrieren, die aktuell nur für die Kartierung der Pflichtlehre genutzt wird. Damit soll den Studierenden auf Grundlage der hier erhobenen Daten und Zuordnungen eine Online-Navigation zur Verfügung gestellt werden, die ihnen bei Interesse an ausgewählten Lernzielen und ggf. unter Angabe bereits besuchter Tutorien sinnvoll mögliche weitere Tutorien vorschlägt.

## Limitationen

In der vorliegenden Arbeit wird die Behandlung der Lernziele des Konsensusstatements „praktische Fertigkeiten“ [6] durch die im Lernzentrum angebotenen Tutorien in einer Behandlungstiefe durch unabhängige Bewerter/-innen klassifiziert. Die Interrater-Reliabilität wird als Korrelation nach Spearman berechnet, weil die Autoren die in Tabelle 1 dargestellte Skala als ordinal skaliert betrachten. Es existiert eine umfangreiche Diskussion in der Literatur zur Frage, ob Skalen wie die hier eingesetzte mit parametrischen statistischen Verfahren behandelt werden können (vgl. [16], [17]). Einige Autoren berechnen für solche und selbst für nominal skalierte Merkmale Intraklassen-Korrelationen (IKK) [18]. Die Berechnung von Korrelationen nach Spearman reduziert die Power der Untersuchung (und damit die Stärke der Korrelation), die Berechnung von IKKs würde stärkere Korrelationen ergeben, könnte aber den  $\alpha$ -Fehler erhöhen [16], [17]. Insofern stellen die hier berechneten Korrelationen eine konservative Schätzung der tatsächlichen Übereinstimmung der Bewerter/-innen dar.

Die erhobenen Gründe für den initialen Dissens in der Klassifikation wurden thematisch gruppiert. Die Analyse dieser Daten folgt keiner qualitativen Methodik.

Die Bewertung, in welcher Tiefe ein Lernziel des Konsensusstatements in einem der bisher 48 Tutorien behandelt wird, erfolgte durch erfahrene studentische Tutoren, weil sie die Abdeckung eines solchen Lernziels in den von ihnen unterrichteten Tutorien fundiert einschätzen können.

Dabei bleibt unklar, inwiefern diese studentische Klassifikation der Einschätzung durch approbierte Ärzte gleicht, die sich dazu allerdings zunächst detailliert praktische Erfahrungen mit dem Tutorienprogramm aneignen müssten.

## Fazit

Studentische Tutorien unterstützen das Selbststudium der Studierenden. Der Abgleich mit dem Konsensusstatement „praktische Fertigkeiten“ ist geeignet, ein bestehendes Unterrichtsangebot systematisch weiter zu entwickeln und „praktische Unterrichtsinhalte entsprechend der Leitlinien auszurichten“ [6]. Aus dem Abgleich mit etablierten Curricula ergeben sich darüber hinaus evidenzbasierte Möglichkeiten für die weitere Entwicklung des Konsensusstatements und damit des NKLM.

## Anmerkung

Die Autoren Dannenberg, Friedrich, Jarczewski und Reinsch haben zu gleichen Teilen beigetragen.

## Danksagung

Die Autoren danken Dr. Henrike Hölzer, Berlin, Dr. Claudia Kiessling, MPH, München und Dr. Kai Schnabel, MME, Bern sowie zwei anonymen Reviewern für ihre konstruktive und kollegiale Kritik am Manuskript. Besonderer Dank gilt Sebastian Schubert, Berlin für seine konstruktive Mitarbeit an Studiendesign und Manuskript. Wir möchten zudem Rudi Mörgeli für die Übersetzung des Manuskripts danken.

## Interessenkonflikt

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

WB leitet das Lernzentrum der Charité und ist Mitglied der GMA und deren Ausschüsse „Praktische Fertigkeiten“ und „Methodik der Ausbildungsforschung“. KAD, TF, AJ und AKR sind studentische Tutor/-innen des Lernzentrums und Studierende der Charité. OA leitet die Abteilung für Curriculumsorganisation und ist Mitglied der GMA sowie des GMA-Ausschusses „Methodik der Ausbildungsforschung“.

## Literatur

- Hudson JL, Tonkin AL. Clinical skills education: outcomes of relationships between junior medical students, senior peers and simulated patients. *Med Educ.* 2008;42(9):901-908. DOI: 10.1111/j.1365-2923.2008.03107.x
- Yudkowsky R, Otaki J, Lowenstein T, Riddle J, Nishigori H, Bordage G. A hypothesis-driven physical examination learning and assessment procedure for medical students: initial validity evidence. *Med Educ.* 2009;43(8):729-740. DOI: 10.1111/j.1365-2923.2009.03379.x
- Frank JR, Danoof D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach.* 2007;29(7):642-647. DOI: 10.1080/01421590701746983
- Bloch R, Bürgi H. The Swiss catalogue of learning objectives. *Med Teach.* 2002;24(2):144-150. DOI: 10.1080/01421590220120759
- Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT). *GMS Z Med Ausbild.* 2009;26(3):Doc35. DOI: 10.3205/zma000627
- Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. Konsensusstatement "Praktische Fertigkeiten im Medizinstudium" – ein Positionspapier des GMA-Ausschusses für praktische Fertigkeiten. *GMS Z Med Ausbild.* 2011;28(4):Doc58. DOI: 10.3205/zma000770
- Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach.* 2001;23(2):123-137. DOI: 10.1080/01421590120036547
- Bundesministerium für Gesundheit. Approbationsordnung für Ärzte vom 27. Juni 2002. *Bundesgesetzbl.* 2002;I(44):2417.
- Locke EA, Latham GP. A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice Hill; 1990.
- Braun OL. Berufliche Zielklarheit im Studium. In: Müller GF (Hrsg). *Lebenslanges Lernen.* Landau: Knecht; 2001. S.261-276.
- Gollwitzer PM. Zielbegriffe und -theorien in der heutigen Psychologie. In: Pawlik K (Hrsg). *Bericht über den 39. Kongress der Deutschen Gesellschaft für Psychologie in Hamburg 1994.* Göttingen: Hogrefe; 1995. S.295-300.
- Kähler R. Individuelle Einflussfaktoren auf Studienzufriedenheit und persönliche Ziele von Medizinstudierenden der Charité – Universitätsmedizin Berlin. Berlin: Dissertation an der Charité; 2010. Zugänglich unter/available from: [http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS\\_derivate\\_00000007656/Dissertation\\_Ragna\\_Kaehler.pdf](http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_00000007656/Dissertation_Ragna_Kaehler.pdf)
- Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. *Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA).* Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House, 2010, Doc10gma13.
- Trainingszentrum für Ärztliche Fertigkeiten. *Tutorienangebot.* Berlin: Charite Universitätsmedizin; 2012. Zugänglich unter/available from: <http://taef.charite.de>
- Kraft R, Haase C. Auswertung des studentischen Fragebogens zu den TÄF-Tutorien und des Trainingszentrums für ärztliche Fertigkeiten im Sommer 2011. Berlin: Charite Universitätsmedizin; 2011. Zugänglich unter/available from: <http://taef.charite.de>
- Jamieson S. Likert scales: how to (ab)use them. *Med Educ.* 2004; 38(12):1212-1218. DOI: 10.1111/j.1365-2929.2004.02012.x
- Norman G. Likert scales, levels of measurement and the "laws" of statistics. *Adv in Health Sci Educ.* 2010;15(5):625-632. DOI: 10.1007/s10459-010-9222-y

18. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. *Acad Med.* 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13

**Korrespondenzadresse:**

Wolf E. Blaum  
Charité - Universitätsmedizin Berlin, Abteilung für  
Curriculumsorganisation, Lernzentrum, Berlin,  
Deutschland  
wolf.blaum@charite.de

**Bitte zitieren als**

Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie. *GMS Z Med Ausbild.* 2012;29(4):Doc58.  
DOI: 10.3205/zma000828, URN: urn:nbn:de:0183-zma0008285

**Artikel online frei zugänglich unter**

<http://www.egms.de/en/journals/zma/2012-29/zma000828.shtml>

**Eingereicht:** 15.01.2012

**Überarbeitet:** 22.05.2012

**Angenommen:** 20.06.2012

**Veröffentlicht:** 08.08.2012

**Copyright**

©2012 Blaum 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.

# The practical use of the consensus statement on practical skills in medical school – a validation study

## Abstract

**Objective:** The importance of the acquisition of practical medical skills during medical school is increasing. With the consensus statement “Practical Skills,” developed by the GMA as part of the National Competency-Based Learning Objective Catalogue for Medicine (NKLM), a reference frame was created for the procurement of such skills. This frame consists of 290 learning objectives divided by “organ system,” type (core or elective learning objective), current stage of medical education and level of instruction.

By comparing a large and well evaluated range of student tutorials with the consensus statement, one can analyze the practical benefit of the statement, as well as evaluate the tutorial program for completeness.

**Methods:** In the first stage, four evaluators in two groups independently classified all consensus statement’s learning objectives by each of the 48 tutorials currently offered. The inter-rater reliability among the evaluators of each group was calculated both collectively, and according to each organ system. In the second stage, disagreements in the classification were resolved through discussion and consensus decision-making. The coverage of the learning objectives by the tutorials, in the required level of instruction, was then analyzed separately by learning objective type and organ system. Reasons for any initial dissent were recorded and grouped thematically.

**Results:** The correlation between the classifications of the two evaluators was moderately significant. The strength of this correlation, and thus the precision of individual learning goals wording, varied according to organ system. After a consensus was reached, the results show that the offered tutorials covered 66% of all learning objectives, as well as 74% of the core objectives. The degree of coverage differed according to organ system and stage of medical education.

**Conclusion:** The consensus statement is suitable to systematically analyze and develop teaching units. The comparison with established curricula also offers possibilities for further development of the consensus statement, and therefore also of the NKLM.

**Keywords:** skills, practical skills, clinical skills, medical education, peer-teaching, curriculum, curricular mapping, learning objectives

## Introduction

The importance of the acquisition of practical (medical) skills in medical school has increased in recent years [1], [2]. At the same time, medical curricula are increasingly relying on outcome definitions and/or (national) learning objective catalogues [3], [4]. In Germany, the Medical Education Society (GMA), in conjunction with the German Medical Faculty Association (MFT), has been working since 2009 on the development of a National Competency-Based Learning Objective Catalogue for Medicine (NKLM) [5]. As part of this project, the consensus statement “Practical Skills in Medical School” was recently

published, which comprises 289 learning objectives, divided into core/elective goals and 16 organ systems. The consensus statement “should have a formative effect on the faculties, to bring the content of their practical teachings in accordance to the guidelines” [6].

The current validation study analyzes the suitability of the consensus statement as a national reference, to “map” a comprehensive curriculum – here the tutorials offered by the Learning Center of the Charité.

“Curriculum Mapping” includes the transparent, well-defined presentation of educational content and objectives, as well as their correlation and thematic association. It renders aspects such as structure, completeness, relevance, complexity, coherence, and curricular organization in a clear and understandable manner, and is also

Wolf E. Blaum<sup>1,2</sup>

Katja A. Dannenberg<sup>2</sup>

Torsten Friedrich<sup>2</sup>

Anne Jarczewski<sup>2</sup>

Anne-Katrin Reinsch<sup>2</sup>

Olaf Ahlers<sup>1,2</sup>

1 Charité - Universitätsmedizin  
Berlin, Klinik für  
Anästhesiologie mit  
Schwerpunkt operative  
Intensivmedizin, Berlin,  
Deutschland

2 Charité - Universitätsmedizin  
Berlin, Abteilung für  
Curriculumsorganisation,  
Lernzentrum, Berlin,  
Deutschland

recommended by the Association for Medical Education in Europe AMEE (see [7]). In this process, a mapping of each (medical) curriculum is essential for three reasons:

1. The medical schools are accountable, to the public and to the legislators, for the competency of their graduates to practice medicine (§41 ÄAPPO) [8].
2. Clarity of purpose and adherence to the objectives contribute to student satisfaction and performance [9], [10], [11], [12].
3. The integration of teaching content in the curriculum provides a reference for teachers and students, thus enhancing the quality of instruction and student examination performances [13].

The Learning Center of the Charité, which emerged from the former "Training Center for Medical Skills," runs an extensive program of peer-teaching tutorials for the conveyance of medical skills [14]. The Learning Center currently employs 19 student tutors, who receive regular training in the performance of medical procedures, as well as in didactics. These tutors independently determine the demand for new tutorials, and are responsible for their development under the supervision of experts. Each year, about 500 tutorials are offered on 48 topics. These classes are visited by approximately 4 500 students, and the participation is voluntary and free of charge. All tutorials are systematically evaluated by the Evaluation Department of the Charité. The students are extremely satisfied with the content and scope of the tutorials, as well as the effectiveness of the training (median 1 on a seven-point Likert scale) [15]. Based on these evaluations, the tutors independently refine the individual tutorials, as well as the range of the courses offered.

