

**AUS DEM LEHRSTUHL
FÜR ORTHOPÄDIE
PROF. DR. MED. DR. H.C. J. GRIFKA
DER MEDIZINISCHEN FAKULTÄT
DER UNIVERSITÄT REGENSBURG**

**LEHRE UND WISSENSVERMITTLUNG IN DER ORTHOPÄDIE UNTER
BERÜCKSICHTIGUNG DER MODIFIZIERTEN APPROBATIONSORDNUNG**

- EINE BEFRAGUNG AN 506 STUDENTEN UNTERSCHIEDLICHER SEMESTER -

**Inaugural – Dissertation
Zur Erlangung des Doktorgrades
der Medizin**

**der
Medizinischen Fakultät
der Universität Regensburg**

**vorgelegt von
Klaus Bredl**

2009

**AUS DEM LEHRSTUHL
FÜR ORTHOPÄDIE
PROF. DR. MED. DR. H.C. J. GRIFKA
DER MEDIZINISCHEN FAKULTÄT
DER UNIVERSITÄT REGENSBURG**

**LEHRE UND WISSENSVERMITTLUNG IN DER ORTHOPÄDIE UNTER
BERÜCKSICHTIGUNG DER MODIFIZIERTEN APPROBATIONSORDNUNG**

- EINE BEFRAGUNG AN 506 STUDENTEN UNTERSCHIEDLICHER SEMESTER -

**Inaugural – Dissertation
Zur Erlangung des Doktorgrades
der Medizin**

**der
Medizinischen Fakultät
der Universität Regensburg**

**vorgelegt von
Klaus Bredl**

2009

Dekan:

Prof. Dr. med. B. Weber

1. Berichterstatter:

PD. Dr. med. C. Lüring

2. Berichterstatter:

Prof. Dr. med. W. Jilg

Tag der mündlichen Prüfung:

17. Juli 2009

Inhaltsverzeichnis

1.) Einleitung.....	- 8 -
1.1) Aufbau und Prüfungswesen der medizinischen Ausbildung	- 9 -
1.1.1) Rechtliche Bestimmungen zum Aufbau und Prüfungswesen der medizinischen Ausbildung bis zur Einführung der Approbationsordnung.....	- 9 -
1.1.2) Ärztliche Approbationsordnung im Wandel der Zeit	- 11 -
1.1.3) Der Weg zur aktuellen Approbationsordnung und Zulassungsbestimmungen zum Medizinstudium	- 16 -
1.1.4) Aktuelle Approbationsordnung und grundlegende Veränderungen nach der 9. Novellierung der Approbationsordnung	- 17 -
1.1.5) Allgemeines zum Prüfungswesen	- 24 -
1.1.5.1) Funktionen von Prüfungen	- 24 -
1.1.5.2) Testgütekriterien von Prüfungen	- 25 -
1.1.5.3) Klassifikation kognitiver Prozesse	- 29 -
1.1.6) Verschiedene Prüfungsmethoden	- 31 -
1.1.6.1) Schriftliche Prüfungen.....	- 32 -
1.1.6.2) Mündliche Prüfungen	- 35 -
1.1.6.3) Praktische Prüfungen.....	- 38 -
1.1.7) Schriftliche Prüfungsmethoden im Vergleich	- 39 -
1.1.7.1) Vorteile und Nachteile von MC-Fragen.....	- 40 -
1.1.7.2) Vorteile und Nachteile von Offenen Fragen	- 47 -
1.2) Zielsetzung der Arbeit	- 51 -
2.) Material und Methode.....	- 53 -
2.1) Problemstellung.....	- 53 -
2.2) Testentwicklung und Fragebogenerstellung	- 54 -
2.2.1) Teil I des Fragebogens (Zufriedenheit der Studierenden)	- 54 -
2.2.2) Teil II des Fragebogens (Wissensstand im Fachbereich Orthopädie).....	- 55 -
2.2.3) Teil III des Fragebogens (Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)	- 57 -

2.3) Untersuchungsaufbau	- 57 -
2.4) Untersuchungsgut.....	- 58 -
2.5) Durchführung der Befragung	- 59 -
2.6) Auswertung und Analyse.....	- 60 -
2.6.1) Auswertung von Teil I (Zufriedenheit der Studierenden)	- 60 -
2.6.2) Auswertung von Teil II (MC-Fragen).....	- 60 -
2.6.3) Auswertung von Teil II (Offene Fragen)	- 61 -
2.6.4) Auswertung von Teil III (Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)	- 63 -
2.6.5) Auswertung des Vergleichs von MC-Fragen und Offene Fragen	- 64 -
2.6.6) Auswertung des Vergleichs von Prä-Test und Post-Test der Universitäten Regensburg und Heidelberg.....	- 64 -
3.) Ergebnisse	- 65 -
3.1) Rücklauf.....	- 67 -
3.2) Ergebnisse der Fragebogenauswertung für Teil I (Zufriedenheit der Studierenden)	- 67 -
3.3) Ergebnisse der Fragebogenauswertung für Teil II (MC-Fragen zum Wissensstand im Fachbereich Orthopädie).....	- 84 -
3.4) Ergebnisse der Fragebogenauswertung für Teil II (Offene Fragen zum Wissensstand im Fachbereich Orthopädie)	- 88 -
3.5) Ergebnisse der Fragebogenauswertung für Teil III (Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)	- 95 -
3.6) Vergleich der Ergebnisse von Prä- und Post-Test im Praktikum Orthopädie.....	- 103 -
3.7) Ergebnisse für die Multiple-Choice-Fragen im Einzelnen und differenzierte Betrachtung aller Fragen nach inhaltlichen Kriterien	- 105 -

4.) Diskussion	- 114 -
4.1) Zum Stand der Literatur	- 114 -
4.2) Prüfungswesen in der Diskussion	- 117 -
4.3) Vorbemerkungen	- 118 -
4.4) Diskussion der Ergebnisse	- 120 -
4.4.1) Diskussion der Ergebnisse zu Teil I.....	- 120 -
4.4.2) Diskussion der Ergebnisse zu Teil II (MC-Fragen).....	- 126 -
4.4.3) Diskussion der Ergebnisse zu Teil II (Offene Fragen)	- 129 -
4.4.4) Diskussion der Ergebnisse zu Teil III.....	- 134 -
4.4.5) Diskussion der Ergebnisse: Vergleich MC-Fragen und Offene Fragen....	- 140 -
4.4.6) Diskussion der Ergebnisse für die MC-Fragen im Einzelnen	- 143 -
4.4.7) Diskussion der Ergebnisse: Vergleich Prä-Test und Post-Test der unterschiedlichen Universitäten	- 146 -
4.5) Diskussion der Ergebnisse von Prä- und Posttest im Blockpraktikum Orthopädie	- 149 -
4.6) Zusammenfassung	- 152 -
5.) Literaturverzeichnis	- 155 -
6.) Tabellenverzeichnis	- 166 -
7.) Diagrammverzeichnis	- 168 -
8.) Abkürzungsverzeichnis	- 171 -
9.) Anhang.....	- 172 -
10.) Danksagung.....	- 181 -
11.) Lebenslauf	- 182 -

1.) Einleitung

Das Schlagwort „Qualität der Lehre“ gehört seit längerem zum begrifflichen Inventar in den politischen Diskussionen. Dabei wird die „Qualität der Lehre“ oft als Postulat dargestellt, was einen Qualitätsmangel impliziert (27). Auch in der Öffentlichkeit wird auf Grund der bis vor kurzem andauernden Ärzte-Streiks in den Kliniken Deutschlands und auf Grund der neuesten inhaltlichen, strukturellen und prüfungsrelevanten Änderungen des Medizinstudiums verstärkt über das Gesundheitswesen und im speziellen über die medizinische Ausbildung diskutiert.

Diese Diskussionen nahmen bereits 1869 mit der Vereinheitlichung der medizinischen Ausbildung durch die Reichsgewerbeordnung ihren Anfang. Auch 1970 mit der Einführung der Approbationsordnung kamen sie zu keinem Ende und spiegeln auch heute im Jahr 2006 nach der 9. Änderung der Ärztlichen Approbationsordnung die Problematik wider, die medizinische Ausbildung in adäquater Art und Weise dem Wandel der gesellschaftlichen, wissenschaftlichen und technischen Gegebenheiten anzupassen (89).

Am stärksten kristallisiert sich dabei das Problem des Prüfungswesens heraus, das nach der 9. Änderung der Ärztlichen Approbationsordnung in dem Begriff des sog. „Hammerexamens“ seine Klimax findet. Dabei wurde das Prüfungssystem in der Mediziner Ausbildung, das in Gesellschaft und didaktischen Funktionen einen hohen Stellenwert einnimmt (84, 85) und von den Veränderungsprozessen und Reformbestrebungen in großem Umfang betroffen ist, vielfach hinsichtlich Strukturierung, Dauer, Inhalt, Art, Zusammensetzung einer Prüfung aus mehreren Teilprüfungen und Prüfungsmethoden geändert. Auf Grund dieser Tatsache sind die verschiedenen Prüfungsmethoden auch Objekt vieler Untersuchungsreihen, die die einzelnen Methoden im Hinblick auf ihre Qualität analysieren. Abgesehen von der für die Studierenden wohl bedeutsamsten Frage, welche Anzahl an Staatsexamina nun am sinnvollsten sei, tritt dabei auch der Vergleich zwischen den in den USA entwickelten, seit 1950 weltweit eingeführten und immer stärker verbreiteten Multiple-Choice-Fragen (MC-Fragen) und den Offenen Fragen in den Vordergrund.

Im folgenden Kapitel wird sowohl der geschichtliche Hintergrund und aktuelle Stand der rechtlichen Bestimmungen zur medizinischen Ausbildung in Deutschland abgehandelt, als auch auf die verschiedenen Prüfungsmethoden mit ihren qualitativen Vor- und Nachteilen eingegangen.

1.1) Aufbau und Prüfungswesen der medizinischen Ausbildung

1.1.1) Rechtliche Bestimmungen zum Aufbau und Prüfungswesen der medizinischen Ausbildung bis zur Einführung der Approbationsordnung

Durch die gesetzlichen Vorgaben zur ärztlichen Ausbildung werden die Zulassungsbedingungen, die Dauer und der Aufbau des Studiums, der Lehrplan und die Prüfungsordnungen reglementiert. Mit jeder bisherigen gesetzlichen Neuordnung kam es auch zu einer Novellierung der Prüfungsordnung, die sich auf Art der Prüfungsmethoden, Zeitpunkt der Prüfung, Inhalt der Prüfung, Zulassungsbestimmungen und Bestehensgrenzen, Anzahl der geprüften Fächer, Benotungspraxis und auf die Durchführung der Prüfungsabschnitte bezog.

Eine chronologische Übersicht über die gesetzlichen Neuregelungen bis zur 7. Novellierung der Approbationsordnung im Jahr 1989 wurde von J. Grifka in seinem Beitrag „Stellenwert der Lehre in der Medizinerbildung“ angefertigt (27):

Mit der Reichsgewerbeordnung von 1869 wurde angesichts der stetig steigenden Anzahl von Medizinstudenten eine reichseinheitliche Regelung der Ausbildung zum Arzt erstellt. Ebenso legte man damit fest, dass der Bundesrat „Vorschriften über den Nachweis der Befähigung“ erlässt und dass es zur Niederlassung als Arzt im deutschen Reich der Approbation bedurfte (10, 36). Die erste reichseinheitliche Prüfungsordnung wurde 1883 erlassen. Deren Bestimmungen zu Folge dauerte das Studium neun Semester und fand mit dem Staatsexamen seinen Abschluss. Die Approbation erfolgte unmittelbar an das Staatsexamen.

Die nächste Novellierung der Prüfungsordnung fand 1901 statt. Die Zulassungsbestimmungen änderten sich darin, dass nun nicht mehr nur den Absolventen von Gymnasien, sondern auch den Absolventen von Realschulen der Zugang zum Medizinstudium erlaubt war. Ebenso verlängerte sich das Studium auf zehn Semester und ein aus sieben Prüfungsabschnitten bestehendes Staatsexamen wurde eingeführt. Im Anschluss an die zehn Semester folgte eine einjährige praktische Tätigkeit als Medizinalpraktikant (bis 1939), in der vier Monate mit der Behandlung von inneren Krankheiten vorgeschrieben waren (10). Dies führte letztendlich zu einer Verlängerung des Studiums von 4 ½ auf 6 Jahre. Den Inhalt des Studiums betreffend versuchte man bereits damals mit dieser Novellierung den Diskussionen um das

richtige Verhältnis von Theorie und Praxis gerecht zu werden. Außerdem kam es auch zur Eingliederung neuer Fachgebiete und Prüfungsfächer in den Unterricht.

Zu einer erneuten Änderung der Prüfungsordnung kam es 1924. Darin wurde erstmalig eine Trennung des Studiums in vier Semester Vorklinik und sechs Semester Klinik und die Erweiterung des Staatsexamens von sieben auf 14 Prüfungsabschnitte beschlossen (10, 36).

1927 kam es mit der Verlängerung der Vorklinik von vier auf fünf Semester zu einer Neufassung des Studienablaufs. Die sechs Semester Klinik wurden beibehalten. 1933 novellierte man teilweise die Prüfungsordnung, indem eine Teilung der Vorprüfung in einen naturwissenschaftlichen und einen anatomisch-physiologischen Abschnitt vorgenommen wurde (10, 36).