## Research Objectives

The aim of the current study is the mapping of the Learning Center's peer-teaching tutorials, and a comparison with the consensus statement "Practical Skills in Medical School". The comparison should answer three questions:

1. Is the consensus statement suited to re-/structure a comprehensive and well evaluated range of tutorials, and therefore to "bring the content of their practical teachings in accordance to the guidelines" [6], as the statement strives to do?
2. For the purposes of the consensus statement, is the tutorial program of the Learning Center complete, or are there gaps or unintended redundancies?
3. Does the comparison with the consensus statement suggest that the tutorials should be visited in a particular order?

## Methods

### Comparison of the Tutorial Program with the Consensus Statement

The consensus statement describes 289 learning objectives, 232 of which are defined as core objectives. Each learning objective is assigned to at least one of 16 organ systems. The objective 275 is assigned to the organ system Growth/Ageing and Emergency. This learning objective was recorded twice in this study, and assigned to both "Growth/Ageing" and "Emergency Medicine" organ systems, leading to a total of 290 learning objectives.

In one of three stages of the medical education, the target level of instruction for each objective is defined as one of three levels. These stages include the beginning of the first elective clinical rotation, the beginning of the practical year, and the beginning of residency. The level of instruction is described as

1. "was given demonstration";
2. "performed under supervision"; and
3. "skilled performance" (see [6]).

In order to compare the learning objectives of the consensus statement with the tutorials, the depth at which the tutorial deals with a particular objective was classified. The classification was based on the Level of Instruction Scale of the consensus statement, which was slightly modified for this study, and is shown in Table 1.

In order to depict the level of instruction of the 290 learning objectives in the 48 tutorials, a total of 13 920 (48 x 290) classifications were needed.

### Classification by the Evaluators

Four student tutors with long term experience in peer-teaching at the Learning Center were divided into two evaluation groups:

The authors AKR and AJ classified the parameters of 129 learning objectives, in all 48 available tutorials, for the following organ systems: respiratory, blood/immunology, gastrointestinal, cardiovascular, nervous, psyche, sensory organs, and Growth/Ageing. The authors KAD and TF classified 161 learning objectives, also in all 48 tutorials, for the following organ systems: musculoskeletal, endocrine, communication, emergency medicine, soft skills, urinary/sexual, dermatological, and comprehensive skills. In the first step, the evaluators classified the objectives independently from each other, allowing for the calculation of the inter-rater reliability. In the second step, classification discrepancies were resolved through discussion and consensus decision-making, until a final classification was achieved. This classification was then used to determine the percentage of the national learning objectives that were covered by the currently available tutorials. Reasons for any initial dissent were recorded.

**Table 1: Level of Instruction Scale used to classify the level of instruction at which a learning objective was covered in the tutorials. Modified according to [6].**

Level of Instruction	Classification Key	Consensus Statement Key
0	The objective was not covered in the tutorial.	Not applicable
1	The skill described in the learning objective was demonstrated in the tutorial, including a theoretical background discussion.	Demonstration was performed, and the theoretical background was discussed
2	The skill described in the learning objective was performed under supervision during the tutorial, using model-systems where applicable.	Performance under supervision, minimum of a few trials
3	The skill described in the learning objective was routinely performed in the tutorial, using model-systems where applicable. Tutorial participants are able to list indications and ramifications.	Skilled performance, adequate use, and known ramifications

## Data Analysis

The classifications were gathered in Libre Office 3, and the coverage of the learning objectives by the tutorials was calculated separately by organ system, type of objective (core or elective), as well as current stage of medical education.

In order to calculate the inter-rater reliability, the correlation between evaluators was determined separately for each group using Spearman's correlation with SPSS19. Additionally, the correlation was calculated separately by organ system, in order to examine to which extent the consensus statement's wording for the learning objectives left a margin for interpretation.

The reasons for initial dissent among the evaluators in the classification procedure were recorded and grouped thematically, so as to identify issues and aid the further development of the consensus statement.

## Results

### Quality of Objective Formulation in the Consensus Statement

Group 1 analyzed parameters for 129 learning objectives in 48 tutorials, for a total of 6 192 classifications, and reached a moderately significant agreement between the evaluators ( $r=0.66$ ).

Group 2 analyzed parameters for 161 learning objectives in 48 tutorials, for a total of 7 728 classifications, and reached a modest, but significant agreement between the evaluators ( $r=0.31$ ).

The congruence between tutors varied in relation to organ system, and the intensity of this correlation was in some cases substantial: the highest congruence rate was reached in the 192 classifications of the "Psyche" organ system ( $r=1.0$ ), and the lowest in the 624 classifications

of the "Urinary/Sexual" organ system ( $r=0.2$ ). Table 2 shows the number of classifications and the consistency of the ratings for each organ system. As seen in Table 2, the level of congruence is independent from the number of learning objectives, and thus also the number of classifications, per organ system.

### Coverage of the Learning Objectives

The current tutorial program covers 65.9% of all learning goals, and 73.7% of the core objectives listed in the consensus statement. Some of the learning goals are covered in nearly all tutorials: 42 tutorials deal with the objective "Regard to occupational health and safety in the personal medical activities and responsibilities, per example, needle stick injury, working in an ergonomic manner, recapping, wearing gloves, ..." (Soft Skills); 39 tutorials deal with the objective "Being able to explain the procedures to the patient during an examination" (Communication).

The level of instruction to which an objective must be covered, as required by the consensus statement, increases along with stage of medical education. The coverage of these objectives by the tutorial program, to the consensus statements' required depth and at higher stages of medical education, decreases: the objectives that must be covered, to the required depth, by the time of the first elective clinical rotation amounts to 63% of all learning objectives (and 70.3% of the core objectives). The target level of instruction was achieved by the beginning of the practical year for 48.6% of the objectives (53% of the core objectives), and 39.3% of the objectives (41.8% of the core objectives) by the beginning of the continuing education.

The coverage of the learning goals varied, in some cases considerably, in relation to organ system. The tutorials offered covered 90% of the learning objectives (100% of core objectives) for the Nervous System, but only 30% of

**Table 2: Number of classifications and consistency of the ratings in relation to organ system. \*Determined via Spearman's Correlation; \*\*Significance level of  $p < 0.01$ .**

Organ System	Number of Classifications	$r^*$
Psyche	192	1.0**
Nervous	480	0.85**
Cardiovascular	1104	0.72**
Sensory Organs	1920	0.69**
Respiratory	480	0.62**
Growth/Ageing	1008	0.55**
Endocrine	96	0.5**
Gastrointestinal	624	0.48**
Blood/Immunology	480	0.41**
Emergency Medicine	1680	0.37**
Comprehensive Skills	1728	0.31**
Communication	672	0.29**
Musculoskeletal	528	0.27**
Soft Skills	1152	0.27**
Dermatological	624	0.23**
Urinary/Sexual	1104	0.2**

the goals (50% core objectives) for the Respiratory System.

Table 3 shows the coverage of learning objectives by the tutorials in relation to organ system and stage of education. For all organ systems, there is a decrease in the coverage of objectives (to the required depth) as the student level of education increases.

## Key Issues for the Further Development of the Consensus Statement

The disagreements recorded during the classification procedure were attributed to one of three causes: ambiguous formulation of learning objectives, unspecific scope of the objective, or objective-independent causes. The most frequent problem listed was vague formulation of the consensus statement. For example, the evaluators had difficulty determining what types of infusion lines were included in objective 90 "Establishing lines" (German: "Zugänge anlegen"; usually refers to starting an IV line, but can include central, arterial, intraosseous lines, etc), or how exactly to interpret objective 115, "Indications and regulations for technical examinations."

A second major cause for inconsistencies in the classification process was the vague scope of some objectives. While the learning objective "Active and passive examination of the upper extremities including shoulder, elbow and wrist, as well as fingers and thumb (especially inspection, locating anatomical landmarks, establishing range of motion, and functionality of the joints)" is extremely precise and well defined, others, such as the objective "Neuroradiological examinations methods," leave a significant gap in terms of type and extent of the skills to be acquired.

The objective-independent causes refer to the consensus statement as a whole. For example, it was unclear to the evaluators whether practice on simulators, which was specifically mentioned in some learning objectives, implied its exclusion for all other objectives.

## Discussion

### Consensus Statement

The tutorial program in the Learning Center of the Charité was established in 1999, and has been continually evolving. Since the beginning, the tutorials have been systematically evaluated, and now the results of the seventh generation of tutors demonstrate their experience and influence. The analysis of the consensus statement against this extensive and well established program can help appraise the statement's validity. The current tutorial program covers a significantly higher proportion of core learning objectives than the total list of objectives. This is true, without exceptions, for individual organ systems as well as for the consensus statement as a whole. The prioritization of the tutorial content supports, and is congruent with, the priorities set by the consensus statement. Furthermore, the complexity of the learning objectives increases along with the three set stages of medical education. The coverage of these objectives by the tutorials, at the required depth, decreases both overall and according to organ system. The level of complexity required by the consensus statement matches that of the tutorials, which supports the statement's requirements. The stated objectives differ as to the practicality of their integration in specific courses, as suggested by the inconsistency between evaluators in relation to organ system,

**Table 3: Coverage of learning objectives by the tutorial program in relation to organ system and stage of education.**

Organ System	Core Objectives				All Learning Objectives			
	Number of Objectives	Covered before Clinical Electives	Covered before Practical Year	Covered before Continuing Education	Number of Objectives	Covered before Clinical Electives	Covered before Practical Year	Covered before Continuing Education
Respiratory	6	50 %	33 %	33 %	10	30 %	20 %	20 %
Musculoskeletal	8	75 %	75 %	62 %	11	73 %	73 %	64 %
Blood/Immunology	9	78 %	44 %	33 %	10	70 %	40 %	30 %
Endocrine	2	50 %	50 %	50 %	2	50 %	50 %	50 %
Gastrointestinal	8	63 %	25 %	0 %	13	54 %	31 %	15 %
Communication	14	100 %	100 %	93 %	14	100 %	100 %	93 %
Emergency Medicine	35	77 %	63 %	57 %	35	77 %	63 %	57 %
Soft Skills	23	74 %	65 %	61 %	24	71 %	63 %	58 %
Urinary/Sexual	12	75 %	75 %	67 %	24	67 %	63 %	54 %
Dermatological	9	44 %	22 %	11 %	13	31 %	15 %	8 %
Cardiovascular	18	50 %	22 %	11 %	23	48 %	26 %	17 %
Nervous	8	100 %	50%	13 %	10	90 %	50 %	20 %
Comprehensive Skills	31	77 %	50 %	55 %	36	72 %	64 %	53 %
Psyche	4	25 %	25 %	25 %	4	25 %	25 %	25 %
Sensory Organs	28	79 %	36 %	14 %	40	63 %	30 %	15 %
Growth/Ageing	17	35 %	35 %	29 %	21	33 %	33 %	29 %

shown in table 2. The following themes were listed as possible causes: “Ambiguous formulation of learning objectives,” “Unspecific scope of the objective,” and “Objective-independent causes.” A systematic review of these issues can contribute to the further development of the consensus statement. It would be also interesting to investigate whether the conveyance of practical skills, as established by the faculty in the mandatory curriculum, are complete and free of redundancy by the standards of the consensus statement.

## Tutorial program

In light of the results, gaps and overlaps in the tutorials could be identified, and the data was further used to restructure the tutorial program.

By comparing the level of instruction of a particular learning objective in the tutorials with the level required at specific milestones by the consensus statement, it is possible to recommend to students the tutorial participation in a specific order. The consensus statement is therefore suitable for the restructure of existing curricula. We have started to integrate the tutorial curriculum in the existing courses and teaching platform of the Charité [13], which has so far only been used to chart mandatory courses. This encompasses the development of a student online navigation system, which uses the data collected from this study to suggest tutorials based on the student’s interest, or as a meaningful follow-up to the tutorials already visited.

## Limitations

In the current study, the coverage of the learning objectives of the consensus statement “Practical Skills” [6] by

the tutorials from the Learning Center of the Charité was independently classified. The inter-rater reliability was calculated using Spearman’s correlation, as the authors considered the scale shown in table 1 to be an ordinal scale. There is an extensive discussion in the literature as to whether a scale such as the one used in this study can be appraised using parametric statistics (see [16], [17]). Some authors evaluate such scales, and even nominal scales, using intraclass correlations (ICC) [18]. The use of Spearman’s correlation reduces the power of the examination (and therefore the strength of the correlation), and although the use of ICCs provide stronger correlations, it can potentially increase the  $\alpha$ -error [16], [17]. Therefore, the correlation shown here represents solely a conservative estimate of the actual evaluator congruency.