Basierend auf der Reichsärzteordnung von 1935 erfolgte 1939 mit der Bestallungsordnung als „Verordnung zur Durchführung und Ergänzung der Reichsärzteordnung“ eine grundlegende Veränderung der Mediziner Ausbildung. Die Dauer des Studiums wurde wieder auf vier Semester Vorklinik und sechs Semester Klinik gekürzt. Ab sofort musste ein Krankenpflegedienst von sechs Monaten abgeleistet werden, dessen Dauer 1942 auf vier Monate reduziert wurde. Neben diesem Krankenpflegedienst musste während des Studiums sowohl eine Fabrik- oder Landdienstzeit als auch eine Pflichtfamulatur von je sechs Wochen absolviert werden. Zu den Neuerungen zählten auch die Einführung von Pflichtvorlesungen und Pflichtübungen sowie die Erweiterung des Staatsexamens von 14 auf 19 Prüfungsfächer. Nach Beendigung des Studiums folgten eine Pflichtassistentenzeit und ein Landvierteljahr, die den Absolventen auf die Ausübung des ärztlichen Berufes in eigener Praxis vorbereiten sollte. Während des Zweiten Weltkrieges wurden zahlreiche neue Änderungen erlassen, welche aber nicht nur auf Grund der damaligen politischen Situation, sondern auch wegen der unmöglichen Umsetzung in die Praxis von den Universitäten nicht eingehalten werden konnten (10, 47).

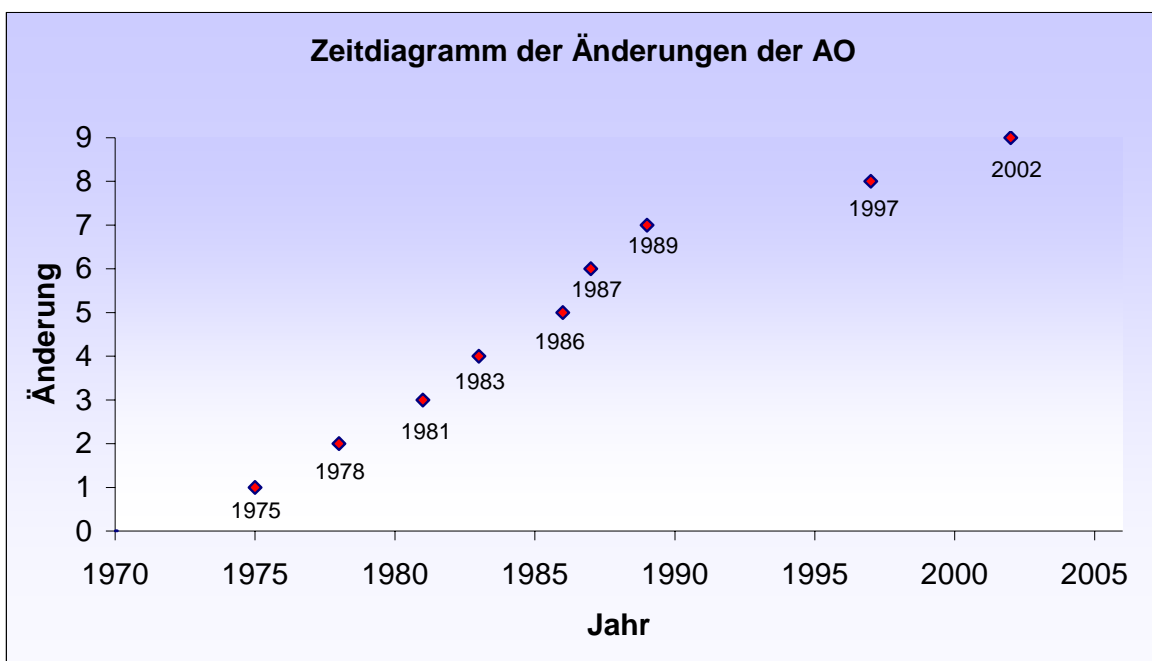
Beschlossen durch die Bestallungsordnung der Bundesrepublik von 1953 kam es zur vollständigen Übertragung der Studienplangestaltung und der Durchführung der Prüfungen auf die Fakultäten, die auch eine möglichst geringe Zahl der Pflichtvorlesungen umsetzen sollten. Die Dauer des Studiums wurde mit fünf Semestern Vorklinik und sechs Semestern Klinik festgelegt, der Krankenpflegedienst mit acht Wochen und die Famulatur mit drei Monaten. Das Staatsexamen setzte sich aus 19 Prüfungsfächern zusammen und bestand aus praktisch-mündlichen Prüfungen,

deren Benotung nach dem Schulnoten-Prinzip erfolgte. Anschließend an das Medizinstudium folgte eine zweijährige Vorbereitungszeit als Medizinalassistent, wobei mindestens sechs Monate Innere Medizin und mindestens je vier Monate Gynäkologie und Chirurgie abzuleisten waren (11). Die Bestallungsordnung von 1953 wurde durch zahlreiche Änderungen in den Jahren 1955, 1957, 1958, 1963 und 1965 ergänzt, bis 1970 mit der Einführung der Approbationsordnung für Ärzte (AppOÄ) eine wesentliche Neuordnung auftrat.

1.1.2) Ärztliche Approbationsordnung im Wandel der Zeit

Noch vor Einführung der Approbationsordnung formulierte 1968 der Ministerialdirigent im Bundesministerium für Gesundheitswesen, Herr Dr. med. Reinhard Rachold, bei seinem Einführungskapitel zur Bestallungsordnung für Ärzte: „Wir sind uns dabei bewusst, dass das künftige Niveau des deutschen Arztes nicht entscheidend von diesen Bestimmungen abhängt, sondern von den Lehrern, in deren Hände die Ausbildung gelegt ist (10).“

Diagramm 1: Zeitdiagramm der Änderungen der ärztlichen Approbationsordnung



Am 28.10.1970 erfolgte mit der Einführung der Approbationsordnung für Ärzte (AppOÄ) eine völlige Neuregelung der medizinischen Ausbildung, die im Wesentlichen am 01.10.1972 in Kraft trat (5). Diese beinhaltete auch detaillierte Übergangsregelungen, deren Umsetzung teilweise mehrere Jahre dauerte (54). Mit der Aufteilung des Studiums in zwei Jahre Vorklinik und drei Jahre Klinik, der Einführung des Praktischen Jahres, dem Verzicht auf Pflichtvorlesungen und der Einführung von Pflichtpraktika sowie der Einführung von schriftlichen Multiple-Choice-Prüfungen, die erstmals im Sommersemester 1974 durchgeführt wurden, ist die Approbationsordnung von 1970 in den meisten ihrer damals entscheidenden Punkte auch heute noch gültig. Zu den damaligen Änderungen gehörten auch die Einführung von Studienjahren anstatt von Semestern, das verpflichtende Ableisten von zwei Monaten Krankenpflegedienst vor Ablegen der ärztlichen Vorprüfung und von zwei Monaten Famulatur vor Antritt zum 2. Abschnitt der ärztlichen Prüfung. Die Prüfungen wurden lediglich mit bestanden oder nicht bestanden bewertet, wobei die Bestehensgrenze für die MC-Fragen bei 50% festgesetzt wurde. Um schon damals dem Verlangen nach einer mehr praxisorientierten Ausbildung nachzukommen, wurde in Paragraph 2 Abs. 2 geregelt, dass der Unterricht, soweit zweckmäßig, nicht am einzelnen Fachgebiet, sondern am Lehrgegenstand ausgerichtet und deshalb problemorientiert durchgeführt werden sollte (5). In Mainz wurde das „Institut für medizinische und pharmazeutische Prüfungsfragen“ (IMPP) gegründet, der eine vertragliche Regelung der Bundesländer zu Grunde liegt. In den Zuständigkeitsbereich des IMPP fallen die Erstellung und fortlaufende Bearbeitung der Prüfungsfragen sowie die Herstellung, Versendung und Auswertung der Fragebögen. Diese Aufgaben sind in Artikel 2 des „Abkommens über die Einrichtung und Finanzierung des IMPP vom 14.10.1970“ reglementiert. Nach der Auflösung der örtlichen Prüfungsausschüsse der Hochschulen wurden deren Kompetenzen den neugegründeten Landesprüfungsämtern übertragen (8, 54). Als kritischer Punkt des von der Lehrvermittlung losgelösten Lernverhaltens wurde vor allem der Inhalt von Paragraph 2 Satz 1 gesehen. Darin heißt es: „Die Hochschule hat eine Ausbildung zu vermitteln, die es dem Studierenden ermöglicht, den Wissensstoff und die Fähigkeiten zu erwerben, die in den in dieser VO vorgesehenen Prüfungen gefordert werden“ (5). Der Marburger Bund nahm dazu folgende Stellungnahme: „Die Ausbildung des Arztes an der Universität wird damit auf das Niveau einer reinen Prüfungsvorbereitung im Stil eines Repetitoriums festgelegt. Anleitung zum selbstständigen, wissenschaftlichen Denken ist in diesem Ausbildungsgang nicht mehr möglich (64). Die 1. Änderung der Approbationsordnung erfolgte am 21.05.1975, in der

eine verpflichtende Famulatur in der Praxis beschlossen wurde. Am 24.02.1978 wurde anhand der 2. Änderung der Approbationsordnung die Dauer der Famulatur auf vier Monate erhöht, ein Leistungsnachweis für Pflichtpraktika eingeführt und die Bestehensgrenze von 50% auf 60% angehoben, welche aber erst im Herbst 1979 verwirklicht wurde. Ebenso setzte man eine Klausel fest, mit der man sich die Möglichkeit der Wiedereinführung von Pflichtvorlesungen beibehielt. Nach dem Skandalphysikum im März 1981 reagierte man darauf mit der 3. Änderung der Approbationsordnung vom 15.07.1981, in der die Bestehensgrenze wieder auf 50% gesenkt wurde. Mit der 4. Änderung der Approbationsordnung vom 19.12.1983 erfolgte eine Wiedereinführung der Benotung der Staatsexamina. Grundlegendere Veränderungen traten mit der 5. Änderung der Approbationsordnung vom 15.12.1986 auf. Die schriftliche Prüfung nach dem Praktischen Jahr wurde abgeschafft, gleichzeitig wurde die Einführung des „Arzt im Praktikum“ (AiP) mit einer Dauer von zwei Jahren durchgesetzt. Unterricht in kleineren Gruppen und Bed-Side-Teaching mit maximal fünf Studenten wurden ebenfalls gesetzlich geregelt. Kurze Zeit später folgte am 28.05.1987 die 6. Änderung der Approbationsordnung, deren Novellierung die Verschiebung des AiP um ein Jahr mit Dauer von 18 Monaten, die Ermittlung durchschnittlicher Prüfungsleistungen und eine neue Wiederholungsregelung beinhaltete. Da die 5. und 6. Änderung in so einem kurzen Zeitverhältnis aufeinander folgten, wurde am 14.07.1987 eine Neufassung der Approbationsordnung erstellt, in der beide berücksichtigt wurden. Bei der Vorstellung der 7. Änderung der Approbationsordnung wurde betont, dass diese Verordnung lediglich als Sofortmaßnahme zu verstehen sei, um offensichtliche Mängel des Medizinstudiums schnellstmöglich zu beheben (58). Gesetzlich verabschiedet wurde die 7. Änderung der Approbationsordnung am 21.12.1989. Darin wurde als Paragraph 1 eine genaue Zieldefinition aufgenommen, wonach das Studium auf einen eigenverantwortlich und selbstständig handelnden Arzt abzielt (77). Ebenso wurde beschlossen, dass in der Vorklinik nur noch eine begrenzte Teilnehmerzahl in den Seminaren zulässig ist und im klinischen Abschnitt die Patientendemonstrationen mit maximal acht Studenten und die Patientenuntersuchungen mit maximal drei Studenten durchgeführt werden dürfen. Bereits vor der gesetzlichen Verabschiedung der 7. Änderung der Approbationsordnung wurde vom damaligen Bundesministerium für Jugend, Familie, Frauen und Gesundheit eine Sachverständigenkommission einberufen, die 1993 nach 3 ½ Jahren dem Bundesministerium für Gesundheit eine Kurzfassung der Arbeit mit neuen Ideen für eine grundlegende Reform des Medizinstudiums vorlegte (43). In

seiner Eröffnungsrede des 96. Deutschen Ärztetages im Mai 1993 erklärte Bundesminister Horst Seehofer: „Ich sage Ihnen heute zu, dass es noch in dieser Legislaturperiode eine Neufassung der Approbationsordnung geben wird“ (3). Letztendlich wurde die 8. Änderung der Approbationsordnung am 17.12.1997 von der Bundesregierung beschlossen. Der Stand nach der 8. Änderung der Approbationsordnung wird im Folgenden detaillierter dargestellt, um den Vergleich mit der aktuellen Approbationsordnung nach der 9. Änderung im nächsten Unterkapitel genauer ziehen zu können.