The initial reasons for dissent among the evaluators during the classification procedure were grouped thematically. No qualitative methods were regarded during the data analysis.

Experienced student tutors, as experts on their own tutorials, were deemed most apt to estimate whether the 48 tutorials cover, and to what depth, the objectives of the consensus statement. It remains unclear how the classification by students would relate to a classification by licensed physicians, though the latter would first require extensive practical experience in the tutorial program.

## Conclusion

The student-led tutorials support student independent studies. The comparison with the consensus statement “Practical Skills” is suitable to systematically analyze and develop teaching units and to “bring the content of their

practical teachings in accordance to the guidelines" [6]. The comparison with established curricula offers evidence-based possibilities for further development of the consensus statement, and therefore also of the NKLM.

## Note

The authors Dannenberg, Friedrich, Jarczewski and Reinsch contributed equally.

## Acknowledgement

The authors thank Dr. Henrike Hölzer, Berlin, Dr. Claudia Kiessling, MPH, Munich, and Dr. Kai Schnabel, MME, Bern, as well as two anonymous reviewers for their constructive and collegial critic of the manuscript. Special thanks to Sebastian Schubert, Berlin, for his constructive involvement in the study design and the manuscript. We would also like to thank Rudi Mörgeli for the translation of the manuscript.

## Competing interests

The authors declare that they have no competing interests.

WB is the head of the Learning Center of the Charité, a member of the GMA, and of the GMA committees "Practical Skills" and "Educational Research Methods." KAD, TF, AJ and AKR are student tutors of the Learning Center, and students of the Charité. OA is head of the Department of Curricular Organization, a member of the GMA, as well as a member of the GMA committee "Educational Research Methods."

## References

- Hudson JL, Tonkin AL. Clinical skills education: outcomes of relationships between junior medical students, senior peers and simulated patients. *Med Educ.* 2008;42(9):901-908. DOI: 10.1111/j.1365-2923.2008.03107.x
- Yudkowsky R, Otaki J, Lowenstein T, Riddle J, Nishigori H, Bordage G. A hypothesis-driven physical examination learning and assessment procedure for medical students: initial validity evidence. *Med Educ.* 2009;43(8):729-740. DOI: 10.1111/j.1365-2923.2009.03379.x
- Frank JR, Danof D. The CanMEDS initiative: implementing an outcomes-based framework of physician competencies. *Med Teach.* 2007;29(7):642-647. DOI: 10.1080/01421590701746983
- Bloch R, Bürgi H. The Swiss catalogue of learning objectives. *Med Teach.* 2002;24(2):144-150. DOI: 10.1080/01421590220120759
- Hahn EG, Fischer MR. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM) für Deutschland: Zusammenarbeit der Gesellschaft für Medizinische Ausbildung (GMA) und des Medizinischen Fakultätentages (MFT). *GMS Z Med Ausbild.* 2009;26(3):Doc35. DOI: 10.3205/zma000627
- Schnabel KP, Boldt PD, Breuer G, Fichtner A, Karsten G, Kujumdshiev S, Schmidts M, Stosch C. Konsensusstatement "Praktische Fertigkeiten im Medizinstudium" – ein Positionspapier des GMA-Ausschusses für praktische Fertigkeiten. *GMS Z Med Ausbild.* 2011;28(4):Doc58. DOI: 10.3205/zma000770
- Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach.* 2001;23(2):123-137. DOI: 10.1080/01421590120036547
- Bundesministerium für Gesundheit. Approbationsordnung für Ärzte vom 27. Juni 2002. *Bundesgesetzbl.* 2002;I(44):2417.
- Locke EA, Latham GP. A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice Hill; 1990.
- Braun OL. Berufliche Zielklarheit im Studium. In: Müller GF (Hrsg). *Lebenslanges Lernen.* Landau: Knecht; 2001. S.261-276.
- Göllwitzer PM. Zielbegriffe und -theorien in der heutigen Psychologie. In: Pawlik K (Hrsg). *Bericht über den 39. Kongress der Deutschen Gesellschaft für Psychologie in Hamburg 1994.* Göttingen: Hogrefe; 1995. S.295-300.
- Kähler R. Individuelle Einflussfaktoren auf Studienzufriedenheit und persönliche Ziele von Medizinstudierenden der Charité – Universitätsmedizin Berlin. Berlin: Dissertation an der Charité; 2010. Zugänglich unter/available from: [http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS\\_derivate\\_00000007656/Dissertation\\_Ragna\\_Kaehler.pdf](http://www.diss.fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_00000007656/Dissertation_Ragna_Kaehler.pdf)
- Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. *Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA).* Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House, 2010, Doc10gma13.
- Trainingszentrum für Ärztliche Fertigkeiten. Tutorienangebot. Berlin: Charite Universitätsmedizin; 2012. Zugänglich unter/available from: <http://taef.charite.de>
- Kraft R, Haase C. Auswertung des studentischen Fragebogens zu den TÄF-Tutorien und des Trainingszentrums für ärztliche Fertigkeiten im Sommer 2011. Berlin: Charite Universitätsmedizin; 2011. Zugänglich unter/available from: <http://taef.charite.de>
- Jamieson S. Likert scales: how to (ab)use them. *Med Educ.* 2004; 38(12):1212-1218. DOI: 10.1111/j.1365-2929.2004.02012.x
- Norman G. Likert scales, levels of measurement and the "laws" of statistics. *Adv in Health Sci Educ.* 2010;15(5):625-632. DOI: 10.1007/s10459-010-9222-y
- Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. *Acad Med.* 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13

### Corresponding author:

Wolf E. Blaum  
Charité - Universitätsmedizin Berlin, Abteilung für  
Curriculumsorganisation, Lernzentrum, Berlin,  
Deutschland  
[wolf.blaum@charite.de](mailto:wolf.blaum@charite.de)

### Please cite as

Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie. *GMS Z Med Ausbild.* 2012;29(4):Doc58. DOI: 10.3205/zma000828, URN: <urn:nbn:de:0183-zma0008285>

**This article is freely available from**

<http://www.egms.de/en/journals/zma/2012-29/zma000828.shtml>

**Received:** 2012-01-15

**Revised:** 2012-05-22

**Accepted:** 2012-06-20

**Published:** 2012-08-08

**Copyright**

©2012 Blaum 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.

# Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche

## Zusammenfassung

**Einleitung:** Sowohl für die Curriculumsentwicklung und -kartierung als auch zur Orientierung innerhalb der seit Jahren zunehmenden Fülle von Lernressourcen in der medizinischen Ausbildung stellt das semantische Web („Web 3.0“) ein niedrigschwelliges, effektives Hilfsmittel dar, das es ermöglicht, inhaltlich verwandte Elemente über Systemgrenzen hinweg zu identifizieren. Voraussetzung dafür ist der Einsatz eines geeigneten strukturierten Vokabulars zur maschinenlesbaren, inhaltlichen Beschreibung von Objekten, um die bisher notwendige manuelle Verknüpfung durch eine automatisch erzeugte, inhaltsbasierte bzw. semantische Verknüpfung zu ersetzen.

Ziel dieser Arbeit ist es, existierende Taxonomien und Ontologien zur Annotation medizinischer Lernressourcen und -inhalte zusammenzustellen und anhand ausgewählter Kriterien zu vergleichen sowie auf ihre Eignung zum Einsatz im eingangs beschriebenen Kontext zu überprüfen.

**Methoden:** Anhand einer systematischen Literaturrecherche wurden existierende Taxonomien und Ontologien zur Beschreibung medizinischer Lernressourcen identifiziert. Für jedes so identifizierte strukturierte Vokabular wurden mittels Websuche und/oder Kontakt zu den Herausgebern Thema, Struktur, Sprache, Umfang, Wartung und Technik der Taxonomie/Ontologie ermittelt und deren Eignung für den Einsatz im semantischen Web überprüft.

**Ergebnisse:** In 20 identifizierten Publikationen wurden 14 strukturierte Vokabulare identifiziert, die sich in Sprache, Umfang, Aktualität und Wartung zum Teil sehr stark unterscheiden.

Keines der identifizierten Vokabulare erfüllte die erforderlichen Kriterien zur inhaltlichen Beschreibung medizinischer Ausbildungsinhalte und Lernressourcen im deutschsprachigen Raum.

**Diskussion:** Auf dem Weg zum Web 3.0 stellen Auswahl und Einsatz eines geeigneten, deutschsprachigen Vokabulars zur maschinenlesbaren, inhaltlichen Beschreibung von Objekten ein relevantes Problem dar. Mögliche Lösungsansätze umfassen die Neuentwicklung, die Übersetzung und/ oder die Kombination vorhandener Vokabulare, ggf. mit einer teilweisen Übersetzung englischsprachiger Vokabulare.

**Schlüsselwörter:** medizinische Ausbildung, semantisches Web, Web 3.0, Taxonomie, Ontologie, curriculares Mapping, Curriculumskartierung, e-learning, neue Medien

**Wolf E. Blaum**<sup>1,2</sup>

**Anne Jarczewski**<sup>2</sup>

**Felix Balzer**<sup>1,2</sup>

**Philip Stötzner**<sup>2</sup>

**Olaf Ahlers**<sup>1,2</sup>

1 Charité - Universitätsmedizin Berlin, Campus Charité Mitte und Campus Virchow-Kliniken, Klinik für Anästhesiologie mit Schwerpunkt operative Intensivmedizin, Berlin, Deutschland

2 Charité - Universitätsmedizin Berlin, Abteilung für Curriculumsorganisation, Lernzentrum, Berlin, Deutschland

## Autorenschaft

Die Autoren Blaum und Jarczewski haben zu gleichen Teilen beigetragen.

## Einleitung

Die seit Jahren zunehmende Fülle von insbesondere online verfügbaren Lernressourcen in der medizinischen Ausbildung kann derzeit kaum sinnvoll genutzt werden,

weil es für Lehrende und Lernende an einem systematischen und einfachen Zugang zu deren Inhalten fehlt [1]. Die stetig steigende Zahl von verfügbaren Quellen stellt Studierende zunehmend vor die Schwierigkeit, thematisch passende, qualitativ hochwertige und didaktisch sinnvolle Angebote, welche zudem ihrem Ausbildungsniveau entsprechen, in einem Meer von Möglichkeiten zu identifizieren.

Eine beachtliche Zunahme von Ressourcen erlebt dabei insbesondere das computerbasierte Lernen. In einem kürzlich erschienenen Review zum Effekt von computer-

basierten virtuellen Patienten – einer Teilmenge der computerbasierten Ressourcen – haben Cook et al. allein zu diesem Thema 698 Publikationen identifiziert [2]. Gleichzeitig wird das PC-basierte Lernen mit der Immatrikulation der „Generation Y“ zunehmend wichtiger [3]. Lehrende und Planende stehen darüber hinaus vor der Herausforderung, die inhaltlichen Zusammenhänge in den von ihnen verantworteten Curricula darzustellen (zu kartieren) [4], was bisher eine manuelle Verknüpfung von curricularen Elementen voraussetzt.

## Semantisches Web

Besonders die Entwicklung des Internets hin zum semantischen Web (Web 3.0) birgt ein erhebliches Potential zur inhaltlichen Verknüpfung von Lernressourcen ebenso wie zum Einsatz in der Curriculumskartierung [5], [6].

Während die Inhalte des Internets in seiner Entstehung im Wesentlichen bestimmt waren durch von wenigen Autoren bereit gestellte Informationen für ein breites, konsumierendes Publikum, ist das aktuelle „Web 2.0“ gekennzeichnet durch vom Nutzer erzeugte Inhalte [7]. Den Daten beider Versionen ist gemein, dass die darin enthaltenen Informationen ausschließlich von Menschen verstanden und interpretiert werden können. Jede inhaltliche Verknüpfung der Daten setzt demnach eine menschliche Aktion voraus.

Der Begriff „Web 3.0“ (auch „semantic Web“ genannt) bezeichnet ein Konzept zur Speicherung von Daten im Netz, welches durch Hinzufügen von maschinenlesbaren Metainformationen (z. B. mittels strukturierter Vokabulare) die Möglichkeit bietet, inhaltliche Verknüpfungen zwischen Datenobjekten von Maschinen automatisch generieren bzw. aufdecken zu lassen [8]. Voraussetzung dafür ist eine maschinell lesbare, inhaltliche Beschreibung der in den Daten enthaltenen Informationen [6], [9].

Eine geeignete Art der inhaltlichen Beschreibung ist der Einsatz strukturierter Vokabulare. Dabei wird jede Ressource mit definierten Begriffen eines strukturierten Vokabulars verschlagwortet.

## Taxonomie und Ontologie

Taxonomien und Ontologien stellen formalisierte Vokabulare zur inhaltlichen Beschreibung (semantische Annotation) von Objekten, wie zum Beispiel Lernressourcen, dar [10]. Durch die darüber hinaus in Taxonomien hinterlegten Beziehungen der Vokabeln untereinander kann eine (manuell angestoßene) automatische Suche Lernressourcen mit inhaltlich ähnlichen Beschreibungen auch an voneinander unabhängigen Orten identifizieren [1], [11], [12].

Die Beziehungen der Vokabeln einer Taxonomie sind in der Regel monohierarchisch, jeder Vokabel sind also exakt eine (oder keine) übergeordnete Vokabel („ist ein“- oder „ist Teil von“-Beziehung) und beliebig viele nachgeordnete Vokabeln zugeordnet. Ein klassisches Beispiel einer streng monohierarchischen Taxonomie ist die International Classification of Diseases <http://www.dimdi.de/>

<http://www.dimdi.de/static/de/klassi/index.htm>. Ontologien hingegen bilden auch polyhierarchische Strukturen mit der Zuordnung einer Vokabel zu mehr als einer übergeordneten Vokabel ab. Grundlegender Unterschied zu Taxonomien ist dabei, dass keine sprachlichen Konstrukte, sondern logisch definierte Formalismen dafür verwandt werden, Gegenstände und Zusammenhänge der Realwelt zu definieren. Durch diese Unabhängigkeit von sprachlichen Eigenschaften stellen Ontologien ein solides Instrumentarium für semantische Beschreibungen dar und können so eine große Hilfe in der Standardisierung von Begrifflichkeiten bieten [13].

Ein bekanntes Beispiel einer polyhierarchischen Ontologie mit Äquivalenzbeziehungen sind die Medical Subject Headings <http://www.nlm.nih.gov/mesh/>, die zur Beschreibung von in Medline indizierten Publikationen verwendet werden. Abbildung 1 verdeutlicht die Unterschiede zwischen Taxonomie und Ontologie.

Taxonomien und Ontologien sind prinzipiell geeignet, inhaltliche Zusammenhänge auch zwischen Objekten herzustellen, die in voneinander unabhängigen Systemen existieren. So können sie beispielsweise helfen, notwendige aber aufwendige manuelle Verknüpfungen zwischen verschiedenen Outcomeframeworks [14] oder Lernzielkatalogen [15], wie sie unter anderem in einer Minderheit der Ansätze zur Curriculumskartierung genutzt werden [5], [16], zumindest entscheidend vorzubereiten, wenn nicht sogar zu automatisieren.

In der Vorbereitung auf die hier dargestellte Arbeit konnten die Autoren keine allgemein akzeptierten Kriterien zur Beschreibung und zum Vergleich strukturierter Vokabulare identifizieren.

## Fragestellung

Bei der Auswahl einer geeigneten Taxonomie bzw. Ontologie sehen sich medizinische Fakultäten einer Vielzahl von unterschiedlich umfangreichen, gewarteten und verfügbaren Angeboten gegenüber. Ziel dieser Arbeit ist es, existierende Taxonomien und Ontologien zur Annotation medizinischer Lernressourcen und -inhalte zusammenzustellen, anhand hierfür entwickelter Kriterien zu vergleichen und auf ihre Eignung für den Einsatz im semantischen Web zu untersuchen.

## Methoden

Im ersten Schritt wurde eine Literaturrecherche in den Datenbanken der GMS (German Medical Sciences) <http://www.egms.de/dynamic/de/index.htm> im Juni 2011 durchgeführt, um Publikationen zu identifizieren, die entweder mindestens eine Taxonomie oder Ontologie für die medizinische Ausbildung oder deren dortigen Einsatz beschreiben. Die Suchbegriffe sind in Tabelle 1 links zusammengestellt.

Bei dieser Recherche konnten fünf strukturierte Vokabulare identifiziert werden.

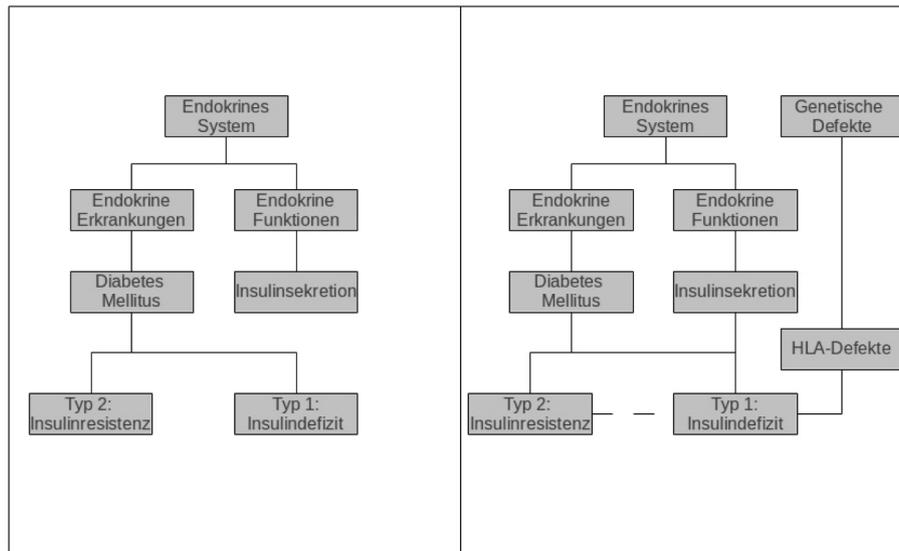


Abbildung 1: In einer Taxonomie (links) ist jedem Element genau ein übergeordnetes Element zugeordnet (Zuordnungen als durchgezogene Linien). In einer Ontologie (rechts) kann ein Element mehrere übergeordnete Elemente haben und zudem mit anderen Elementen assoziiert (gestrichelte Linie) sein.

Tabelle 1: Suchbegriffe der systematischen Literaturrecherche (Zahlenangaben: zur Analyse verwendete Publikationen/Gesamtzahl der Treffer)

Deutsch – vor allem GMS	5/37	English	15/564
medizinische Taxonomie	2/9	Medical education ontology	6/37
medizinische Ontologie	3/12	Taxonomy medical education german	0/93
deutschsprachige Taxonomie	0/2	Taxonomy medical education elearning	5/238
Taxonomie medizinische Lehre	0/2	taxonomy learning resources	1/127
Ontologie medizinische Lehre	0/6	medical education semantic network	2/15
Taxonomie medizinische Lernressourcen	0/3	medical education curriculum mapping taxonomy	0/18
Taxonomie medizinische Lerninhalte	0/3	taxonomy medical education learning objectives	1/36

Die Autoren extrahierten aus diesen Publikationen im Konsens Kriterien zur Beurteilung der Eignung eines Vokabulars für die Annotation von Elementen der medizinischen Ausbildung.

Diese extrahierten Kriterien umfassen:

- **Thema des Vokabulars:** es sollen Ausbildungsformate und -methoden und biomedizinische Inhalte beschreibbar sein.
- **Struktur:** Das Vokabular soll durch eine Vielzahl an unterschiedlichen Nutzern zur Beschreibung von Inhalten einsetzbar sein und deshalb Synonym- und Äquivalenzbeziehungen zwischen Termini unterstützen.
- **Sprache:** Die zu identifizierenden Vokabulare sollen von deutschsprachigen Lehrenden und Lernenden einsetzbar sein.
- **Umfang:** Das Vokabular soll einerseits fein granular genug sein, um Inhalte und Methoden des Medizinstudiums im deutschsprachigen Raum differenziert zu beschreiben, andererseits die annotierenden Nutzer aber nicht durch zu feine Granularität zu irrelevanten Unterscheidungen zwingen.
- **Wartung:** Um nachhaltig einsetzbar zu sein, sollte das Vokabular durch eine feste Institution oder Organisation wenigstens jährlich gewartet sein.
- **Technik:** Datenformat, Verfügbarkeit und Urheberrecht des Vokabulars sollten einen kostengünstigen Einsatz

zur webbasierten Beschreibung von Ressourcen und Elementen medizinischer Curricula ermöglichen.

Im zweiten Schritt wurde die systematische Literaturrecherche im Juni und Juli 2011 auf die Datenbanken Medline <http://www.ncbi.nlm.nih.gov/pubmed> und GMS <http://www.egms.de/dynamic/de/index.htm> unter Einsatz von 7 zusätzlichen, englischsprachigen Suchbegriffen ausgeweitet. Die zusätzlichen Suchbegriffe sind in Tabelle 1 rechts zusammengestellt. Anhand der Abstracts der gefundenen Artikel wurden alle Publikationen mit Bezug zur medizinischen Ausbildung, die ein formalisiertes Vokabular beschreiben, zur weiteren Analyse ausgewählt und die in der Publikation beschriebene(n) Taxonomie(n) und Ontologie(n) identifiziert.

Zu jedem derart identifizierten strukturierten Vokabular wurde eine Websuche durchgeführt sowie die Herausgeber zu Aktualität, Nutzung, Größe und technischen Details wie Datenformat befragt.

Wo kein Herausgeber des Vokabulars zu eruiieren war, wurden die Autoren der beschreibenden Publikationen angeschrieben.

Für jedes strukturierte Vokabular wurden so Sprache, Hierarchieform, Art der Querverweise, Thematik/Fachgebiet, Größe, Aktualität und Pflege, Format und Verfügbarkeit ermittelt.

## Ergebnisse

Es konnten 601 Publikationen mit der beschriebenen Suchstrategie gefunden werden. Anhand der Abstracts dieser Publikationen wurden 20 Artikel identifiziert, die die Entwicklung oder den Einsatz eines strukturierten Vokabulars in der medizinischen Ausbildung beschreiben. In den 20 identifizierten Artikeln werden 14 verschiedene strukturierte Vokabulare beschrieben. Vier davon sind explizit für die Beschreibung von Ausbildungsformaten und -methoden entwickelt.

Mit dem „Unified Medical Language System“ (UMLS) existiert darüber hinaus eine Metaontologie, die ein semantisches Netz von über 130 einzelnen Terminologien umfasst.

## Struktur, Anwendungsgebiete und Sprache

Die 14 identifizierten Vokabulare konnten einem von drei primären Anwendungsgebieten zugeordnet werden. Diese umfassen „Ausbildungsformate und -methoden“, „biomedizinische Inhalte“ sowie „Dokumentation und Verwaltung“ (siehe Tabelle 2). Im Wesentlichen unterscheiden sich die Vokabulare jedoch in ihrem Aufbau durch die Art der Verknüpfungen zwischen ihren Termini: Die polyhierarchischen Ontologien einerseits umfassen in aller Regel nicht nur Eltern-Kind-Beziehungen zwischen den Termini, sondern ermöglichen auch Äquivalenz- und Assoziationsverweise. Die hier beschriebenen monohierarchischen Taxonomien andererseits lassen regelhaft nur eine Form von Verknüpfung ihrer Termini zu und sind allenfalls teilweise um eine Verknüpfung vom Typ „Synonym“ ergänzt. Die einem Vokabular zugrunde liegende Struktur richtet sich bei den hier Betrachteten wesentlich nach dem Zweck und den durch das Vokabular zu beschreibenden Inhalten. Die eher zur Verwaltung und Dokumentation entwickelten Vokabulare folgen einer monohierarchischen Struktur, die Ontologien zur Beschreibung von Ausbildungsformaten und -methoden sind, ebenso wie jene zur Beschreibung biomedizinischer Inhalte, polyhierarchisch strukturiert.

Weitere zentrale Unterscheidungsmerkmale innerhalb dieser Anwendungsgebiete, wie Sprache und Umfang, sind ebenfalls in Tabelle 2 aufgeführt.

Drei der identifizierten Vokabulare zur Beschreibung von Ausbildungsformaten und -methoden liegen primär in Englisch, „Topics for Indexing Medical Education“ (TIME) darüber hinaus auch in Französisch vor. Die TIME-Ontologie ist zur inhaltlichen Beschreibung medizinischer Lehr- und Lerninhalte konzipiert, mit der Taxonomie der „Medical Education Taxonomy Research Organization“ (METRO) können vor allem Methoden und Prozesse der medizinischen Ausbildung annotiert werden. Der „British Education Thesaurus“ (BET) ist ein allgemeiner Ausbildungsthesaurus ohne spezifischen Bezug zur Medizin.