Das Ziel der Neufassung war (17),

- I) Die Ausbildung an die veränderten Anforderungen in der gesundheitlichen Versorgung anzupassen;
- II) Verzahnung von klinischen und theoretischen Ausbildungsinhalten über die gesamte Studienzeit zu gewährleisten;
- III) Die Ausbildungsinhalte auf das Wissen zu konzentrieren, welches der Arzt als Grundlage für seine spätere berufliche Tätigkeit und die Fähigkeit zur Weiterbildung braucht;
- IV) Die berufspraktische Ausbildung grundlegend zu verbessern, wobei die Zahl der Studenten mit den verfügbaren Ausbildungskapazitäten in Übereinstimmung gebracht werden muss.

Die Umsetzung von Punkt I) sollte durch eine verstärkte Berücksichtigung der ganzheitlichen Sicht des Patienten, eine verbesserte Zusammenarbeit zwischen den Versorgungssektoren und den Berufsangehörigen, eine verbesserte Koordination der Behandlung im ambulanten Bereich und eine größere Bedeutung für die Komponenten Prävention, Gesundheitsförderung, Medizin des Alterns und Gesundheitsökonomie erreicht werden.

Das in Punkt II) vorgegebene Ziel realisierte man durch die gesetzliche Einführung von mindestens 126 Stunden an Seminaren mit klinischen Inhalten in den ersten vier Semestern. Somit wurde die Einbeziehung klinischer Inhalte in das Grundlagenstudium verwirklicht.

Hinsichtlich Punkt III) führte man eine grundlegende Veränderung der Staatsprüfung am Ende des Studiums durch. An die Stelle der Prüfung fächerspezifischen Wissens trat die Prüfung fächerübergreifenden und gegenstandsbezogenen Wissens.

Anhand von Punkt IV) kam es zur Verbesserung des Unterrichts am Krakenbett (UaK) als zentrales praxisbezogenes Ausbildungselement. Die strukturelle Umsetzung wurde durch die Verringerung der Gruppengröße der Studierenden, die am Patienten ausgebildet werden, von drei auf zwei bei der Patientenuntersuchung und von acht auf sechs bei der Patientendemonstration verbindlich gewährleistet. Weiterer Bestandspunkt war die Einführung von Blockpraktika, die den verstärkten Praxisbezug unterstützen sollte.

Die Strukturierung des Studiums betreffend umfasste die ärztliche Ausbildung nach geltendem Recht (§1 Abs.2 ÄAppO) ein Hochschulstudium der Medizin von mindestens sechs Jahren, aufgeteilt in einen vorklinischen Teil von zwei Jahren und einen klinischen Teil von vier Jahren, wobei das letzte Jahr des Studiums eine zusammenhängende praktische Ausbildung in Krakenanstalten von 48 Wochen beinhaltete (§3 ÄAppO). Verpflichtend waren eine Ausbildung in Erster Hilfe (§5 ÄAppO), ein Krankenpflegedienst von zwei Monaten (§6 ÄAppO), eine Famulatur von vier Monaten (§7 ÄAppO) und eine auf das Studium folgende 18monatige Tätigkeit als Arzt im Praktikum (§34a-d ÄAppO). Die ärztliche Ausbildung umfasste folgende staatliche Prüfungen (§10 ÄAppO): Die Ärztliche Vorprüfung nach einem Studium von mindestens zwei Jahren und die Ärztliche Prüfung, die in drei Abschnitte aufgeteilt war. Der Erste Abschnitt erfolgte nach einem Studium der Medizin von einem Jahr nach Bestehen der Ärztlichen Vorprüfung, der Zweite Abschnitt nach Bestehen des Ersten Abschnitts und einem Studium der Medizin von drei Jahren nach Bestehen der Ärztlichen Vorprüfung und der Dritte Abschnitt nach einem Studium der Medizin von einem Jahr nach Bestehen des Zweiten Abschnitts der Ärztlichen Prüfung. Diese Prüfungen legte man nach der ÄAppO vor den von den Bundesländern eingerichteten Landesprüfungsämtern für Medizin ab (§9 ÄAppO). Die Ärztliche Vorprüfung setzte sich aus einem schriftlichen und einem mündlichen Teil zusammen, der Erste Abschnitt der Ärztlichen Prüfung wurde schriftlich, der Zweite Abschnitt schriftlich und mündlich und der Dritte Abschnitt mündlich und praktisch durchgeführt (§13 ÄAppO).

Ebenso verabschiedete man eine rechtliche Grundlage für die Durchführung von Modellstudiengängen in der medizinischen Ausbildung. Anhand dieser Modellklausel vom 11.02.1999 konnten innovative Entwicklungen und alternative Studiengänge auf den Weg gebracht werden, auch wenn eine Abweichung zu einigen ansonsten zwingenden Vorgaben der Approbationsordnung bestand. Vorreiter dieser Neuregelung war die Charité in Berlin, deren entwickeltes Pilotprojekt sich an

internationalen Vorbildern wie Harvard und der kanadischen Reform-Universität McMaster orientierte. Die private Universität Witten/Herdecke und die Universitäten in Aachen, Heidelberg, Köln und Bochum folgten kurz darauf mit einem eigenen Modellstudiengang.

Bereits kurz nach Verabschiedung der 8. Änderung der Approbationsordnung lag dem damaligen Bundesgesundheitsminister Horst Seehofer ein Entwurf der 9. Änderung der Approbationsordnung vor. Dieser basierte auf einer Arbeitsgemeinschaft bestehend aus dem Hochschulausschuss der Kultusministerkonferenz und der Zentralen Vergabestelle für Studienplätze (ZVS). Diesem Änderungsvorschlag der Bund-Länder-Kommission wurde bald ein Gegenentwurf des Präsidiums des Medizinischen Fakultätentages (MFT) entgegengestellt, was die Umsetzung einer erneuten Novellierung hinauszögerte (81).

1.1.3) Der Weg zur aktuellen Approbationsordnung und Zulassungsbestimmungen zum Medizinstudium

2001 kündigte die damalige Bundesgesundheitsministerin Ulla Schmid an, dass noch in der laufenden Legislaturperiode (also im Jahr 2002) die 9. Verordnung zur Änderung der Approbationsordnung für Ärzte (ÄAppO) verabschiedet werde. Vor dem Deutschen Ärztetag am 22.05.2001 in Ludwigshafen sagte Frau Schmid: „Bundeskanzler Gerhard Schröder wird die Reform des Medizinstudiums zur Chefsache machen“ (81). Am 26.04.2002 stimmte das Bundesratsplenium der 9. Änderung der Approbationsordnung mit großer Mehrheit zu und am 27.06.2002 erging der Beschluss der neuen ÄAppO. Damit konnte das Medizinstudium nach der neuen Ausbildungsordnung zum Wintersemester 2003/2004 an den Medizinischen Fakultäten begonnen werden, nachdem der frühere Novellierungsversuch 1997 wegen des Widerstandes der Kultusminister und zuletzt der Finanzminister der Länder gescheitert war. Mit dieser Neuordnung der medizinischen Ausbildung sollte eine der umfangreichsten Reformen des Studiums seit 1970 bewirkt werden (69).

Auch die Zulassungsbestimmungen zum Medizinstudium wurden geändert. Seit Beginn des Wintersemesters 2005/2006 bestehen folgende Zulassungsquoten für Medizin: 20% Abiturnotenbeste, 20% Wartezeit und 60% Auswahl durch die jeweilige Hochschule. Diese Quoten gelten nach Abzug der Vorabquoten für Ausländer,

Härtefälle, Zweitstudienbewerber, zukünftige Sanitätsoffiziere und die besondere Hochschulzulassung. Zum Wintersemester 2005/2006 wurden in der Medizin 4,4-mal mehr Bewerber/-innen als Studienplätze (Studienplatzanzahl = 8413) verzeichnet (52).

1.1.4) Aktuelle Approbationsordnung und grundlegende Veränderungen nach der 9. Novellierung der Approbationsordnung

Im Folgenden werden die wichtigsten novellierten rechtlichen Bestimmungen der aktuellen Approbationsordnung dargestellt und in Vergleich mit den bisherigen gestellt (6, 7):

Erster Abschnitt: Die Ärztliche Ausbildung (§§ 1-7)

§1 - Ziele und Gliederung der ärztlichen Ausbildung

In Absatz 1 werden nun die Teilbereiche des Ausbildungszieles deutlicher formuliert. Ziel ist die Vermittlung von Kenntnissen, Fähigkeiten und Fertigkeiten, die für eine umfassende Gesundheitsversorgung der Bevölkerung erforderlich sind. Ebenso wurden ergänzend zu den bisherigen Fächern interdisziplinäre Querschnittsbereiche in die Ausbildung integriert, die für alle Studenten obligatorisch sind. Die Studenten setzen sich jetzt auch mit dem Querschnittsbereich „Geschichte, Theorie, Ethik der Medizin“ (GTE) auseinander, in dem sie anhand eines benoteten Scheins ihr Wissen nachweisen müssen (80).

Die Strukturierung des Studiums wurde in Absatz 2 neu reglementiert. Der Krankenpflegedienst wurde von zwei auf drei Monate verlängert und es gibt jetzt neue Regelungen zur Teilbarkeit des Praktikums. Wichtigster Änderungspunkt ist die Zusammenfassung der bisherigen drei Teile der Ärztlichen Prüfung zu einer Prüfung, auf die im nächsten Absatz genauer eingegangen wird. In Absatz 3 wird die Reglementierung der ärztlichen Prüfung abgehandelt. Dort heißt es:

„Die Ärztliche Prüfung nach Absatz 2 Nr.6 wird abgelegt:

- 1) der Erste Abschnitt der Ärztlichen Prüfung nach einem Studium der Medizin von zwei Jahren

und

2) der Zweite Abschnitt der Ärztlichen Prüfung nach einem Studium der Medizin von vier Jahren einschließlich eines Praktischen Jahres nach Absatz 2 Satz 1 Nr. 1 nach Bestehen des Ersten Abschnitts der Ärztlichen Prüfung. Die in § 27 genannten Fächer und Querschnittsbereiche werden von den Universitäten zwischen dem Ersten Abschnitt der Ärztlichen Prüfung und dem Beginn des Praktischen Jahres geprüft“ (7).

Somit fällt der Termin für den Zweiten Abschnitt der Ärztlichen Prüfung auf die Zeit nach dem Praktischen Jahr. Erstmals werden auch die neuen Querschnittsbereiche erwähnt, sowie der Zeitraum für die Prüfung des Lernerfolges in ihnen und den übrigen Fächern festgelegt.

§ 2 – Unterrichtsveranstaltungen

Auf Grund von Absatz 1 werden nun die Vorlesungen in den Hintergrund gedrängt und mehr Gewichtung auf die Praxis und die Seminare gelegt. Für die Hochschulen bedeutsam ist, dass diese nicht mehr bei der Ankündigung der Veranstaltungen darauf hinweisen müssen, dass der Besuch dieser Unterrichtsveranstaltung das Erreichen des Ausbildungszieles fördert. Die in Anlage 1 festgelegte Dauer der zu leistenden Veranstaltungen wird von 624 auf 630 Stunden nur leicht erhöht, jedoch wird die bisherige starre zeitliche Regelung für die einzelnen Veranstaltungen verlassen. Absatz 2 ist prinzipiell neu. Darin wird der Schwerpunkt des fachübergreifenden Denkens neu erwähnt und vor allen anderen als explizites Unterrichtsziel benannt. Der Unterricht im Studium soll fachübergreifendes Denken fördern und soweit zweckmäßig problemorientiert am Lehrgegenstand ausgerichtet sein. Auch wird die Einführung integrierter Lehrveranstaltungen und Seminare von insgesamt 154 Stunden vor dem ersten Abschnitt der Ärztlichen Prüfung zur Vermittlung klinischer Inhalte und Bezüge festgelegt, deren Inhalt einschließlich Teilnahme- und Erfolgsnachweis in § 2 Absatz 7 Satz 1 geregelt ist. Die Gruppengröße beim sog. Bed-Side-Teaching wird mit sechs Studierenden festgesetzt. In Absatz 3 und 4 ist die Stundenzahl am Krankenbett mit 476 Stunden reglementiert und die Blockpraktika werden in Dauer, Inhalt und Anteil an den praktischen Übungen auf mindestens 20% für den klinischen Teil festgelegt. Absatz 8 schreibt vor, dass bis zum Ersten Abschnitt der Ärztlichen Prüfung und bis zum Beginn des Praktischen Jahres jeweils ein Wahlfach belegt werden muss. Dabei kann für den ersten Abschnitt aus den hierfür angebotenen Wahlfächern der Universität frei gewählt werden, für den Zweiten Abschnitt ist ein Stoffgebiet zu wählen,

dass in Anlage 3 aufgelistet ist und auch von der Universität angeboten wird. Absatz 9 schreibt vor, dass die Lehrveranstaltungen regelmäßig auf ihren Erfolg zu evaluieren sind und die Ergebnisse bekannt gegeben werden müssen. Mit diesem neu eingeführten Absatz wird vor allem durch den Zwang zur Veröffentlichung der Ergebnisse der Wettbewerb zwischen den Hochschulen gefördert. Somit wird die Beurteilung der Lehre durch die Studierenden zu einem wichtigen Instrument bei der Werbung der Studierenden für die Hochschulen, da diese nach dem 2005 ausgelaufenen Staatsvertrag über die zentrale Vergabe von Studienplätzen trotz mangelhafter Betreuung und folglich schlechter Bewertung keine Garantie mehr haben, genügend Studienanfänger zugewiesen zu bekommen, was sich auch negativ auf den Erhalt der Personaldecke auswirkt.