Mit der „Ontology of Bio-Medical Educational Objectives“ (OBEO) steht darüber hinaus eine deutschsprachige On-

tologie zur semantischen Annotation vor allem von Lernzielen in der medizinischen Ausbildung zur Verfügung. Gleichzeitig sind die für die Annotation biomedizinischer Inhalte vorgesehenen Vokabulare teilweise sehr umfangreich und / oder stark spezialisiert. Lediglich vier von ihnen sind für die umfassende Beschreibung der Medizin entwickelt und fokussieren nicht auf ein Themengebiet wie beispielsweise die Anatomie, Onkologie oder Pharmazie (siehe Tabelle 2), sondern eignen sich zur Beschreibung eines umfassenden Teils der biomedizinischen Inhalte.

## Umfang, Wartung und Technik

Während der BET die umfangreiche Annotation von allgemeinen Ausbildungsinhalten ermöglicht, sind die Möglichkeiten zur Beschreibung medizinspezifischer Inhalte mit diesem Thesaurus eingeschränkt. Wie aus Tabelle 2 ersichtlich, sind die anderen drei identifizierten Vokabulare (METRO, OBEO und TIME) zur Beschreibung von Ausbildungsformaten und -methoden weniger umfangreich als die Vokabulare zur Beschreibung von biomedizinischen Inhalten oder Inhalten der Verwaltung und Dokumentation.

Umgekehrt sind die vier identifizierten Vokabulare zur umfassenden Beschreibung biomedizinischer Inhalte teilweise sehr umfangreich, wie aus Tabelle 2 ersichtlich. Um aus diesen Vokabularen zur entweder inhaltlich-biomedizinischen oder methodisch-didaktischen Beschreibung von Objekten ein je nach Anwendungsgebiet geeignetes Vokabular wählen zu können, sind in Tabelle 3 Daten zu Aktualität und Pflege, Datenformat und Kosten für diese sieben Ontologien – drei zur Beschreibung von Ausbildungsformaten und -methoden, vier zur Beschreibung biomedizinischer Inhalte – zusammengestellt.

## Diskussion

Die Nutzung strukturierter Vokabulare, wie sie Taxonomien und Ontologien darstellen, ermöglicht den maschinellen Vergleich und Austausch von Daten zwischen Institutionen und Systemen [17] und den Vergleich unterschiedlicher Inhalte mit automatischer Identifikation von inhaltlichen Schnittmengen zwischen zwei Informationsmengen: Dadurch wird beispielsweise der automatische Abgleich eines medizinischen Curriculums mit beliebigen Outcomerframeworks möglich [15] – eine Funktion, die nicht zuletzt im Rahmen von Akkreditierungen von besonderem Interesse sein dürfte [4].

Da die Vokabulare prinzipiell zur formalen inhaltlichen Beschreibung beliebiger Objekte einsetzbar sind, können neben klassischen Ressourcen wie Büchern, Medien und Programmen auch Kurse, Lernziele oder Prüfungsfragen und jede Form von Onlineressource annotiert und damit Ausgangspunkt oder Ergebnis einer Suche nach thematisch verwandten Objekten werden. Einige wenige Implementierungen curricularer Kartierung greifen diese Möglichkeit auf, um die inhaltlichen Zusammenhänge von

Tabelle 2: Anwendungsgebiete, Sprache und Umfang identifizierter Vokabulare. Angaben zum Umfang laut Beschreibungen des Herausgebers. Mit ° markierte Vokabulare sind in der Metaontologie UMLS enthalten.

<b>Polyhierarchische Ontologien zur Beschreibung von Ausbildungsformaten und -methoden</b>				
Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz				
Kürzel	Name	Thema	Sprache(n)	Umfang
BET	British Education Thesaurus	Ausbildungsinhalte aller Art	Eng	Keine Angaben
METRO	Medical Education Taxonomy Research Organization	Prozesse & Methoden in medizinischer Ausbildung	Eng	ca. 180 Termini
OBEO	Ontology of Bio-Medical Educational Objectives *	Medizinische Ausbildung (v. a. Lernziele)	Deu	eigene: 72 Klassen, 12 Objekt-Relationen; Grundlage: Upper-Level Ontologie BioTop (375 Klassen, 76 Objekt-Relationen), Toplevel-Ontologie DOLCE-Life (37 Klassen, 70 Objekt-Relationen)
TIME	Topics for indexing Medical education	Biomedizinische Inhalte, Medizinische Lehr-/ Lerninhalte	Eng, Frz	1.500 Termini, 2.400 Querverweise
<b>Polyhierarchische Ontologien biomedizinischer Inhalte</b>				
Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz				
Kürzel	Name	Thema	Sprache(n)	Umfang
FMA	Foundational Model of Anatomy	Anatomie	Eng, (Deu), weitere	75.000 Klassen, 120.000 Termini, 2.100.000 Querverweise, 168 Beziehungsarten
-	Gene Ontology*	Biowissenschaften	Eng	34.439 Termini
MeSH	Medical Subject Heading°	Biomedizinische Inhalte	Eng, Deu	Englisch: 26.142 Hauptschlagwörter, 181.141 Synonyme Deutsch: 26.142 Hauptschlagwörter, 57.656 Synonyme
NCI Thesaurus	National Cancer Institute Thesaurus*	Onkologie	Eng	34.000 Konzepte, 200.000 Querverweise
RePORT	Research Portfolio Online Reporting Tools	Biomedizinische Inhalte	Eng	Bezug aus mehreren Datenbanken (eRA databases, Medline, PubMed Central, the NIH Intramural Database, iEdison)
-	RxNorm °	Pharmazeutika	Eng	keine genauen Angaben, jedoch <u>alle</u> verschreibungspflichtige Medikamente der USA enthalten
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms°	Biomedizinische Inhalte, Klinik	Eng, (Deu), weitere	311.000 Konzepte, 800.000 Termini, 1.360.000 Querverweise
UMLS	Unified Medical Language System	Biomedizinische Inhalte	Eng, Deu, weitere	Semantisches Netz von >130 Einzelterminologien (unter anderem °); > 1.179.000 Konzepte, > 4.174.000 Termini
<b>Monohierarchische Taxonomien für Verwaltung und Dokumentation</b>				
Verweise: Hierarchie, Synonyme (nur #)				
ICD	International Classification of diseases°	Klassifikation der Krankheiten und verwandter Gesundheitsprobleme	Eng, Deu	70.000 Kodierungen
LOINC	Logical Observation Identifier Names and Codes#°	Untersuchungs-/ Testergebnisse aus Labor und Klinik	Eng, (Deu)	7 Kategorien, 100.000 Datensätzen (Deu: 3800)
OPS	Operationen- und Prozedurenschlüssel	Operationstechniken/ Prozeduren	Deu	6 Kapitel, 65 Bereiche, ca. 28.300 Prozedurenklassen

Tabelle 3: Aktualität und Pflege, Datenformat, Kosten und Weblinks von Vokabularen medizinischer Inhalte.

<b>Polyhierarchische Ontologien zur Beschreibung von Ausbildungsformaten und -methoden</b>					
Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz					
<b>Kürzel</b>	<b>Name</b>	<b>Aktualität &amp; Pflege</b>	<b>Datenformat</b>	<b>Kosten</b>	<b>Weblinks</b>
METRO	Medical Education Taxonomy Research Organization	Seit 2005 keine Weiterentwicklung	Download – XML, Website (Blogspot)	Open Source	<a href="http://metro2.blogspot.de/">http://metro2.blogspot.de/</a>
OBEO	Ontology of Bio-Medical Educational Objectives*	„work-in-progress“ Entwurf (Universitätsklinikum Freiburg, Dtl.)	Download – OWL	Open Source	<a href="http://www.imbi.uni-freiburg.de/ontology/obeo/obeo.owl">http://www.imbi.uni-freiburg.de/ontology/obeo/obeo.owl</a>
TIME	Topics for indexing Medical education	„work-in-progress“ (University of Ottawa, Kanada)	Download – XML, Textformat, Web based	Open Source	<a href="http://www.time-item.org">www.time-item.org</a> (derzeit offline)
<b>Polyhierarchische Ontologien biomedizinischer Inhalte</b>					
Verweise: Hierarchie, Assoziation (nicht *), Äquivalenz					
<b>Kürzel</b>	<b>Name</b>	<b>Aktualität &amp; Pflege</b>	<b>Datenformat</b>	<b>Kosten</b>	<b>Weblinks</b>
MeSH	Medical Subject Heading	jährlich (NLM, USA; DIMDI, Dtl.)	Download/ CD-ROM – XML, CSF, Textformat; Web based (MeSH-Browser)	Open Source (Englisch, nach Registrierung NLM), kostenpflichtig (Deutsch)	<a href="http://www.nlm.nih.gov/mesh/">http://www.nlm.nih.gov/mesh/</a> (Englisch) <a href="http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm">http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm</a> (Deutsch)
RePORT	Research Portfolio Online Reporting Tools	wöchentlich (NIH, USA)	Download – XML, CSV; Web based (RePorter)	Open Source	<a href="http://report.nih.gov/">http://report.nih.gov/</a>
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms	halbjährlich (IHTSDO, Dänemark)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	kostenpflichtig für Nicht-Mitgliedsstaaten (Dtl.) (nach Registrierung UMLS/IHTSDO)	<a href="http://www.ihtsdo.org/snomed-ct/">http://www.ihtsdo.org/snomed-ct/</a>
UMLS	Unified Medical Language System	jährlich (NLM, USA)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	Open Source (nach Registrierung NLM, Copyright)	<a href="http://www.nlm.nih.gov/research/umls/">http://www.nlm.nih.gov/research/umls/</a>

Elementen eines Curriculums aufzuzeigen [5], [15]. Durch die maschinenlesbare inhaltliche Beschreibung von Lernmodulen in e-learning-Systemen mit Hilfe von Taxonomien könnten die Nutzer solcher Module automatisch auf thematisch verwandte Module in denselben (oder auch anderen) e-learning-Plattformen verwiesen werden. Formalisierte inhaltliche Beschreibungen von Ressourcen sind zudem flexibel zu durchsuchen und leicht zu warten [8].

Der Einsatz solcher Vokabulare setzt die konsequente Annotation aller neuen Objekte mittels des definierten Vokabulars voraus, was einen Mehraufwand beim Erstellen neuer Inhalte bedeutet. Gleichzeitig entfällt die Notwendigkeit der manuellen Verknüpfung neuer Inhalte. Das semantische Beschreiben von Objekten wird also erst mit zunehmendem Umfang an Inhaltselementen effizient.

Im deutschsprachigen Raum konnte keine ad hoc nutzbare Taxonomie oder Ontologie zur Annotation von Ressourcen und Inhalten der medizinischen Ausbildung identifiziert werden.

Die in deutscher Sprache verfügbaren strukturierten Vokabulare sind entweder - wie etwa OBEO - bisher nicht fein granular genug, um eine sinnvolle inhaltliche Unterscheidung zwischen Lernressourcen zu ermöglichen oder zwar sehr fein granular, aber eher für die inhaltliche Annotation von Patientendaten (SNOMED-CT) oder biomedizinischen Publikationen (MeSH) geeignet.

Die teilweise erheblichen Unterschiede im Umfang zwischen den identifizierten Vokabularen, die vor allem aus den sehr spezifischen Einsatzgebieten wie etwa der Beschreibung von Unterrichtsformaten (METRO), resultieren, werfen die Frage auf, ob die Granularität ausreichend fein zur differenzierten Annotation von Inhalten und Methoden der medizinischen Ausbildung ist.

Die in englischer Sprache verfügbaren, speziell für die Domäne der medizinischen Ausbildung entwickelten Ontologien, wie TIME oder METRO, müssten zudem zunächst ins Deutsche überführt werden. Dabei entstehen Probleme in der Auswahl inhaltlicher Entsprechungen. So findet der englischsprachige Begriff „medical education“ eine deutsche Entsprechung nur in der Trias aus medizinischer Aus-, Fort- und Weiterbildung, die im Englischen wiederum eher den Begriffen „graduate-“, „continous-“ und „postgraduate-“ medical education entsprechen.

Ein weiteres Problem einer Übersetzung ist die Frage, ob und wie eine Übersetzung an die mögliche Weiterentwicklung und Pflege des originalsprachlichen Vokabulars gebunden werden kann.

Eine alternative Neuentwicklung einer primär deutschsprachigen Ontologie ist jedoch in jedem Fall mit erheblichem Aufwand verbunden.

Auffällig bei den identifizierten und in Tabelle 3 zusammengefassten Vokabularen ist insbesondere, dass alle wenigstens online verfügbar sind und teilweise direkt Technologien des Internets wie etwa XML als Datenformat einsetzen. Obwohl sie nicht primär für den Einsatz im Web entwickelt wurden, sind sie demnach jedoch zumin-

dest dafür geeignet. Die technischen Grundlagen auf dem Weg zum Web 3.0 sind damit gelegt [9].

## Limitationen

Wie jede Literaturrecherche ist diese Arbeit limitiert durch Suche und Auswahl der verwendeten Publikationen.

Darüber hinaus besteht die Möglichkeit, dass strukturierte Vokabulare existieren, zu denen keine in den verwendeten Datenbanken indizierten Publikationen vorliegen.

In der Vorbereitung auf die hier dargestellte Arbeit konnten die Autoren keine allgemein akzeptierten Kriterien zur Beschreibung und zum Vergleich strukturierter Vokabulare identifizieren. In einem ersten Schritt wurden deshalb zunächst in einer Teilmenge der durchsuchten Datenbanken Taxonomien identifiziert und aus diesen Publikationen die von den jeweiligen Autoren selbst verwendeten Kriterien zur Beschreibung ihres strukturierten Vokabulars ermittelt. Aus diesen Kriterien wurden für diese Arbeit im Konsens solche extrahiert, die den Autoren geeignet erscheinen, Taxonomien für die medizinische Ausbildung miteinander zu vergleichen. In der Auswahl dieser Kriterien haben die Autoren sich um die Generalisierbarkeit der Ergebnisse auf die Situation an den deutschsprachigen medizinischen Fakultäten bemüht. Lokal könnten andere oder weitere, hier nicht behandelte Kriterien von Bedeutung sein.

## Fazit

Die inhaltliche Beschreibung (semantische Annotation) von medizinischen Lernressourcen bietet die Möglichkeit, inhaltlich verwandte Elemente über Systemgrenzen hinweg zu identifizieren. Bisher existiert keine geeignete deutschsprachige Taxonomie oder Ontologie, um eine systematische Ressourcenbeschreibung für Medizinstudierende in Deutschland zu realisieren und das Potential des Web 3.0 zur Unterstützung der Curriculumskartierung nutzen zu können. Mögliche Ansatzpunkte sind die Neuentwicklung eines geeigneten strukturierten Vokabulars, die Übersetzung existierender englischsprachiger Vokabulare oder eine Kombination aus bestehenden Teilkabularen, ggf. mit einer teilweisen Übersetzung englischsprachiger Vokabulare.

## Anmerkung

WB leitet das Lernzentrum der Charité und ist Mitglied der GMA und deren Ausschüsse „Praktische Fertigkeiten“ und „Methodik der Ausbildungsforschung“. AJ und PS sind studentische Tutor/-innen des Lernzentrums und Studierende der Charité. FB ist Co-Entwickler der OBEO. OA leitet die Abteilung für Curriculumorganisation der Charité und ist Mitglied der GMA und deren Ausschuss „Methodik der Ausbildungsforschung“.

## Danksagung

Die Autoren danken Dr. Ullrich Woermann, Bern sowie zwei anonymen Reviewern für Ihre kollegiale und konstruktive Kritik am Manuskript und Rudi Mörgli für die Übersetzung des Manuskripts ins Englische.

## Interessenkonflikt

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

## Literatur

- Holzer M, Pfähler M, Hege I, Fischer M. Wer sucht, der soll finden! - Ein Überblick über Verschlagwortung und Suche medizinischer Lerninhalte. *GMS Med Inform Biom Epidemiol.* 2006;2(3):Doc20. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/mibe/2006-2/mibe000039.shtml>
- Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. *Acad Med.* 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13
- Junco R, Mastrodicasa J. Connecting to the Net Generation: What Higher Education Professionals Need to Know About Today's Students. Washington, DC: National Association of Student Personnel Administrators; 2007.
- Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach.* 2001;23(2):123-137. DOI: 10.1080/01421590120036547
- Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House; 2010. Doc10gma13. DOI: 10.3205/10gma013
- Berners-Lee T, Hendler J, Lassila O. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Sci Am.* 2001;284:34-43. DOI: 10.1038/scientificamerican0501-34
- Alby T. Web 2.0. Konzepte, Anwendungen, Technologien. München: Hanser Verlag; 2007.
- Berners-Lee T, Fischetti M. Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventors. New York: HarperBusiness; 2006.
- Segaran T, Evans C, Taylor J. Programming the semantic web. Sebastopol, CA: O'Reilly & Associates; 2009.
- Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. *Med Teach.* 2007;29(7):655-659. DOI: 10.1080/01421590701615808
- Willett TG, Marshall KC, Broudo M, Clarke M. It's about TIME: a general-purpose taxonomy of subjects in medical education. *Med Educ.* 2008;42(4):432-438. DOI: 10.1111/j.1365-2923.2008.03012.x
- Boeker M, Schober D, Schulz S, Balzer F. Ontology of Bio-Medical Educational Objectives (OBEO): ein Vorschlag für eine Ontologie medizinischer Lernziele. *GMS Med Inform Biom Epidemiol.* 2010;6(2):Doc11. DOI: 10.3205/mibe000111
- Stenzhorn H, Schulz S, Boeker M, Smith B. Adapting Clinical Ontologies in Real-World Environments. *J Univers Comput Sci.* 2008;14(22):3767-3780.
- Ellaway R, Evans P, McKillop J, Cameron H, Morrison J, McKenzie H, Mires G, Pippard M, Simpson J, Cumming A, Harden R, Guild S. Cross-referencing the Scottish Doctor and Tomorrow's Doctors learning outcome frameworks. *Med Teach.* 2007;29(7):630-635. DOI: 10.1080/01421590701316548
- Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" - eine Validierungsstudie. *GMS Z Med Ausbild.* 2012;29(4):Doc58. DOI: 10.3205/zma000828
- Willett TG. Current status of curriculum mapping in Canada and the UK. *Med Educ.* 2008;42(8):786-793. DOI: 10.1111/j.1365-2923.2008.03093.x
- Uschold M, Gruninger M. Ontologies: Principles, Methods and Applications. *Knowl Eng Rev.* 1996;11:93-155. DOI: 10.1017/S0269888900007797

### Korrespondenzadresse:

Wolf E. Blaum  
Charité - Universitätsmedizin Berlin, Abteilung für  
Curriculumsorganisation, Lernzentrum, Charitéplatz 1,  
10117 Berlin, Deutschland, Tel.: +49 (0)30/450-576003,  
Fax: +49 (0)30/450-576952  
[wolf.blaum@charite.de](mailto:wolf.blaum@charite.de)

### Bitte zitieren als

Blaum WE, Jarczewski A, Balzer F, Stötzner P, Ahlers O. Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche. *GMS Z Med Ausbild.* 2013;30(1):Doc13.  
DOI: 10.3205/zma000856, URN: urn:nbn:de:0183-zma0008568

### Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2013-30/zma000856.shtml>

**Eingereicht:** 29.06.2003

**Überarbeitet:** 05.09.2012

**Angenommen:** 12.10.2012

**Veröffentlicht:** 21.02.2013

### Copyright

©2013 Blaum 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.

# Towards Web 3.0: Taxonomies and ontologies for medical education - a systematic review

## Abstract

**Introduction:** Both for curricular development and mapping, as well as for orientation within the mounting supply of learning resources in medical education, the Semantic Web ("Web 3.0") poses a low-threshold, effective tool that enables identification of content related items across system boundaries. Replacement of the currently required manual with an automatically generated link, which is based on content and semantics, requires the use of a suitably structured vocabulary for a machine-readable description of object content.

Aim of this study is to compile the existing taxonomies and ontologies used for the annotation of medical content and learning resources, to compare those using selected criteria, and to verify their suitability in the context described above.

**Methods:** Based on a systematic literature search, existing taxonomies and ontologies for the description of medical learning resources were identified. Through web searches and/or direct contact with the respective editors, each of the structured vocabularies thus identified were examined in regards to topic, structure, language, scope, maintenance, and technology of the taxonomy/ontology. In addition, suitability for use in the Semantic Web was verified.

**Results:** Among 20 identified publications, 14 structured vocabularies were identified, which differed rather strongly in regards to language, scope, currency, and maintenance.

None of the identified vocabularies fulfilled the necessary criteria for content description of medical curricula and learning resources in the German-speaking world.

**Discussion:** While moving towards Web 3.0, a significant problem lies in the selection and use of an appropriate German vocabulary for the machine-readable description of object content. Possible solutions include development, translation and/or combination of existing vocabularies, possibly including partial translations of English vocabularies.

**Keywords:** medical education, semantic web, web 3.0, taxonom, ontology, curricular mapping, curriculum charting, e-learning, new media

**Wolf E. Blaum**<sup>1,2</sup>

**Anne Jarczewski**<sup>2</sup>

**Felix Balzer**<sup>1,2</sup>

**Philip Stötzner**<sup>2</sup>

**Olaf Ahlers**<sup>1,2</sup>

1 Charité - University Medicine Berlin, Campus Charité Mitte and Campus Virchow-Klinikum, Department of Anesthesiology and Intensive Care Medicine Campus, Berlin, Germany

2 Charité - University Medicine Berlin, Department of Curriculum Management, Learning Center, Berlin, Germany

## Authorship

The authors Blaum and Jarczewski contributed equally to the study.

## Introduction

At the present time, mounting supply of available learning resources in medical education, particularly online, can hardly be used effectively by teachers and learners due to the lack of a simple and systematic mode of access to its contents [1]. Due to the constantly increasing number of available sources, students are confronted with the challenge of identifying, in a sea of possibilities, themat-

ically appropriate, high-quality, and didactically significant material that corresponds to their level of education. Computer-based learning in particular has experienced a substantial increase of resources. In a recent review in the topic of computer-based virtual patients - a mere subset of the computer-based resources - Cook et al. identified 698 publications on this single issue [2]. At the same time, the PC-based learning is becoming increasingly important with the arrival of the "Generation Y" [3]. Teachers and planners are also being challenged to present (to map) the contextual relationships in the curriculum for which they are responsible [4], a task that until now required only a manual linking of curricular elements.

## Semantic Web

Notably, development of the Internet towards the Semantic Web (Web 3.0) presents considerable potential for the contextual linking of learning resources, as well as applications in curricular mapping [5], [6].

While at first content of the Internet was essentially determined by a few authors, who provided information to a broad, consuming audience, the current "Web 2.0" is characterized by user-generated content [7]. A common feature in data from both versions is that their information can only be understood and interpreted by humans. Hence, any contextual relation between data requires human interaction.

The term "Web 3.0" (also "Semantic Web") refers to a network data storage system which, through the addition of machine-readable meta-information (per example, via structured vocabularies), allows machines to automatically generate and discover contextual relations between data objects [8]. This innovation requires a machine-readable description of the information contained in data [6], [9].

A suitable type of content description is the use of structured vocabularies. In this case, each individual resource is tagged with the defined terms of a structured vocabulary.

## Taxonomy und Ontology

Taxonomies and ontologies provide controlled vocabularies for content description (semantic annotation) of objects, such as, per example, learning resources [10]. Relationship between the terms of a vocabulary are stored within a taxonomy, allowing for an (manually initiated) automated search to identify similar content descriptions among a multitude of learning resources, independent of their location [1], [11], [12].

Relationships between vocabulary terms of a taxonomy are typically monohierarchical, thus each term has exactly one (or no) parent term ("is" - or "is a part of" - relationship), as well as an arbitrary number of subordinate (child) terms. A classic example of a strictly monohierarchical taxonomy is the International Classification of Diseases <http://www.dimdi.de/static/de/klassi/index.htm>. Conversely, ontologies may also utilize polyhierarchical structures, so that a single vocabulary term may have multiple parent terms. The fundamental difference to taxonomies is that, instead of language constructs, logically defined formalisms are used for the definition of objects and relationships in the real world. Due to this independence from linguistic properties, ontologies represent a robust toolkit for semantic descriptions, and can greatly facilitate the standardization of terminology [13].

A well known example of a polyhierarchical ontology with equivalence relationships are the Medical Subject Headings <http://www.nlm.nih.gov/mesh/>, which are used to describe publications indexed in Medline. Figure 1 clarifies the differences between taxonomy and ontology.

Taxonomies and ontologies are fundamentally suitable to create contextual relationships between objects, even if they are located in separate, independent systems. They may also further assist preparation, if not complete automation, of the necessary – but burdensome – manual linking between different outcome frameworks [14] or learning objective catalogs [15], as they are generally used in a minority of approaches to curriculum mapping.

To the best of our knowledge, generally accepted criteria for description and comparison of structured vocabularies are currently not available.

## Research Question

When choosing an appropriate taxonomy or ontology, medical faculties are faced with a multitude of offers that differ in scope, maintenance, and availability. Aim of this study is to compile existing taxonomies and ontologies used for annotation of medical learning resources and content, to compare those using purpose-built criteria, and to examine their suitability for use in the Semantic Web.

## Methods

A literature search in the databases of the GMS (German Medical Sciences) <http://www.egms.de/dynamic/de/index.htm> was initially carried out in June 2011, in order to identify publications describing at least one taxonomy/ontology with current or potential use in the medical field. Searched terms are summarized in Table 1, on the left column.