§ 3 – Praktisches Jahr

Laut Absatz 2 Satz 2 ist es möglich, das Praktische Jahr aus wichtigen Gründen zu unterbrechen. Die bereits abgeleisteten Teile werden angerechnet, sofern sie nicht länger als zwei Jahre zurückliegen.

§ 4 – Durchführung des Praktischen Jahres in außeruniversitären Einrichtungen

Die wichtigste Neuerung dieses Paragraphen ist, dass eine stärkere Einbeziehung außeruniversitärer Einrichtungen in das Praktische Jahr vollzogen wurde. Die Innere und Chirurgie-Abteilungen müssen nur noch 60 Behandlungsplätze mit unterrichtsgerechten Patienten vorweisen, statt der bisher geforderten 80 Betten.

§ 6 – Krankenpflagedienst

Hier ist in Absatz 1 eine Änderung vorgenommen worden. Die bisherige Dauer des Krankenpflagedienstes wurde von zwei Monaten auf drei Monate verlängert.

§ 13 – Art und Bewertung der Prüfung

Bei den Bewertungen der Leistungen sind in Absatz 2 folgende Prüfungsnoten zu verwenden: „sehr gut“ (1) = eine hervorragende Leistung, „gut“ (2) = eine Leistung, die erheblich über den durchschnittlichen Anforderungen liegt, „befriedigend“ (3) = eine Leistung, die in jeder Hinsicht durchschnittlichen Anforderungen gerecht wird,

„ausreichend“ (4) = eine Leistung, die trotz ihrer Mängel noch den Anforderungen genügt und „nicht ausreichend“ (5) = eine Leistung, die wegen erheblicher Mängel den Anforderungen nicht mehr genügt. Eine Bewertung mit „ungenügend“ (6) = eine unbrauchbare Leistung, ist nicht mehr möglich.

In Absatz 3 wurde der Umstand des Bestehens einer Prüfung aus schriftlichem und mündlich-praktischem Teil novelliert. Ein Abschnitt gilt nur dann als bestanden, wenn sowohl der schriftliche als auch der mündlich-praktische Teil bestanden sind. Somit ist die Regelung, dass ein schlechter „mangelhafter“ (Note 5) Prüfungsteil mit einem „guten“ (Note 2) oder „sehr guten“ (Note 1) Prüfungsteil ausgeglichen werden kann, obsolet. Bei Nichtbestehen eines Prüfungsteiles muss und darf nur noch der nichtbestandene Prüfungsteil wiederholt werden. Außerdem ist es nicht erlaubt, Scheine aus dem klinischen Abschnitt schon vor Bestehen des ersten Abschnitts zu erwerben. Ausnahme hierbei bilden die Modellstudiengänge, die vom Ersten Abschnitt der Ärztlichen Prüfung befreit sind.

Dritter Abschnitt: Die Ärztliche Prüfung (§§ 22-33)

Der Erste Abschnitt der Ärztlichen Prüfung (§§ 22-26)

§ 22 – Inhalt des Ersten Abschnitts der Prüfung

In Absatz 1 wird der schriftliche Teil des Ersten Abschnitts der Ärztlichen Prüfung abgehandelt. Dabei wurde das Fach Biochemie zu Biochemie/Molekularbiologie erweitert und somit auch die Molekularbiologie als Prüfungsgegenstand in die neue Approbationsordnung aufgenommen. Laut Absatz 2 wird die mündlich-praktische Prüfung nicht mehr wie bisher in zwei zufällig ausgewählten Fächern aus Anatomie, Physiologie, Biochemie und Psychologie / Medizinische Soziologie durchgeführt, sondern jeder Prüfling wird mündlich-praktisch in den drei Fächern Anatomie, Biochemie/Molekularbiologie und Physiologie geprüft. Die mündliche Prüfung muss gemäß Absatz 3 mit praktischen und klinischen Inhalten verknüpft werden.

§ 24 – Mündlich-praktischer Teil der Prüfung

In Absatz 1 wird nun nicht mehr die Gesamtprüfungsdauer angegeben, sondern die Prüfungsdauer pro Prüfling, die mindestens 45min, höchstens 60min beträgt.

§ 25 – Bewertung der Prüfungsleistungen

Hier wurde geändert, dass die Noten für den schriftlichen und den mündlich-praktischen Teil addiert werden und die Summe durch zwei geteilt wird. Folglich gehen beide Teile zu gleichen Teilen in die Endnote ein.

§ 26 – Zeugnis

Neben der Note des schriftlichen Teils, des mündlich-praktischen Teils und der Gesamtnote befindet sich jetzt auch die Note des Wahlfaches auf dem Zeugnis, die jedoch keine Auswirkung auf die Endnote hat.

Der Zweite Abschnitt der Ärztlichen Prüfung (§§ 27-33)

§ 27 – Zulassung zum Zweiten Abschnitt der Ärztlichen Prüfung

Neuerungen fanden hier unter anderem im Absatz 3 statt. Dort wird reglementiert, dass von den Leistungsnachweisen der 22 Hauptfächer mindestens drei fächerübergreifend gestaltet werden müssen, wobei an einem mindestens drei verschiedene Fächer beteiligt sein sollen. Eine Erkennung der beteiligten Fächer ist auf dem Schein erforderlich. Anhand von Absatz 5 ist die Benotung der Leistungsnachweise und auch der Nachweise über die Beteiligung an den Blockpraktika vorgeschrieben. Diese Noten erscheinen gesondert auf dem Zeugnis nach dem Muster der Anlage 12 zu dieser Verordnung, jedoch ohne Einfluss auf die Endnote zu haben.

§ 29 – Schriftlicher Teil der Prüfung

In den 3 Absätzen wird festgelegt, dass es sich um ein fallbezogenes, fächerübergreifendes und problemorientiertes Examen im Antwort-Wahl-Verfahren (multiple choice) handelt, das aus 320 Fragen besteht, die an drei aufeinander folgenden Prüfungstagen mit je fünf Stunden Prüfungszeit zu beantworten sind.

§ 30 – Mündlich-praktischer Teil der Prüfung

Gemäß Absatz 1 wird die mündlich-praktische Prüfung an zwei Tagen durchgeführt, wobei die praktischen Aufgaben am ersten Tag abgeleistet werden. Die Dauer der Prüfung wird nicht mehr mit der Gesamtprüfungsdauer angegeben, sondern mit

mindestens 45 min und höchstens 60 min pro Prüfling festgesetzt, wobei maximal vier Prüflinge eine Prüfungsgruppe bilden.

§ 31 – Bewertung der Prüfungsleistungen

Die Note für den Zweiten Abschnitt der Ärztlichen Prüfung wird gemäß § 25 ermittelt.

§ 33 – Gesamtnote und Zeugnis für die Ärztliche Prüfung

Entsprechend der Neuordnung der staatlichen Prüfungen wird die Gesamtnote laut Absatz 1 berechnet: Der Zahlenwert für den Ersten Abschnitt der Ärztlichen Prüfung und der mit zwei vervielfachte Zahlenwert für den Zweiten Abschnitt der Ärztlichen Prüfung werden addiert und die Summe durch drei geteilt. Somit ergibt sich eine Gesamtnote, die bis auf die zweite Stelle hinter dem Komma angegeben wird.

Bereits kurz nach der Zustimmung des Bundesrats zur 9. Novellierung der ÄAppO hat am 29.05.2002 der 105. Deutschen Ärztetag in Rostock die Bundesregierung dringend aufgefordert, die sich an das Studium anschließende Phase „Arzt im Praktikum“ (AiP) unverzüglich durch eine Änderung der Bundesärzteordnung (BÄO) abzuschaffen (68). Diese Forderung zeigte Wirkung und am 30. September 2004 wurde der 1988 eingeführte 18monatige AiP abgeschafft (9). Ebenfalls nach der 9. Änderung der Approbationsordnung wurde noch eine zweite wichtige Veränderung vorgenommen: So ist nun eine Vorverlegung des Praktischen Jahres um zwei Monate möglich, damit nach dem Praktischen Jahr wenigstens zwei Monate zur Vorbereitung auf den Zweiten Abschnitt der Ärztlichen Prüfung zur Verfügung stehen (51).

Da diese neuen Änderungen nicht alle ad hoc durchgeführt werden können und nicht für alle Studenten jeden beliebigen Semesters die gleichen Auswirkungen zeigen, gibt es sog. Übergangsbestimmungen, die sich hauptsächlich auf die Absolvierung der Prüfungen beziehen. Dieser Rechtsrahmen wird in § 43 der neuen Approbationsordnung auf zwei Druckseiten in sieben Absätzen geregelt. Studenten, die ihr Studium im Wintersemester 2003/2004 oder später begonnen haben, absolvieren das ganze Studium nach neuem Recht. Für Studenten, die ihr Studium vorher begonnen haben, sind vier Gruppen zu unterscheiden (70):

- Wer die ärztliche Vorprüfung am 1.Oktober 2003 noch nicht abgelegt hat, hat diese spätestens bis zum 30.April 2006 nach altem Recht abzulegen. Nach dem Weiterstudium gilt dann für das zweite ärztliche Examen das neue Recht.
- Wer am 1.Oktober 2003 die ärztliche Vorprüfung schon bestanden hat, hat bis spätestens 1.Oktober 2005 das erste ärztliche Examen nach altem Recht zu absolvieren. Nach Weiterstudium gilt für ihn hinsichtlich des zweiten Examens das neue Recht.
- Wer am 1.Oktober 2003 das erste ärztliche Examen bereits bestanden hat, hat bis 1.Oktober 2005 das zweite und dritte ärztliche Examen zu absolvieren. Nach diesem Stichtag ist das zweite Examen nach neuem Recht zu abzulegen.
- Wer bereits das zweite Examen am 1.Oktober 2003 nach altem Recht bestanden hat, muss auch das dritte Examen nach altem Recht ablegen.

Eine Problematik ergibt sich durch die Anerkennung von Leistungsnachweisen, die nach altem Recht erbracht wurden und den Anforderungen des neuen Rechts nicht oder nicht vollständig (z.B. mangels Benotung) entsprechen. Hier gilt nach § 43 Abs. 2 Satz 8 und 9 der Approbationsordnung für Ärzte, dass das Landesprüfungsamt solche Leistungsnachweise, soweit sie durch den Wechsel des Rechts bedingt sind, anerkennen kann.

Die wesentlichen Veränderungen durch die 9. Novellierung der Approbationsordnung und den Nachbesserungen seien noch einmal zusammenfassend dargestellt:

- Drei Monate Krankenpflegepraktikum (bisher 2 Monate)
- Fächerübergreifender und problemorientierter Unterricht
- Belegung je eines Wahlpflichtfaches vor dem Ersten und Zweiten Abschnitt der Ärztlichen Prüfung
- Benotung aller belegten Fächer und entsprechende Dokumentation im Abschlusszeugnis (jedoch ohne Einflussnahme auf die Note)
- Ärztliche Prüfung in zwei Examina: Erster Abschnitt nach vier Semestern, Zweiter Abschnitt nach dem Praktischen Jahr
- Bei Nichtbestehen eines Prüfungsteils Wiederholung nur des nichtbestandenen (schriftlichen oder mündlich-praktischen) Teils

- Vorverlegung des Praktischen Jahres um zwei Monate
- Abschaffung des Arzt im Praktikum

Prof. Dr. med. Gebhard von Jagow, Präsident des Medizinischen Fakultätentages (MFT) gab bei einer Sitzung in Frankfurt am Main folgendes Statement ab: „Bei der vollständigen Umsetzung der Reform handelt es sich um eine Agenda 2009/2010“ (83).

1.1.5) Allgemeines zum Prüfungswesen

1.1.5.1) Funktionen von Prüfungen

Die große Bedeutung von Hochschulprüfungen für Studierende, Hochschulen und Gesellschaft wird bei Trost (84) thematisiert. Ebenso fasst Trost die Funktionen von Hochschulexamina in drei Punkten zusammen (84):

Erstens ist es Aufgabe der Examina, die Studierenden über Ihren Kenntnis-, Leistungs- und Fähigkeitsstand zu informieren und eventuelle Wissenslücken aufzuzeigen. Diese Funktion der Rückmeldung wird besonders von den studienbegleitenden Prüfungen erfüllt. Der zweite Punkt bei Trost betrifft die Rückmeldung über die Effektivität der Ausbildung, die die Fakultäten anhand der Examina erlangen. Drittens sollen Prüfungen ein bestimmtes Basiswissen und Basiskönnen der Absolventen garantieren und damit die Gesellschaft vor fehlender Kompetenz der akademisch ausgebildeten Personen schützen.