Five structured vocabularies were identified.

The authors extracted criteria used to assess the suitability of a vocabulary for annotation of elements of medical education from these publications in consensus.

These extracted criteria include:

- *Vocabulary theme*: training formats and methods, as well as biomedical content, should be described.
- *Structure*: The vocabulary should support content description by a variety of different users, and therefore support synonym and equivalence relationships between terms.
- *Language*: German-speaking teacher and students should be able to use the respective vocabulary.
- *Scope*: The vocabulary should be specific enough to differentiate the contents and methods used in medical studies in the German-speaking world; but conversely, not so detailed as to force the users to annotate using irrelevant distinctions.
- *Maintenance*: In order to remain current, the vocabulary should be updated by a specified institution or organization at least once a year.
- *Technology*: Data format, availability, and copyright of the vocabulary should allow for a cost-effective use in description of web-based resources and elements of medical curricula.

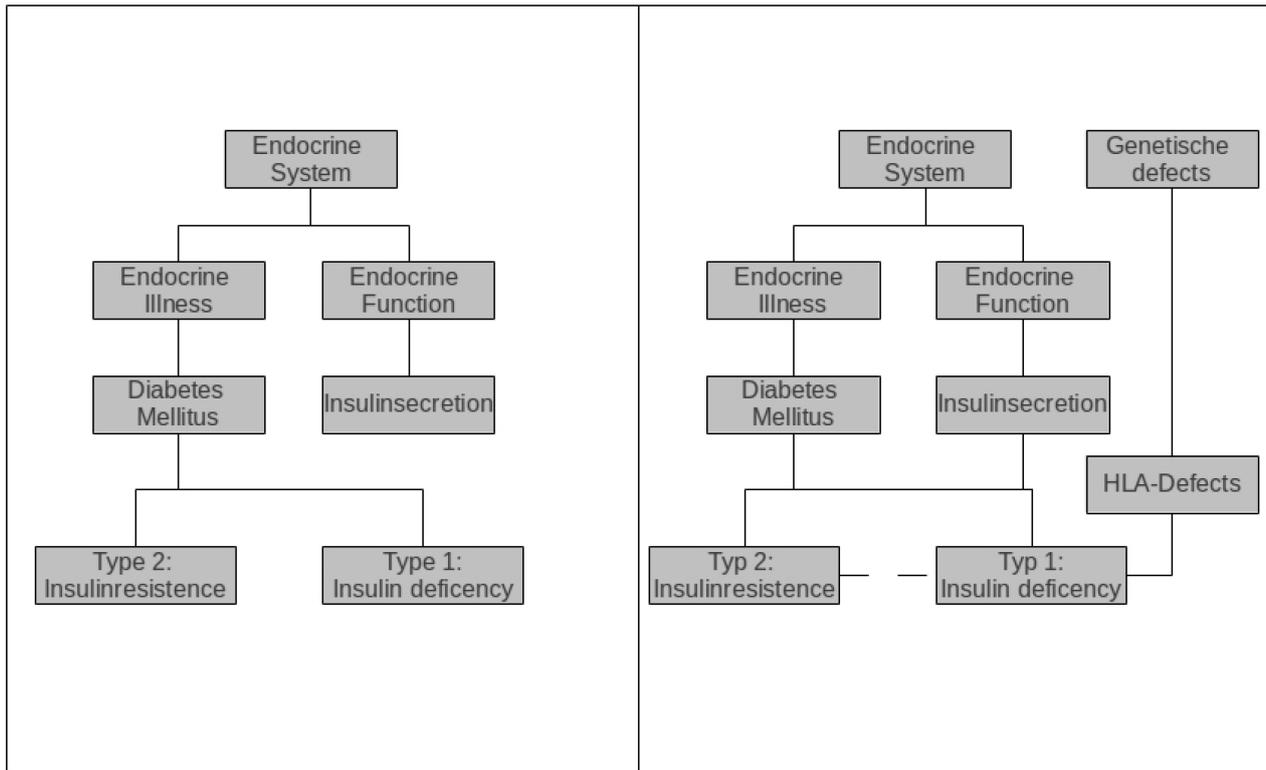


Figure 1: In a taxonomy (left), each element is assigned exactly one parent term (solid lines). In an ontology (right), an element may not only be assigned multiple parent terms, but may also be associated with other elements (dashed line).

Table 1: Terms used in the systematic literature search (Figures: publications used in the analysis / total number of hits)

German – mainly GMS	5/37	English	15/564
medizinische Taxonomie (medical taxonomy)	2/9	Medical education ontology	6/37
medizinische Ontologie (medical ontology)	3/12	Taxonomy medical education german	0/93
deutschsprachige Taxonomie (German language taxonomy)	0/2	Taxonomy medical education elearning	5/238
Taxonomie medizinische Lehre (taxonomy medical teachings)	0/2	taxonomy learning resources	1/127
Ontologie medizinische Lehre (ontology medical teachings)	0/6	medical education semantic network	2/15
Taxonomie medizinische Lernressourcen (taxonomy medical educational resources)	0/3	medical education curriculum mapping taxonomy	0/18
Taxonomie medizinische Lerninhalte (taxonomy medical educational content)	0/3	taxonomy medical education learning objectives	1/36

The second step took place in June and July 2011, and involved an expansion of the systematic literature search in Medline <http://www.ncbi.nlm.nih.gov/pubmed> and in GMS <http://www.egms.de/dynamic/de/index.htm>, using 7 supplementary English terms. The additional terms searched are summarized in Table 1, on the right column. Based on the abstracts of the search results, all publications describing a formal vocabulary related to medical education were selected for further analysis, and the cited taxonomies and ontologies were identified. A web search was conducted for each of the structured vocabularies thus identified, and their respective editors were contacted and questioned regarding the actuality, use, size and technical details (such as data format) of their vocabularies. If the editor of a vocabulary could not be identified, the authors of the publication describing the respective vocabulary were then contacted.

Each structured vocabulary was analyzed in regard to language, hierarchical design, type of cross-referencing, topic/subject area, size, currency, maintenance, format, and availability.

## Results

The described search strategy identified 601 publications. Based on the abstracts of these publications, 20 articles were identified describing the development or use of a structured vocabulary in medical education. Among the 20 articles found, 14 different structured vocabularies were described. Four of these were explicitly designed for description of training formats and methods. Furthermore, a Metaontology comprising a semantic network of more than 130 different terminologies is

available under the "Unified Medical Language System" (UMLS).

## Structure, Applications and Language

The 14 identified vocabularies could be classified into one of three primary applications. These applications include "Training Formats and Methods", "Biomedical Content", and "Administration and Documentation" (see Table 2). In essence, vocabulary structures differ in the method by which their terms are connected: polyhierarchical ontologies include not only the parent-child relationships between terms, but also allow for analysis of equivalency and association references. Conversely, the described monohierarchical taxonomies allow only one type of relationship between its terms, and are at best supplemented by a "synonym" type link. The underlying structure of a vocabulary depends substantially on its purpose or on the content that the vocabulary is designed to describe. Vocabularies primarily developed for documentation and administration tend to follow a monohierarchical structure, while ontologies meant to describe training formats and methods, as well as those describing biomedical content, are usually polyhierarchically structured. Further essential distinguishing features within these areas of application, such as language and scope, are listed in Table 2.

Three of the identified vocabularies designed to describe training formats and methods are mainly in English, such as "Topics for Indexing Medical Education" (TIME), and secondarily in French. The TIME-ontology is designed for description of medical teaching and learning content, while the "Medical Education Taxonomy Research Organization" (METRO) taxonomy is mainly suitable for the annotation of methods and processes of the medical education. The "British Education Thesaurus" (BET) is a general education thesaurus without any specific reference to medicine.

Additionally, an ontology in German language is available through the "Ontology of Bio-Medical Educational Objectives" (OBEO), which is used primarily for the semantic annotation of learning objectives in medical training.

Vocabularies intended for annotation of biomedical content are partially very extensive and/or specialized. Only four of them were designed for a comprehensive medical description, not focusing on a single thematic area, such as anatomy, oncology, or pharmacology (see Table 2), but rather fit to describe extensively a broad area of the biomedical content.

## Scope, Maintenance, and Technology

While BET allows extensive annotation of general educational content, the possibility of describing specific medical content with this thesaurus is restricted. As seen in Table 2, the other three identified vocabularies (METRO, OBEO, and TIME), which have been designed for description of training formats and methods, are significantly less extensive than vocabularies used for describing

biomedical content, or content administration and documentation.

Conversely, the four identified vocabularies for comprehensive description of biomedical content are sometimes very extensive (see Table 2).

Table 3 summarizes data on updates and support, data format, and costs for the seven identified ontologies - three for the description of training formats and methods, and four for the description of biomedical content.

## Discussion

Use of structured vocabularies, such as taxonomies and ontologies, enables automatic comparison and exchange of data between institutions and systems [17]. It also allows automatic comparison of multiple sets of information, which automatically identifies overlapping content: a function that would, for instance, enable automatic adjustment of a medical curriculum with any possible outcome frameworks [15] - a particularly interesting feature in terms of accreditation.

Since the vocabularies can in principle be used for the formal content description of any object, they may be used - in addition to traditional resources such as books, media, and programs - to annotate courses, learning objectives, exam questions, or any other form of online resource. As such, they provide a starting point or search results for thematically related objects. A few curriculum mapping implementations utilize this opportunity to show contextual relationships between elements in a curriculum [5], [16]. Users of such modules can be automatically referred to related modules in the same (or another) e-learning platform by machine-readable description of learning module contents in e-learning systems, which is based on taxonomies.

In addition, formalized content description of resources is easy to browse and to maintain [8].

Use of such vocabularies requires consistent annotation of all new objects using a defined vocabulary, which signifies an additional effort when creating new content. Conversely, this eliminates the need for manually linking new content. Semantic description of objects, therefore, can only be effective with an increasing amount of content elements.

No German language based, purpose built taxonomy or ontology for the annotation of resources and medical education content could be identified.

The structured vocabularies available in German - such as OBEO - are not yet specific enough to allow for a meaningful distinction among learning resources, but rather appropriate for the content annotation of patient data (SNOMED-CT) or biomedical publications (MeSH).

Considerable differences in the scope of the identified vocabularies - which resulted primarily from very specific applications, such as the description of teaching formats (METRO) - raise the question of whether their granularity is fine enough for a differentiated annotation of content and methods in medical education.

**Table 2: Applications, language and extent of the identified vocabularies. Details regarding scope were provided by the individual editors. Vocabularies marked with ° are included in the UMLS Metaontology.**

<b>Polyhierarchical Ontologies for the Description of Training Formats and Methods</b>			
References: hierarchy, (no*) association, equivalency			
Acronym	Name	Topic	Language
BET	British Education Thesaurus	All forms of educational content	English
METRO	Medical Education Taxonomy Research Organization	Processes & Methods in medical education	English
OBEO	Ontology of Bio-Medical Educational Objectives *	Medical education (mainly learning objectives)	German
TIME	Topics for Indexing Medical Education	Biomedizinische Inhalte, Medizinische Lehr-/ Lerninhalte	English, French
<b>Polyhierarchical Ontologies for Biomedical Content</b>			
References: hierarchy, (no*) association, equivalency			
Acronym	Name	Topic	Language
FMA	Foundational Model of Anatomy	Anatomy	Eng. (Ger), etc
-	Gene Ontology*	Biosciences	English
MeSH	Medical Subject Heading°	Biomedical Content	English, German
NCI Thesaurus	National Cancer Institute Thesaurus*	Oncology	English
RePORT	Research Portfolio Online Reporting Tools	Biomedical Content	English
-	RxNorm °	Pharmaceutics	English
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms°	Biomedical Content, Clinics	Eng. (Ger), etc
UMLS	Unified Medical Language System	Biomedical Content	Eng. (Ger), etc
<b>Mono-hierarchical Taxonomies for Administration and Documentation</b>			
References: hierarchy, synonyms (only #)			
Acronym	Name	Topic	Language
ICD	International Classification of Diseases [Check German Version]°	Classification of diseases and related health problems	English, German
LOINC	Logical Observation Identifier Names and Codes#	Examination/test results from clinic/laboratory	English, (German)
OPS	Operationen- und Prozedurenschlüssel	Surgical techniques and procedures	German

Tabelle 3: Aktualität und Pflege, Datenformat, Kosten und Weblinks von Vokabularen medizinischer Inhalte.