Van den Bussche (85) differenziert die bei Trost (84) schon erkennbare Aufteilung der Funktionen von Examina in didaktische und gesellschaftliche weiter:

1) Gesellschaftliche Funktionen von Prüfungen

Mittels dem Ausdruck „Fächerlegitimation und Disziplinförderung“ weist Van den Bussche auf den engen Zusammenhang zwischen Prüfungsordnung und Status der einzelnen Disziplinen innerhalb der Fakultäten hin, kurz: „Ein Fach das nicht geprüft wird, ist kein Fach“ (85). Eine weitere Funktion von Prüfungen ist die Selektion und die Statuszuweisung, womit eine Regulierung des ärztlichen Arbeitsmarktes realisiert wird.

Ebenso schaffen Prüfungen Voraussetzungen für einen Kompetenznachweis, wodurch den Absolventen ein Mindestmaß an Kenntnissen, Fertigkeiten und Fähigkeiten bescheinigt wird.

2) Didaktische Funktionen von Prüfungen

Durch Prüfungen ist eine Steuerung der Lehre und des Lernens möglich, indem Dozenten ihren Lernplan und Studierende ihr Lernverhalten und ihre Lernstrategie entsprechend ausrichten. Prüfungen haben sowohl für die Lehrenden als auch für die Lernenden eine Rückkopplungsfunktion, wodurch die Lehre und das Lernen verbessert und präzisiert werden kann.

Es lässt sich somit feststellen, dass die Bedeutung von Prüfungen und Examina nicht nur im individuellen und universitären Bereich anzusiedeln ist, sondern weit über diese Bereiche hinausreicht und auch für die Gesellschaft von direkter und praktischer Relevanz ist. Die Forderung, dass Prüfungen bestimmten Anforderungen hinsichtlich ihrer Qualität als Messinstrumente genügen müssen, ist folglich berechtigt.

1.1.5.2) Testgütekriterien von Prüfungen

Die in Kapitel 1.1.5.1 deklarierten gesellschaftlichen und didaktischen Funktionen müssen von Prüfungen erfüllt werden. Da Prüfungen eine weitreichende Bedeutung besitzen, erscheint es ratsam, die Qualität von Prüfungsverfahren zu evaluieren, um diese zu sichern oder gegebenenfalls verbessern zu können. Im Folgenden sind die wichtigsten Anforderungen dargestellt, denen Prüfungen als Messinstrumente genügen müssen.

Prüfungen als Testverfahren unterliegen sogenannten Testgütekriterien. Diese Testgütekriterien wurden in verschiedenen Arbeiten für Prüfungen in der medizinischen Ausbildung festgehalten (4, 39, 42, 67, 84, 85).

Zu den zentralen Testgütekriterien zählen Objektivität, Zuverlässigkeit bzw. Reliabilität und Validität. Zu den weiteren Kriterien gehören etwa jene der Ökonomie, der Praktikabilität und der Akzeptanz. Im Einzelnen haben diese Gütekriterien folgende Bedeutung:

1) Objektivität

Eine Prüfung weist dann einen hohen Grad an Objektivität auf, wenn zwei voneinander unabhängige Prüfer zu demselben Ergebnis kommen. Der subjektive Einfluss des Prüfenden, bzw. Auswertenden soll dabei in der Durchführung und Auswertung der Prüfung sowie in der Interpretation der Prüfungsergebnisse möglichst klein gehalten werden (84, 85). Das Prüfungsergebnis kommt aufgrund möglichst genau festgelegter Prüfungs- und Bewertungskriterien zustande, die für alle Prüflinge in gleicher Weise gelten. Somit entspricht die Forderung nach Objektivität dem Gerechtigkeits- und Gleichheitsprinzip.

2) Zuverlässigkeit bzw. Reliabilität

Die Reliabilität ist ein Maß für die Messgenauigkeit einer Prüfung. Die Reliabilität ist dann besonders hoch, wenn bei wiederholter Prüfung dasselbe Ergebnis zustande kommt. Schumacher (75) definiert Reliabilität als das Vermögen eines Messinstruments, Messungen über die Zeit reproduzierbar zu erhalten. Die Reliabilität einer Prüfungsmethode wird bestimmt, indem man in einer repräsentativen Gruppe, für die der Test konstruiert wurde, mindestens zweimal dieselbe Prüfung durchführt und dann den Korrelationskoeffizienten aus den beiden Prüfungsergebnissen ermittelt. Eine weitere Möglichkeit der Erfassung der Reliabilität besteht darin, nur einen Test durchzuführen, diesen aber in zwei Hälften aufzuteilen und den Korrelationskoeffizienten aus den Ergebnissen der Teil-Teste zu bestimmen.

Die Reliabilität bzw. Zuverlässigkeit lässt sich entsprechend dieser verschiedenen Möglichkeiten der Bestimmung auch verschieden definieren (84):

- a) als Stabilität der Ergebnisse, d.h. als Übereinstimmung der Ergebnisse desselben Kandidaten bei wiederholter Bearbeitung derselben Prüfung
- b) als Paralleltest-Reliabilität, d.h. als Übereinstimmung der Ergebnisse desselben Kandidaten in verschiedenen Prüfungsversionen
- c) als innere Konsistenz einer Gesamtprüfung oder Teilprüfung.

3) Validität

Unter Validität versteht man die Gültigkeit und hier insbesondere die inhaltliche Gültigkeit oder Kontentvalidität eines Prüfungsverfahrens. Eine Prüfung wird dann als inhaltlich valide angesehen, wenn der Prüfungsstoff diejenigen Kenntnisse, Fähigkeiten oder Fertigkeiten stichprobenartig erfasst, die in der Ausbildung vermittelt werden sollen. Bei Klauer findet sich in diesem Zusammenhang auch der Begriff der Lehrzielgültigkeit (39, S. 86; 40, S. 17ff). Dabei besteht jedoch die Schwierigkeit, zu einem einheitlichen Konsens darüber zu kommen, welche Kenntnisse, Fähigkeiten und Fertigkeiten in der medizinischen Ausbildung tatsächlich vermittelt werden sollen, d.h., welche Kenntnisse, Fähigkeiten und Fertigkeiten für die spätere ärztliche Tätigkeit überhaupt relevant sind.

Einer weiteren Definition nach sind Prüfungen dann valide, wenn sie tatsächlich das messen, was sie messen sollen (85). Dabei tritt jedoch das Problem auf, ein Testverfahren zu entwickeln, das nicht irgendeine Fähigkeit oder irgendein Wissen misst, sondern gezielt bestimmte Fertigkeiten erfasst.

Eine Unterscheidung zwischen drei Aspekten von Validität findet sich in der angloamerikanischen Literatur (72, 75, 84, 85):

a) content validity

Der Begriff der content validity entspricht der Kontentvalidität und bezeichnet das Maß, in dem eine Prüfung misst, was zuvor in der Ausbildung als Lehrziel erfasst und vermittelt wurde (Lehrzielgültigkeit, s.o.). Voraussetzung für die Konstruktion einer Prüfungsmethode ist hierbei eine detaillierte, logische und klare Beschreibung des angestrebten Lernbereichs und damit des Prüfungsgebietes, den der Test erfassen soll. Da ein einzelner Test niemals das Gesamtwissen abfragen kann, muss weiterhin gewährleistet sein, dass der Test eine repräsentative Stichprobe des zu erfassenden Lernstoffs darstellt.

b) criterion validity

Die criterion validity bezeichnet die Gültigkeit, die einer Prüfungsmethode innerhalb des konstruierten Prüfungssystems zukommt. Kriterium ist die Übereinstimmung der Ergebnisse einer Einzelprüfung mit den Ergebnissen anderer Prüfungsmethoden bzw.

der Gesamtprüfung. Voraussetzung für diese „kriterienbezogene Validierung“ ist eine genaue Definition der einzelnen Bereiche ärztlicher Kompetenz und Qualifikation, um entsprechend dieser Teilbereiche Prüfungsmethoden zur Erfassung derselben zu entwickeln.

c) construct validity

Unter construct validity versteht man das Maß, in dem ein Test misst, was er seiner Konstruktion zufolge zu messen vorgibt. Bei jeder Prüfungsmethode wird eine bestimmte Art, bzw. eine bestimmte Art der Anwendung von Wissen, Fähigkeiten und Fertigkeiten getestet. Ausgangspunkt dafür ist, dass zur Lösung einer Aufgabe oder eines Problems verschiedene kognitive Prozesse und auch andere Fähigkeiten notwendig sind, die auf den ersten Blick nicht erkennbar sind und der Bewältigung der Prüfung zugrunde liegen. Als Beispiel lässt sich hier die Konstruktion „Lösung klinischer Probleme“ (clinical problem solving) anführen. Dabei ist davon auszugehen, dass diese Fähigkeit im Laufe der medizinischen und vor allem der klinischen Ausbildung mit wachsender klinischer Erfahrung zunehmend entwickelt wird. Diese Fähigkeit des clinical problem solving liegt somit dem Lösungsprozess zugrunde, wobei sie nicht direkt beobachtbar ist. Eine Erfassung ist nur auf indirektem Weg möglich, indem man diese Fähigkeit gedanklich konstruiert.

4) Praktikabilität, Aufwand und Kosten

Hierbei kommt die aus dem Wirtschaftsbereich bekannte Kosten-Nutzen-Analyse zum Einsatz, d.h. es wird das Verhältnis zwischen Aufwand und Nutzen eines Prüfungsverfahrens bestimmt (67, 84, 85). Dabei weisen zum Beispiel schriftliche Prüfungen im Vergleich mit mündlichen Prüfungen den Vorteil auf, dass viele Kandidaten zugleich geprüft werden können, so dass eine Einsparung an Personal und Zeit vorliegt, was sich letztendlich kostengünstig auswirkt. Das Kriterium der Praktikabilität weist auf die Durchführbarkeit hin, da nicht jede Prüfungsmethode sich von einer Prüfungssituation auf eine andere übertragen lässt. Somit müssen bei der Konstruktion einer Prüfung die Situation des Umfeldes sowie auch die Gegebenheiten vor Ort berücksichtigt werden.

5) Akzeptanz

Unter Akzeptanz versteht man den Grad an Zustimmung, den eine Prüfungsmethode bei Studierenden, Prüfern und der Öffentlichkeit erfährt und der für die Brauchbarkeit einer Prüfung eine grundlegende Rolle spielt (84). Das Kriterium der Akzeptanz ist hauptsächlich bei der Neueinführung von Prüfungsmethoden von Bedeutung, da die Akzeptanz mit zunehmender Gewöhnung erfahrungsgemäß wächst.

Eine Rangfolge der Testgütekriterien von Prüfungen wurde von der American Psychological Association (4) erstellt. Dabei ist die Erfüllung des vorausgehenden Kriteriums notwendige Voraussetzung für das nächstfolgende Kriterium:

Objektivität – Reliabilität – Validität – Praktikabilität – Nutzen – Übertragbarkeit.

Zum Beispiel ist die Reliabilität eine notwendige, aber keine hinreichende Bedingung für die Gültigkeit einer Prüfung. Erst wenn das Kriterium Validität auch erfüllt ist, kann man von einer gültigen bzw. sinnvollen Prüfung ausgehen, da in diesem Fall auch nicht relevanter Prüfungsstoff zuverlässig geprüft werden kann (85).

Der Rangfolge der APA (4) nach stellt somit die Übertragbarkeit das höchste Testgütekriterium dar, da alle übrigen Kriterien erfüllt sein müssen, um eine Prüfungsmethode auf eine andere Testsituation übertragen zu können.

Letztendlich sollte eine Prüfungsmethode die genannten Testgütekriterien in einem ausreichenden Maß erfüllen. Zur Bestimmung des Maßes, in welchem die verschiedenen Prüfungsmethoden diesen Kriterien gerecht werden, dient die Meta-Evaluation von Prüfungsverfahren als die Evaluation von Evaluationsmethoden. Die verschiedenen Prüfungsmethoden werden in Kapitel 1.1.6 dargestellt und hinsichtlich ihrer Testgüte miteinander verglichen.

1.1.5.3) Klassifikation kognitiver Prozesse

Wie in Kapitel 1.1.5.2 unter dem Begriff „construct validity“ bereits aufgeführt, werden anhand der verschiedenen Prüfungsmethoden jeweils eine bestimmte Art, bzw. eine bestimmte Art der Anwendung von Wissen, Fähigkeiten und Fertigkeiten geprüft. Diese Überlegungen zur Gültigkeit von Testverfahren führten zur Analyse der Fähigkeiten

und kognitiven Prozesse, die der Aufgabenlösung zugrunde liegen und die ebenso in alltäglichen Entscheidungsprozessen in der ärztlichen Tätigkeit von großer Bedeutung sind. Auf Grund dieser Überlegungen wurden Untersuchungen und Arbeiten zur Struktur medizinischen Wissens, zu seiner Speicherung und zu seiner Abrufung durchgeführt (14, 16, 56, 62, 75).

Eine Klassifikation kognitiver Prozesse ist bei Bloom (13) und bei McGuire (50) zu finden.

Bloom (13) definiert sechs kognitive Prozess-Niveaus, die entsprechend ihrer Komplexität hierarchisch geordnet sind:

Wissen – Verstehen – Anwendung von Wissen – Analyse – Synthese – Bewertung.

Durch diese hierarchische Gliederung wird impliziert, dass ein kognitiver Prozess auf einem bestimmten Niveau voraussetzt, dass das vorausgehende Niveau beherrscht wird. Soll z.B. eine Bewertung abgegeben werden, muss das Gewusste zunächst verstanden sein, eine korrekte Anwendung durchgeführt worden sein, sorgfältig analysiert und darüber eine Synthese erbracht worden sein.

Als Grundlage für die „National Question Bank of the Association of University Clinical Professors of Australia“ diente die Klassifikation nach McGuire (50), die auf der Einteilung nach Bloom basiert und eine Vereinfachung von dieser darstellt. McGuire ordnet kognitive Prozesse folgenden drei Kategorien zu:

a) Niveau 1

Diese Kategorie impliziert den Abruf und das Wiedererkennen von isolierter Information. Um eine Aufgabe dieser Kategorie lösen zu können, muss der Kandidat in der Lage zu sein, Faktenwissen zu speichern und bei Bedarf abzurufen.

b) Niveau 2

Fragen dieser Kategorie testen Verstehen und Interpretation von Information. Das Lösen einer Frage, die diesem Niveau zugeordnet wird, erwartet das Verstehen von Daten, Auswerten von Daten und Übertragen von Daten in eine andere Form.

c) Niveau 3

In dieser Kategorie wird die Fähigkeit des Kandidaten getestet, Wissen zur Lösung eines spezifischen Problems anzuwenden. Beim Problemlösenden Denken gibt es eine bestimmte Struktur: Analyse des Problems, Abruf von relevantem Faktenwissen, Anwendung des Wissens auf die spezifische Situation.

Diese Klassifikation nach McGuire (50) wird im folgenden Kapitel angewandt, um verschiedene Prüfungsmethoden in ihrer strukturellen Konzeption miteinander zu vergleichen.

1.1.6) Verschiedene Prüfungsmethoden

Prüfungsmethoden lassen sich grob einteilen in schriftliche, mündliche und praktische Prüfungen. Aktuell kommen im Rahmen der medizinischen Ausbildung in Deutschland vor allem Multiple-Choice-Prüfungen als Vertreter der schriftlichen Prüfungen, mündliche und praktische Examensprüfungen zur Anwendung. In den letzten Jahren wurden vor allem im anglo-amerikanischen Bereich zahlreiche neue Prüfungsmodelle entwickelt (Objective Structured Clinical Examination bzw. OSCE, Modified Essay Question, Triple Jump,...), die vermehrt auch Anklang im deutschsprachigen Raum finden.

In den Übersichtsarbeiten „Gutachten über das Prüfungswesen in einer neuen ÄAppO“ (67), „Bericht über die Beurteilungsmethoden in der medizinischen Grundausbildung“ (42) und „Positionspapier zur Reform des Prüfungswesens im Medizinstudium“ (19) findet sich eine vergleichende Gegenüberstellung von Prüfungsmethoden.

In diesem Kapitel werden die am häufigsten verwendeten Prüfungstechniken klassifiziert und beschrieben. Des Weiteren erfolgt ein Vergleich hinsichtlich ihrer wichtigsten Vor- und Nachteile sowie der in Kapitel 1.1.5.2 dargestellten Testgütekriterien.

1.1.6.1) Schriftliche Prüfungen

Aus dieser Kategorie sollen näher dargestellt werden:

Multiple-Choice-Prüfungen, Short-Answer-Question, Essay und Modified-Essay-Question.

Weitere schriftliche Prüfungstechniken, die weniger weit verbreitet sind und auf die hier nicht näher eingegangen werden soll, sind der Progress Test (PT), der SIMP-Test (simulation of initial medical problem solving) und der PMP-Fragetyp (patient management problem solving).

a) Multiple-Choice-Prüfungen

Multiple-Choice Fragen (MC-Fragen) sind Auswahlfragen, von denen mehrere Varianten existieren. Haladyna (28) beschreibt in seinem Buch „Developing und Validating Multiple-Choice Test Items“ zahlreiche Varianten von MC-Fragen. Im Folgenden werden die Typen von MC-Fragen besprochen, die in der medizinischen Ausbildung in Deutschland ihre Anwendung finden. Nach Haladyna besteht die konventionelle MC-Frage aus drei Komponenten: „stem“, „correct choice“ und „several wrong answers, called foils, misleads or distractors“ (28). Dabei wird „stem“ als Reiz für die Antwort definiert („stimulus for the response“), „correct choice“ als die zweifellos eine und einzige richtige Antwort („undeniably one and only right answer“) und „distractor“ als definitive falsche Antwort („unquestionably wrong answer“).

Jede MC-Frage beinhaltet somit vorgefertigte Antwortmöglichkeiten, aus denen der Prüfling die korrekte auszuwählen hat. Je nach Anzahl der vorgegebenen Möglichkeiten und der Antwort als Einfach-, Mehrfach- oder Kombinationsauswahl erfolgt die Kategorisierung der Fragen. Die in Deutschland am meisten verwendeten Formen sind Typ A, Typ B und Typ K.

Typ A besteht aus einem Fragenstamm in Form einer Frage oder einer unvollständigen Aussage und vier oder fünf möglichen Antworten, von denen die richtige zu wählen ist. Bei Typ B handelt es sich um eine Zuordnungsaufgabe: hier sollen einer Reihe von Fragen oder Aussagen die entsprechenden Ergänzungen zugeordnet werden. Typ K besteht wie Typ A aus einem Fragenstamm und Ergänzungen, von denen hier aber mehr als eine richtig sein kann. Aus den vorgegebenen Kombinationen der Richtigantworten soll die zutreffende gewählt werden.

Auf Grund der eindeutigen Auswertung mit dem Computer sind Auswertungsfehler und Prüferfehler auf ein Minimum reduziert (42), wodurch sich MC-Fragen durch eine hohe Objektivität auszeichnen (67). Begründet in der großen Anzahl unabhängiger Fragen weisen MC-Fragen auch eine hohe Reliabilität auf, was einen weiteren wesentlichen Vorteil der MC-Fragen darstellt (46, 67). Ein weiterer Vorteil von MC-Fragen ist die Tatsache, dass in relativer kurzer Zeit ein sehr breites Spektrum an Wissen geprüft werden kann (42).

Bezüglich der Validität von MC-Fragen weist die Literatur keinen Konsens auf, so dass die Validität bis heute noch Gegenstand vieler Untersuchungen ist. Hinsichtlich der Überprüfung reinen Faktenwissens stellen sich MC-Fragen als valide dar, wenn die Voraussetzung erfüllt ist, dass die geprüften Fakten auch tatsächlich ausbildungsrelevant sind (67). „Begründete Zweifel an der Gültigkeit (Validität) des gängigen Multiple-Choice-Verfahrens“ werden im Positionspapier zur Reform des Prüfungswesens im Medizinstudium geäußert (19, S.3), da es fraglich erscheint, ob mit diesem Verfahren ärztliche Kompetenz erfasst werden kann (38, 55, 61, 63).

b) Short-Answer-Question

Short-Answer-Questions sind sehr spezifische Fragen, die der Prüfling in Stichworten innerhalb eines bestimmten Umfangs beantworten soll. Im Bereich der Short-Answer-Questions finden sich Fragen von sehr unterschiedlichem Ausmaß und sehr unterschiedlicher Spezifität, was einen fließenden Übergang von MC-Fragen hin zu den Short Essays impliziert.

Anhand eines vorgefertigten Lösungsschlüssels erfolgt die Auswertung von Short-Answer-Questions. Da die Auswertung per Hand durchgeführt wird, sind Auswertfehler bzw. Differenzen zwischen einzelnen Auswertern nicht auszuschließen. Bei sorgfältiger und genauer Konstruktion der Short-Answer-Questions erreicht die Objektivität einen hohen Standard, liegt aber unter derjenigen von MC-Fragen. Da die Korrektur wie die Auswertung per Hand durchgeführt wird, entsteht ein höherer Arbeitsaufwand.

Hinsichtlich der Reliabilität von Short-Answer-Questions gibt es in der Literatur unterschiedliche Angaben (25, 55), die aber nicht wesentlich unter der Reliabilität von MC-Fragen liegen. Jedoch gibt es in der Literatur kein einheitliches Statement zur „construct validity“, ob nun Offene Fragen andere Fähigkeiten messen als MC-Fragen (55, 63).

Eine weitere Untersuchung zum Thema Multiple Choice Questions und Short Essay Questions wurde von Rothman und Kerenyi durchgeführt (72). Sie kamen zu dem Ergebnis, dass MC-Fragen das spezifische Faktenwissen messen, während Short Essay Questions die Fähigkeit erfassen, eine klare Antwort auf Probleme zu geben, die in kurzen Fallstudien dargestellt sind.

c) Essay

Hierbei handelt es sich um relativ weit gefasste Fragen, die in schriftlich in Form eines kurzen Aufsatzes zu beantworten sind. Anhand von Essays wird das Abrufvermögen von gespeichertem Wissen, das Herstellen von Zusammenhängen, das logische Deduzieren und darüber hinaus die sprachliche Ausdrucksfähigkeit des Kandidaten getestet.

Neben dem Verständnis für grundsätzliche Mechanismen und Zusammenhänge sowie die Anwendung derselben kann bei einer entsprechenden Konzipierung der Fragen auch die Einstellung der Prüflinge getestet werden (67).

Hinsichtlich der Objektivität erreicht der Essay nicht diejenige von MC-Fragen, da die Bewertung trotz gut konzipiertem Auswerteschema von Prüfer zu Prüfer unterschiedlich ausfallen kann und in der Auswertung mehr Fehler auftreten können (67). Da in Essays nur ein relativ kleiner Wissensausschnitt pro Zeiteinheit erfasst wird, finden sich für die Reliabilität meist geringere Werte als für MC-Fragen. Negativ auf die Reliabilität wirkt sich auch das Vorkommen einer unpräzise gestellten Frage aus, da der Prüfling die Intention der Fragestellung auch falsch deuten kann.

Die „construct validity“ zeigt das Problem der Essays auf, da in die Bewertung von Essays auch sprachliche Ausdrucksfähigkeit, Grammatik und Stil sowie das Vermögen zu einer gegliederten Darstellung einfließen. Neben der Problematik, wie diese Fähigkeiten gewertet werden sollen, stellt sich die Frage, welche Bedeutung sie für die zukünftige Tätigkeit als Arzt haben (42, 67). Ein hoher Arbeits- und Zeitaufwand für die Auswertung von Essays muss ebenfalls berücksichtigt werden.

d) Modified-Essay-Question (MEQ)

Diese Prüfungsmethode basiert auf den Essays und wurde entwickelt, um die Objektivität und Praktikabilität der Essays zu verbessern. Ursprünglich wurden die MEQs 1971 durch das „Board of Censors“ des „Royal College of General Practitioners“ entworfen und 1975 von Hodgkin und Knox (35) zur Förderung des Problembezogenen Lernens weiterentwickelt. Die Modified-Essay-Questions sind strukturierte Tests, bei denen die Darstellung eines klinischen Falls den Fragen vorausgeht. Die gestellten Fragen können sowohl Offene Fragen als auch MC-Fragen sein. Vorteile gegenüber den Essays sind die präzisere Fragestellung, der Bezug auf einen vorgegebenen Fall und ein weniger problematisches Korrekturschema. Somit weisen die MEQs eine höhere Objektivität als die Essays auf, aber immer noch niedriger als diejenige von MC-Fragen (67).

Hinsichtlich der Validität erreichen die MEQs gute Werte (19, 67), die „construct validity“ wurde in mehreren Arbeiten nachgewiesen (63). Irwin und Bamber (37) weisen in ihrer Arbeit nach, dass MEQs alle drei von McGuire definierten Prozess-Niveaus erfassen können. Somit testen MEQs sowohl Wiedererkennen und Abruf von Faktenwissen, Verstehen und Interpretation, als auch die Anwendung von Wissen zur Lösung eines spezifischen Problems.

Das Erfassen eines größeren Wissensspektrums und die dem klinischen Entscheidungsprozess sehr ähnliche Struktur der MEQs zählen zu deren Vorteilen. Laut Feletti und Engel (24) bedeutet „die Konstruktion und Auswertung von MEQs eine bemerkenswerte intellektuelle Herausforderung“, abgesehen von der zeitaufwendigen Komponente.

1.1.6.2) Mündliche Prüfungen

Bis zur Einführung der Approbationsordnung 1970 wurden alle Staatsprüfungen in der medizinischen Ausbildung anhand von mündlichen Prüfungen durchgeführt, so dass ihnen in dieser Zeit eine erheblich größere Bedeutung zukam. Heutzutage bilden die mündlichen Prüfungen das zweite Element der in Deutschland in Staatsprüfungen verwendeten Prüfungsarten.

In der mündlichen Prüfung werden Wissen, Ausdrucksfähigkeit und Reaktionsschnelligkeit des Prüflings bemessen. Dabei muss der Prüfling umgehend

und ohne Hilfsmittel auf die Fragen des Prüfers antworten. Je nach Konstruktion und Inhalt der Frage können dabei unterschiedlich komplexe Fähigkeiten getestet werden, vom bloßen Prüfen des Faktenwissens bis hin zum Problemlösenden Denken.