<b>Polyhierarchical Ontologies for the Description of Training Formats and Methods</b>					
References: hierarchy, (no*) association, equivalency					
<b>Acronym</b>	<b>Name</b>	<b>Updates &amp; Support</b>	<b>Data Format</b>	<b>Cost</b>	<b>Web Links</b>
METRO	Medical Education Taxonomy Research Organization	No further development since 2005	Download – XML, Website (Blogspot)	Open Source	<a href="http://metro2.blogspot.de/">http://metro2.blogspot.de/</a>
OBEO	Ontology of Bio-Medical Educational Objectives*	„work-in-progress“ preliminary design (Freiburg University Hospital, Germany)	Download [Check German Version] – OWL	Open Source	<a href="http://www.imbi.uni-freiburg.de/ontology/obeo/obso.owl">http://www.imbi.uni-freiburg.de/ontology/obeo/obso.owl</a>
TIME	Topics for Indexing Medical Education	„work-in-progress“ (University of Ottawa, Canada)	Download – XML, Textformat, Web based	Open Source	<a href="http://www.time-item.org">www.time-item.org</a> (currently offline)
<b>Polyhierarchical Ontologies for Biomedical Content</b>					
References: hierarchy, (no*) association, equivalency					
<b>Acronym</b>	<b>Name</b>	<b>Updates &amp; Support</b>	<b>Data Format</b>	<b>Cost</b>	<b>Web Links</b>
MeSH	Medical Subject Heading	annual (NLM, USA; DIMDI, Germany)	Download/CD-ROM – XML, CSF, Text format; Web based (MeSH-Browser)	Open Source (English, NLM registration required), fee required (German)	<a href="http://www.nlm.nih.gov/mesh/">http://www.nlm.nih.gov/mesh/</a> (English) <a href="http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm">http://www.dimdi.de/static/en/klassi/mesh_umls/mesh/index.htm</a> (German)
RePORT	Research Portfolio Online Reporting Tools	weekly (NIH, USA)	Download – XML, CSV; Web based (RePorter)	Open Source	<a href="http://report.nih.gov/">http://report.nih.gov/</a>
SNOMED-CT	Systematized Nomenclature of Medicine – Clinical Terms	biannual (IHTSDO, Denmark)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	Fee required for non-member states (Germany) (UMLS/ IHTSDO registration required)	<a href="http://www.ihtsdo.org/snome-d-ct/">http://www.ihtsdo.org/snome-d-ct/</a>
UMLS	Unified Medical Language System	annual (NLM, USA)	Download/ CD-ROM, MetamorphoSys (Java); Web based (UMLS Terminology Services (UTS))	Open Source (NLM registration required, Copyright)	<a href="http://www.nlm.nih.gov/research/umls/">http://www.nlm.nih.gov/research/umls/</a>

The ontologies available in English that are purpose built for the domain of medical education, such as TIME and METRO, must first be converted into German. In this case, problems arise in choosing contextually equivalent translations. For instance, the English term "medical education" is equivalent simultaneously to three German terms, namely "medizinischer Ausbildung", "Fortbildung", and "Weiterbildung", which individually refer to the English "graduate", "continuing" and "post-graduate" medical education.

Another problem posed by a translation is the question of whether, and how, a translation can be bound to the original version of the vocabulary in terms of potential developments and updates.

Alternatively, development of an ontology primarily based in the German language is associated with considerable time and effort.

Particularly interesting features among the identified vocabularies, summarized in Table 3, include the fact that all are available online, and at least in part, all utilize direct Internet technologies, such as the XML data format. Although they have not been primarily developed for use in the web, they are at any rate still suitable for this purpose. The technical foundations needed towards Web 3.0 are therefore established [9].

## Limitations

As any literature research, this work is limited by the search and selection of the used publications.

Additionally, existence of structured vocabularies not listed in any of the publications of the searched databases cannot be ruled out.

During preparation of this work, the authors were unable to identify any generally accepted criteria for the description and comparison of structured vocabularies. Therefore, the first step was to identify any medical taxonomy described among the publications of the searched databases. This was followed by a publication analysis based on individual criteria - developed by the respective authors of each publication - used to describe the structured vocabularies. By consensus, the authors of this study then extracted the criteria that seemed most adequate for a comparison of taxonomies for medical training. During criteria selection, the authors have sought to generalize the results, making them suitable for use by German-speaking medical faculties. Alternate or additional criteria not considered in this study may be relevant for local use.

## Conclusion

Content description (semantic annotation) of medical learning resources provides the ability to identify contextually related items across system boundaries. Thus far, there is no suitable taxonomy or ontology in German that may be used to implement a systematic description of resources for Germany's medical students, thus limiting

the potential of the Web 3.0 to support curriculum mapping. Possible approaches include the new development of an appropriate structured vocabulary, the translation of existing English-language vocabularies, or a combination of existing vocabularies sections, possibly with a partial translation of English vocabularies.

## Note

WB is the head of the Learning Center of the Charité, a member of the Medical Education Society (Gesellschaft für Medizinische Ausbildung - GMA), and of the GMA committees "Practical Skills" and "Educational Research Methods." AJ and PS are student tutors of the Learning Center, and students of the Charité. FB is a co-developer of the OBEO. OA is head of the Department of Curricular Management, a member of the GMA, as well as a member of the GMA committee "Educational Research Methods."

## Acknowledgement

The authors thank Dr. Ullrich Woermann, Bern, as well as two anonymous reviewers for their constructive and collegial critic of the manuscript and Rudi Mörgli for translating the manuscript into English.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Holzer M, Pfähler M, Hege I, Fischer M. Wer sucht, der soll finden! - Ein Überblick über Verschlagwortung und Suche medizinischer Lerninhalte. *GMS Med Inform Biom Epidemiol.* 2006;2(3):Doc20. Zugänglich unter/available from: <http://www.egms.de/static/de/journals/mibe/2006-2/mibe000039.shtml>
2. Cook DA, Erwin PJ, Triola MM. Computerized virtual patients in Health Professions Education: a systematic review and meta-analysis. *Acad Med.* 2010;85(10):1589-1602. DOI: 10.1097/ACM.0b013e3181edfe13
3. Junco R, Mastrodicasa J. Connecting to the Net Generation: What Higher Education Professionals Need to Know About Today's Students. Washington, DC: National Association of Student Personnel Administrators; 2007.
4. Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach.* 2001;23(2):123-137. DOI: 10.1080/01421590120036547
5. Ahlers O, Georg W, Blaum W, Stieg M, Hanfler S, Bubser F, Spies C. Der Einsatz einer interdisziplinären, webbasierten Lernzielplattform verbessert sowohl die Unterrichtsqualität als auch die Klausurergebnisse Studierender. Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Bochum, 23.-25.09.2010. Düsseldorf: German Medical Science GMS Publishing House; 2010. Doc10gma13. DOI: 10.3205/10gma013

6. Berners-Lee T, Hendler J, Lassila O. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Sci Am.* 2001;284:34-43. DOI: 10.1038/scientificamerican0501-34
7. Alby T. Web 2.0. Konzepte, Anwendungen, Technologien. München: Hanser Verlag; 2007.
8. Berners-Lee T, Fischetti M. Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventors. New York: HarperBusiness; 2006.
9. Segaran T, Evans C, Taylor J. Programming the semantic web. Sebastopol, CA: O'Reilly & Associates; 2009.
10. Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. *Med Teach.* 2007;29(7):655–659. DOI: 10.1080/01421590701615808
11. Willett TG, Marshall KC, Broudo M, Clarke M. It's about TIME: a general-purpose taxonomy of subjects in medical education. *Med Educ.* 2008;42(4):432-438. DOI: 10.1111/j.1365-2923.2008.03012.x
12. Boeker M, Schober D, Schulz S, Balzer F. Ontology of Bio-Medical Educational Objectives (OBEO): ein Vorschlag für eine Ontologie medizinischer Lernziele. *GMS Med Inform Biom Epidemiol.* 2010;6(2):Doc11. DOI: 10.3205/mibe000111
13. Stenzhorn H, Schulz S, Boeker M, Smith B. Adapting Clinical Ontologies in Real-World Environments. *J Univers Comput Sci.* 2008;14(22):3767-3780.
14. Ellaway R, Evans P, McKillop J, Cameron H, Morrison J, McKenzie H, Mires G, Pippard M, Simpson J, Cumming A, Harden R, Guild S. Cross-referencing the Scottish Doctor and Tomorrow's Doctors learning outcome frameworks. *Med Teach.* 2007;29(7):630-635. DOI: 10.1080/01421590701316548
15. Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. Der praktische Nutzen des Konsensusstatements "praktische Fertigkeiten im Medizinstudium" – eine Validierungsstudie. *GMS Z Med Ausbild.* 2012;29(4):Doc58. DOI: 10.3205/zma000828
16. Willett TG. Current status of curriculum mapping in Canada and the UK. *Med Educ.* 2008;42(8):786-793. DOI: 10.1111/j.1365-2923.2008.03093.x
17. Uschold M, Gruninger M. Ontologies: Principles, Methods and Applications. *Knowl Eng Rev.* 1996;11:93–155. DOI: 10.1017/S0269888900007797

**Corresponding author:**

Wolf E. Blaum  
 Charité - University Medicine Berlin, Department of Curriculum Management, Learning Center, Charitéplatz 1, 10117 Berlin, Germany, Phone: +49 (0)30/450-576003, Fax: +49 (0)30/450/576952  
 wolf.blaum@charite.de

**Please cite as**

Blaum WE, Jarczewski A, Balzer F, Stötzner P, Ahlers O. Auf dem Weg zum Web 3.0: Taxonomien und Ontologien für die medizinische Ausbildung - eine systematische Literaturrecherche. *GMS Z Med Ausbild.* 2013;30(1):Doc13.  
 DOI: 10.3205/zma000856, URN: urn:nbn:de:0183-zma0008568

**This article is freely available from**

<http://www.egms.de/en/journals/zma/2013-30/zma000856.shtml>

**Received:** 2003-06-29

**Revised:** 2012-09-05

**Accepted:** 2012-10-12

**Published:** 2013-02-21

**Copyright**

©2013 Blaum 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.

## **Lebenslauf**

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

## Publikationen

Blaum WE, Schröder T, Dannenberg KA, März M, Hölzer H, Ahlers O, Thomas A. Shame In Medical Education - a randomized study of the acquisition of intimate examination skills. Med Educ 2014; under review

Hautz S, Blaum WE, Feufel MA, Spies, CD.: Comparability of Outcome Frameworks in Medical Education: Implications for framework development. Med Teach 2014; under review.

Ortwein H, Blaum WE, Spies CD: Anesthesiology residents' perspective about good teaching - a qualitative needs assessment. Ger Med Sci. 2014;12:Doc05. doi: 10.3205/000190

Blaum WE, Jarczewski A, Balzer F, Stötzner P, Ahlers O: Towards Web 3.0: Taxonomies and ontologies for medical education – a systematic review. GMS Z Med Ausbild 2013;30(1):Doc13 doi: 10.3205/zma000856

Damanakis A, Blaum WE, Stosch C, Lauener H, Richter S, Schnabel KP: Simulator Network Project Report: A tool for improvement of teaching materials and targeted resource usage in Skills Labs. GMS Z Med Ausbild 2013;30(1):Doc4

Blaum WE, Dannenberg KA, Friedrich T, Jarczewski A, Reinsch AK, Ahlers O. The practical use of the consensus statement on practical skills in medical school – a validation study. GMS Z Med Ausbild. 2012;29(4):Doc58. doi: 10.3205/zma000828

## Danksagung

Ich danke vor allem Frau Univ. Prof. Dr. Claudia Spies für die Anregung zu dem Thema und die Möglichkeit zu dieser Promotion, Ihre engagierte wissenschaftliche und ärztliche Weiterbildung und die finale Korrektur der Dissertationsarbeit. Weiter danke ich meinen Co-Autoren für die konstruktive Zusammenarbeit: Dr. Felix Balzer, Katja Dannenberg, Torsten Friedrich, Anne Jarczewski, Dr. Heiderose Ortwein, Anne Reinsch und Philip Stötzner. Besondere Freude an der Bildungsforschung habe ich auch durch Prof. Dr. Kevin Eva, Dr. Juliane Kämmer, Dr. Claudia Kiessling, Dipl. Psych. Stefan Schauber, Therese Schröder, Fabian Stroben und PD Dr. Anke Thomas gewonnen – Euch allen herzlichen Dank.

Meinen Freunden und vor allem meinen Eltern danke ich für die geduldige Unterstützung meiner Promotion. Auch meiner Tochter Dorle danke ich für die vielen Stunden, die ich mit dem Thema der Arbeit statt mit ihr verbracht habe. Noch mehr aber für all die Zeit, die sie mit mir verbringt.

Besonders danken möchte ich meiner Verlobten, Stefanie Hautz, für ihre ausdauernde und verständnisvolle Unterstützung. Ohne sie wäre ich nie fertig geworden.