Im Vergleich zur schriftlichen Prüfung hat die mündliche Prüfung den Vorteil, dass sie von einer Interaktion zwischen Prüfer und Prüfling bestimmt wird. Anstelle der bloßen Konfrontation mit dem Resultat bei der schriftlichen Prüfung hat der Prüfer bei der mündlichen Prüfung einen direkten Einblick in den Denkprozess des Prüflings und kann in diesen korrigierend oder steuernd eingreifen. Der Vorteil auf Seiten des Prüflings liegt in der unmittelbaren Rückmeldung begründet. Aufgrund der Interaktion zwischen Prüfer und Prüfling lässt sich nicht festlegen, welche Richtung die Prüfung einschlägt. Durch das Ausrichten der weiteren Fragen inhaltlich und hinsichtlich des Schwierigkeitsgrades an den gegebenen Antworten lässt sich der Wissensstand sehr spezifisch und individuell erfassen. Ebenfalls zeichnen sich mündliche Prüfungen durch eine große Vielfalt der Fragestruktur aus, so dass durch gezielte Stichfragen Wissensbereiche abgefragt und aufeinander aufbauende Sachverhalte zusammenhängend geprüft werden können.

Ein Nachteil von mündlichen Prüfungen findet sich in der Objektivität, da sie unter teilweise großen Beurteilerunterschieden leiden und subjektive Faktoren wie Reaktion auf Kleidung, Geschlecht und Attraktivität Einfluss auf das Zustandekommen des Ergebnisses haben (42). Ebenso darf der Stressfaktor für den Prüfling nicht vernachlässigt werden, der sich deutlich im Ergebnis niederschlagen kann (42). Ein weiterer Nachteil ist, dass mündliche Prüfungen nur einen relativ kleinen Ausschnitt des Wissens erfassen, sofern der zeitliche Rahmen eingehalten wird. Ökonomisch gesehen ist der Zeitaufwand sehr hoch, was sich folglich auch negativ auf den Kostenaufwand auswirkt (67).

Zusammenfassend ist festzuhalten, dass mündliche Prüfungen dem Prüfer einen Einblick in die Denkstrategien und kognitiven Prozesse des Prüflings geben, jedoch eine mangelnde Objektivität und Reliabilität aufweisen (42, 67). Um diesem Mangel entgegenzuwirken, wurden Modelle für strukturierte und wiederholbare Prüfungen entwickelt, von denen im Folgenden zwei dargestellt werden sollen: Die standardisierte Mündliche Prüfung und der Triple Jump.

a) Standardisierte Mündliche Prüfung

Der Unterschied zur herkömmlichen mündlichen Prüfung liegt darin, dass die Prüfung nach einer vorgegebenen Struktur abläuft. Dem Kandidaten wird ein Datenblatt mit einer klinischen Fallbeschreibung ausgehändigt, zu der er gezielte Frage stellt und Hypothesen entwirft bezüglich der Krankengeschichte und der Differentialdiagnosen. Er soll zur Diagnosesicherung sinnvoll erscheinende Untersuchungen nennen und Therapiekonzepte entwickeln. Die Leistungen des Kandidaten werden vom Prüfer zu den bestimmten Bereichen anhand einer bestimmten Skala bewertet (67). Gegenüber der herkömmlichen mündlichen Prüfung zeichnet sich die Standardisierte Mündliche Prüfung durch eine höhere Objektivität und Reliabilität aus. Nachteilig ist die aufwendige Vorbereitung dieser standardisierten Prüfungen, die kontrolliert entwickelt und evaluiert werden müssen.

Anzumerken bleibt die Tatsache, dass vor allem in den vorklinischen Fächern die Fallbeispiele keinen effektiven Lerneffekt aufweisen, da in diesem Abschnitt der medizinischen Ausbildung der Schwerpunkt auf medizinischen Grundlagen und naturwissenschaftlichen Zusammenhängen liegt.

b) Triple Jump

Diese Prüfung verläuft in drei Schritten:

1. Schritt: Problemstellung und Hypothesengenerierung
2. Schritt: Lernphase
3. Schritt: Auswertung

Im ersten Schritt wird dem Prüfling ein klinisches Problem zugeteilt, zu dem er nach einer Vorbereitungszeit den Prüfern seine Hypothesen mitteilt. Anschließend definiert er in Kenntnis seiner Wissenslücken seine Lernziele. Im zweiten Schritt werden dem Kandidaten mehrere Stunden Zeit gegeben, um seine selbst gesetzten Lernziele zu erreichen. In der dritten Phase präsentiert der Prüfling seine Ergebnisse und wird von den Prüfern hinsichtlich seines Lernerfolges und seines methodischen Vorgehens bewertet (67).

Die Reliabilität des Triple Jump ist höher als bei unstrukturierten mündlichen Prüfungen (78), hängt aber von der Strukturierung der Auswertungsmodalitäten ab. Da der Triple Jump die Fähigkeiten im klinischen Problemlösungsprozess testet, besitzt er eine hohe

„content validity“ (67). Wie alle mündlichen Prüfungen gibt der Triple Jump auch Aufschluss über die sprachlich-logischen Fähigkeiten des Kandidaten und erlaubt eine direkte Rückkopplung für Prüfer und Prüfling.

1.1.6.3) Praktische Prüfungen

Praktische Prüfungen sind speziell dafür geeignet, durch Verhaltensbeobachtung und Beobachtung von Handlungsabläufen die Fertigkeiten und praktische Kompetenz der Kandidaten zu testen. Da die praktischen Fertigkeiten für die spätere ärztliche Tätigkeit von enormer Bedeutung sind, bedarf es einer angemessenen praktischen Ausbildung und eines entsprechend ausgerichteten Prüfungssystems.

Das „Münchner Curriculare Innovationsprojekt“ (MCIP) (21) definiert den Erwerb praktischer Fertigkeiten als Verlauf auf drei Ebenen:

Ebene A beinhaltet das kognitive Lernen und Verstehen des Prozesses (Wissensrekapitulation, Selbstlernprogramme, Seminare).

Die Einübung sensomotorischer Abläufe findet auf der Fertigkeitsebene B statt (Einübung sensomotorischer Fertigkeiten in sogenannten Übungslabors).

Die Anwendung der erlernten Fertigkeit in der Praxis und deren Perfektion wird von Ebene C verkörpert (Unterricht am Krankenbett).

Somit stellt der Erwerb praktischer Fertigkeiten einen komplexen Prozess dar, dessen Resultat in einer Prüfungssituation durch den Prüfer beobachtet und bewertet werden kann.

Die Wichtigkeit der klinisch-praktischen Prüfung wird von der Kommission der Europäischen Gemeinschaften in seinem „Bericht über die Beurteilungsmethoden in der medizinischen Grundausbildung“ hervorgehoben: „Die klinische Prüfung ist von allen Beurteilungen die wichtigste, denn sie betrifft die wesentlichen Aufgaben eines Arztes.“ (42, S.11).

Nachteil der praktischen Prüfungen sind der organisatorische Aufwand und die zeitaufwendige Durchführung. Ebenso nachteilig wirkt sich auf Grund von erheblichen Unterschieden bezüglich des Prüfungsverlaufes, Schwierigkeitsgrades, Patientenguts und Beurteilung die niedrige Objektivität und Reliabilität aus, speziell bei den

herkömmlichen, unstrukturierten praktischen Prüfungen. Um dem Mangel an Objektivität entgegenzuwirken wurden Modelle für strukturierte praktische Prüfungen entworfen, von denen hier die Objective Structured Clinical Examination (OSCE) zu nennen ist.

Die OSCE wurde als Methode zur Einschätzung klinischer Fähigkeiten entwickelt (32, 33). Nach ihrer Einführung 1975 wurde diese Prüfungsmethode in den verschiedensten Fachrichtungen und Ländern verwendet und fand international Anerkennung (12).

Bei der OSCE stehen die Beobachtung der Prüflinge bei der praktischen Arbeit, die Bewertung anhand von Checklisten und die Simulation des klinischen Alltags im Vordergrund. Die Gesamtprüfung besteht aus ca. 20 Einzelprüfungen mit einer Dauer von jeweils 10-20 Minuten. Die Prüfung findet am Krankenbett statt. Der dabei anwesende Prüfer schätzt die Fertigkeiten des Kandidaten anhand vorgegebener Checklisten ein.

Testgütekriterien und Praktikabilität des OSCE-Verfahrens wurden in verschiedenen Studien getestet (12, 32, 33, 71).

Dabei wurde die Reliabilität des OSCE für ausreichend befunden (71). Ebenso gilt das OSCE-Verfahren als sehr valide (71). Nachteil ist der hohe Kostenaufwand auf Grund der speziellen Entwicklung und der Durchführung und Auswertung (19, 67).

Insgesamt erweist sich das OSCE-Verfahren als praktikabel und übertragbar. Newble (57) beschreibt OSCE als die erfolgreichste Innovation auf dem Gebiet der medizinischen Ausbildung.

1.1.7) Schriftliche Prüfungsmethoden im Vergleich

In den folgenden Kapiteln 3 und 4 beschäftigen sich die eigenen Untersuchungen dieser Arbeit u.a. mit dem Vergleich von Multiple-Choice Fragen (MC-Fragen) und Offenen Fragen. Deswegen wird nun in diesem Kapitel näher auf diese beiden Arten schriftlicher Prüfungsmethoden eingegangen.

Diese Fragetypen werden in Hinblick auf ihre Vor- und Nachteile und auf die ihnen anhaftende spezifische Probleme analysiert.

1.1.7.1) Vorteile und Nachteile von MC-Fragen

Das in den USA entwickelte Multiple-Choice Testverfahren wird dort seit 1950 als schriftliche Prüfungsmethode in der medizinischen Ausbildung verwendet. In der Schweiz erfolgte die Einführung des MC-Verfahrens 1971 (66), in der Bundesrepublik Deutschland 1973 im Zuge der Umsetzung der Approbationsordnung. Aktuell ist das MC-Verfahren weltweit verbreitet.

Das MC-Verfahren mit seinen unterschiedlichen Fragetypen und hinsichtlich der Testgütekriterien wurde in Kapitel 1.1.6.1 bereits dargestellt. Zusammenfassend sei noch einmal gesagt, dass über die hohe Objektivität und Reliabilität dieser Methode weitgehend Einigkeit besteht, jedoch die Validität umstritten ist (s. Kapitel 1.1.6.1). Festzuhalten ist, dass das MC-Verfahren in all seinen Aspekten sorgfältig dokumentiert wurde (82). Im Folgenden werden die Vor- und Nachteile von MC-Fragen dargestellt und diskutiert.

Die wesentlichen Vorteile von MC-Prüfungen sind folgende:

a) Objektivität und Reliabilität

Die hohe Objektivität und Reliabilität von MC-Fragen kommt in der Literatur generell zu einem Konsens (19, 42, 67).

In einer Studie zu den Testgütekriterien von MC-Fragen konnten Levine et al. (60) aufzeigen, dass der Reliabilitätskoeffizient (Maß für die Reliabilität eines Testverfahrens) eine direkte Korrelation mit der Anzahl der gestellten Fragen aufweist. Bei einem Test mit einem Umfang von 150 Fragen wurde z.B. eine Reliabilität von 0,72 gefunden, wohingegen die Reliabilität bei einem Test aus 230 Fragen 0,89 betrug. Sub-Tests von Prüfungen weisen demnach eine niedrigere Reliabilität auf und sind als alleinige Messinstrumente nicht geeignet, zeichnen sich jedoch im Zusammenhang mit anderen Testverfahren durch eine ausreichende Gesamt-Reliabilität aus.

Beim Vergleich der schriftlichen Prüfungsmethoden MC-Fragen und Patient Management Problems (PMPs) in der Studie von Norcini et al. (61) zeigte das Ergebnis, dass die Reliabilität für das Gesamt-Ergebnis gleich oder nur geringfügig kleiner ist als für die MC-Komponente der Prüfung. Daraus lässt sich schließen, dass

eine Erweiterung der Prüfung durch PMPs keine Steigerung der Gesamt-Reliabilität mit sich bringt.

Hinsichtlich der Reliabilität und Validität der verschiedenen MC-Subtypen (Typ A, Typ K, Typ T/F, siehe Kapitel 1.1.6.1) kam Nnodim (60) in seiner Studie zu dem Ergebnis, dass die Gesamt-Reliabilität in den drei Gruppen zwischen 0,71 und 0,85 liegt und sich damit oberhalb des von Lipton und Huxham festgelegten Minimums von 0,70 befindet.

Dass in relativ kurzer Zeit ein umfangreicher Wissensbereich stichprobenartig erfasst werden kann, ist u.a. für die hohe Reliabilität verantwortlich. Elstein (23) beschreibt diesen Vorzug von MC-Fragen folgendermaßen: „..., whereas more MCQs can be asked per unit of time than can any other kind of question (except the true – false ones). So the MCQ format can efficiently sample many clinical situations.“ (23, S. 245). Elstein (23) weist hier auf den Punkt hin, in dem die Schwierigkeit von mündlichen oder praktischen Prüfungen liegt. Soll klinisches Urteilsvermögen oder Problemlösen getestet werden, so muss diese Prüfung extensiv und umfangreich sein, denn das ärztliche Verhalten in klinischen Situationen zeigt große Variabilität: „One or two simulated clinical situations will not do.“ (23, S. 245). Aus Zeitgründen können mündliche oder praktische Prüfungen nur wenige Stichproben des Gesamt-Prüfungsstoffs abfragen, so dass die Zufälligkeit der Auswahl des klinischen Falls in der Prüfung eine größere Rolle spielt.

b) Praktikabilität, Kostenaufwand und Ökonomie

Die Testung von vielen Prüflingen gleichzeitig in relativ kurzer Zeit bei minimaler Anwesenheit von Prüfern ist ein weiterer Grund für die weite Verbreitung von MC-Prüfungen. Ebenso kann auf Grund der eindeutigen Antwortmöglichkeiten die Auswertung per Computer stattfinden. Bei Lennox (45) und Whitby (87) wird dieser Vorteil von MC-Prüfungen als „ease of scoring“ bezeichnet. Elstein (23) geht mit seiner Aussage noch einen Schritt weiter, indem er von einem „enthusiasm for the format and the ease of scoring large numbers of examinations“ (23, S. 245) spricht.

Das „Gutachten über das Prüfungswesen“ (67), das im Auftrag der Bundesministeriums für Gesundheit (BMG) erstellt wurde, kommt zu dem Ergebnis, dass „bei einem vorbereiteten Fragenpool [...] die Kosten des MC-Verfahrens sehr viel geringer sind als bei allen anderen Prüfungsmethoden“ (67, S.15). Norcini et al. (61) beschreiben den finanziellen Aspekt von Prüfungen, an dem sehr qualifizierte

Prüfungsmethoden gescheitert sind, folgendermaßen: „Due to factors such as cost in development and implementation, most of these approaches have proven impractical large-scale use.“ (61, S. 238). Diese Aussage wird von Elstein (23) bestätigt: „they (= essay questions and oral examinations) have been replaced by MCQs largely because of problems of consistency in grading, breadth of domain sampling and cost.“ (23, S. 244).

Aufschluss über die Ökonomie des MC-Testverfahrens ergibt sich durch die Kosten-Nutzen-Analyse: geringe Kosten und ein hoher Nutzen sprechen für einen hohen ökonomischen Wert. Durch die Berechnung des Effektivitätskoeffizienten von Prüfungen, der als Reliabilität pro Prüfungszeit-Einheit definiert ist, lässt sich ebenfalls der Nutzen von Prüfungen bestimmen. Die Auswertung der Arbeit von Norcini et. al (61) zeigte Effektivitätskoeffizienten zwischen 0,63 und 0,90, wobei die Werte für MC-Fragen (0,85-0,88) wesentlich höher lagen als die für die PMPs (0,63-0,70).

Die Nachteile von MC-Prüfungen stellen sich folgendermaßen dar:

a) Validität

Die Validität von MC-Prüfungen ist umstritten.

Die content validity von MC-Fragen ist in erster Linie davon abhängig, ob die Fragen inhaltlich das Wissensspektrum erfassen, das in der Ausbildung vermittelt wurde. In diesem Punkt unterscheiden sich MC-Fragen nicht von anderen Prüfungsmethoden, denn das Maß an Kontentvalidität ist nicht primär eine Eigenschaft des Testverfahrens, sondern liegt in der adäquaten inhaltlichen Konstruktion der Fragen begründet. Somit ist die content validity für den Fall gegeben, dass die im Gegenstandskatalog formulierten Inhalte an den Fakultäten den Studierenden auch tatsächlich vermittelt werden.

Die criterion validity war Gegenstand der Studie von Levine et al. (46). Die Analyse der Auswertung ergab, dass die Ergebnisse der MC-Prüfung stärker mit den Beurteilungen kognitiver Prozesse korrelierten als mit der Beurteilung von praktischen Fertigkeiten oder affektiven Eigenschaften. In Zusammenhang mit den Ergebnissen für die construct validity (s. unten) stellten die Autoren fest, dass MC-Fragen gültige Informationen über einige Bereiche klinischer Kompetenz liefern, dass sie aber einer Ergänzung durch andere Techniken bedürfen, um alle Aspekte klinischer Kompetenz zu erfassen (46).

Die Bewertung der construct validity ergibt sich unter dem Gesichtspunkt der in Kapitel 1.1.5.3 dargestellten Gliederung der kognitiven Anforderungen, die zur Aufgabenbewältigung bewältigt werden müssen. Es herrscht großes Einvernehmen darüber, dass die Validität von MC-Prüfungen hinsichtlich der Überprüfung isolierten Faktenwissens hoch ist, was dem Komplexitäts-Level eins entspricht (67). Dass zwischen dem zugeschriebenen Komplexitäts-Level einer Aufgabe und dem tatsächlichen Schwierigkeitsgrad keine Korrelation besteht, konnte Cox (18) in seiner Studie beweisen. Die Ergebnisse veranlassen Cox, vor einer voreiligen Interpretation von MC-Fragen zu warnen: „The observation [...] warn us against interpretation of problem-solving skills from multiple-choice questions, and against assumptions about how others reach their solutions, correct or incorrect.“ (18, S. 886). Dass schnelles Abrufen von Faktenwissen in der ärztlichen Tätigkeit von großer Bedeutung ist, wird von Elstein (23) beschrieben. Ebenso betont Elstein, dass es wichtig ist zu wissen, wie man Fragen konstruiert, die komplexe kognitive Prozesse erfassen. Wie Cox (18) folgert Elstein aber auch aus seinen Studienergebnissen, dass man letztendlich nicht wissen kann, auf welchem Weg jemand eine MC-Frage gelöst hat: „Any question can be answered either way, depending on the knowledge base of the examinee.“ (23). Auf diesen Nachteil macht auch Anbar (1) aufmerksam: Jeder Prüfer, der Essays oder Offene Fragen ausgewertet habe, wisse, wie wichtig es ist, dem Argumentationsgang des Kandidaten zu folgen. Dieses sei mit MC-Fragen nicht möglich. Mit der gleichen Thematik beschäftigt sich auch die Arbeit von Buckwalter et al. (16). Sie zeigten auf, dass MC-Fragen sehr wohl in der Lage sind, komplexere kognitive Fähigkeiten zu messen, so auch die Interpretation von Daten und Fähigkeiten, Probleme zu lösen (Level 2 und 3). Jedoch mussten Buckwalter et al. einräumen, dass die Konstruktion von komplexen Fragen mit Schwierigkeiten verbunden ist und sich möglicherweise zeitaufwendig und somit auch kostenintensiv gestaltet. Deswegen kann Information meistens auf mehreren vernünftigen Wegen interpretiert werden, und klinische Probleme können auf verschiedene Art und Weise akzeptabel gelöst werden.

Letztendlich führen diese Diskussionen zu der Frage, welche Aspekte ärztlicher Kompetenz mit dem MC-Testverfahren erfasst werden sollen und welche Aspekte tatsächlich erfasst werden. Es steht fest, dass klinische Kompetenz ein komplexes, mehrdimensionales Phänomen darstellt. Um klinische Kompetenz in einer Prüfung messen zu können, muss diese jedoch erst definiert und in ihren verschiedenen Dimensionen analysiert werden.

Eine Untersuchung der klinischen Kompetenz in ihren Komponenten wurde von Hassard et al. (34) durchgeführt. Dabei stellten sich vier voneinander unabhängige Faktoren heraus:

Faktor 1 beinhaltet kommunikative Fähigkeiten und den Umgang mit dem Patienten, erfasst also die Interaktion mit dem Patienten.

Faktor 2 steht in engem Zusammenhang mit der Interpretation von Daten und der Diagnosestellung.

Faktor 3 erfasst das Management im Sinne einer Bewältigung der Aufgaben, die der Patient in seiner Situation an den Kliniker stellt.

Faktor 4 repräsentiert die Fähigkeit, Arbeitshypothesen zu entwerfen.

Hassard (34) kommt zu dem Ergebnis, dass herkömmliche Prüfungen, wie z.B. eine reine MC-Prüfung, die beschriebenen Aspekte ärztlicher Kompetenz nicht in allen Bereichen erfasst und somit einer Ergänzung bedürfen.

b) Auswirkungen auf das studentische Lernverhalten

In diesem Punkt spiegelt sich ein wesentlicher Nachteil von MC-Prüfungen wider. MC-Fragen beeinflussen nachteilig das Lernverhalten der Studierenden, indem sie kurzfristiges und oberflächliches Auswendiglernen von Faktenwissen fördern (19, 66, 74, 88, 89).

Ebenso wird bemängelt, dass es den Studierenden beim Lernen nicht um das eigentliche Ausbildungsziel gehe, sondern nur um ein erfolgreiches Bewältigen und Bestehen der MC-Prüfungen (19). Die Akzeptanz verschiedener Beurteilungsverfahren und ihren Einfluss auf die Lehre und auf die Lernstrategien der Studierenden war Gegenstand der Untersuchung von Pabst (65). In dem Test im Anatomie-Praktikum an der Medizinischen Hochschule Hannover mit Offenen Fragen, mündlichen Prüfungen und MC-Fragen konnte gezeigt werden, dass jedes dieser Testverfahren einen großen Einfluss hatte auf das Interesse der Studierenden für Details und auf das Niveau der Vorbereitung für jede Kursstunde. Nachdem die Studierenden am Ende jeder Stunde eine mündliche Prüfung absolvierten, hatte dies bezüglich des Lernverhaltens zur Folge, dass sie die Vorlesungen besuchten, Lehrbücher zur Vorbereitung lasen und sich differenziert mit der Anatomie und Topographie der verschiedenen Regionen auseinandersetzten (65), anstatt sich auf Wesentliches zu beschränken.

Letztendlich lässt sich festhalten, dass ein ausgewogenes Verhältnis zwischen der Entwicklung klinisch-praktischer Fähigkeiten und der Aneignung theoretischen Wissens in der medizinischen Ausbildung von besonderer Bedeutung ist und ein berechtigtes Anliegen der Ausbilder darstellt. Entscheidendes Kriterium für das Lehrsystem der Fakultäten bzw. für die nationalen Prüfungskomitees sollte sein, dass die zum Einsatz kommenden Beurteilungsverfahren sowohl formal als auch inhaltlich auf die festgesetzten Ausbildungsziele hin abgestimmt werden (44).

c) Verfälschung des Ergebnisses durch vorgegebene Antworthinweise

Der in der englischsprachigen Literatur als „cueing effect“ bezeichnete Hinweiseffekt beschreibt die Tatsache, dass die Kandidaten auf einen Wiedererkennungseffekt reagieren, der das Ergebnis positiv beeinflusst.

In der Studie von McCarthy (48) werden die Ergebnisse von Studierenden verglichen, die über ein simuliertes klinisches Problem mündlich in freier Antwort („uncued“) und schriftlich mit visuellen Antworthinweisen („cued“) geprüft werden. Die Ergebnisse zeigten bessere Werte für den schriftlichen Teil der Prüfung und dort insbesondere für diejenigen Kandidaten, die in der mündlichen Prüfung schlechter abgeschnitten hatten. McCarthy vermutet, dass die mündliche („uncued“) und die schriftliche („cued“) Methode unterschiedliche Fähigkeiten erfassen. Während der Kandidat in der mündlichen Prüfung auf sein erlerntes und gespeichertes Wissen angewiesen ist, kann er sich in der schriftlichen Prüfung anhand vorgegebener Möglichkeiten durch den klinischen Fall „hindurchmanövrieren“. McCarthy kommt zu dem Schluss, dass wichtige Komponenten klinischer Kompetenz, wie z.B. Datensammlung, Anamneseerhebung und körperliche Untersuchung nicht mit Hinweiseffekten versehen sind („uncued“), wohingegen das in der Diagnosestellung sehr wohl der Fall ist („cued“). Dementsprechend seien auch unterschiedliche Verfahren notwendig, diese Fähigkeiten zu testen.

Die drei Prüfungsarten Essay-Fragen, Essay-Fragen mit Antworthinweisen und MC-Fragen waren Ausgangspunkt der Studie von McCloskey und Holland (49). Das Ergebnis zeigte, dass die Studierenden aller Gruppen in den MC-Fragen besser abschnitten als in den Essay-Fragen ohne Hinweis. Formte man die Essay-Fragen dahingehend um, dass sie Antworthinweise ähnlich denen in MC-Fragen enthielten, verbesserten sich die Resultate und erreichten ähnlich gute Werte wie für die MC-

Fragen. Somit ist der entscheidende Unterschied zwischen Essay-Fragen und MC-Fragen im Vorhandensein von Antwortmöglichkeiten begründet.

d) „Encouragement of guessing“

Unter diesem Begriff versteht man die Tatsache, dass MC-Fragen den Kandidaten dazu ermutigen, die Antwort auf eine Frage zu raten. Bei völliger Unwissenheit einer Frage beträgt die Wahrscheinlichkeit 20%, bei fünf Antworten die richtige zu raten. Da für eine falsche Antwort kein Punktabzug erfolgt, lohnt es sich auf jeden Fall, die Antwort zu raten.

Harden et al. (31) analysieren in ihrer Studie die „don't know option“, d.h. die Möglichkeit einer „Ich-weiß-nicht-Antwort“. Ist eine „don't know option“ gegeben, erfolgt bei falscher Antwort ein Punktabzug, bei richtiger Antwort ein Punktgewinn und im Falle der „don't know option“ gar keine Änderung. Ergebnis der Studie war, dass die Kandidaten ihr Resultat verbessern können, wenn sie alle ausgelassenen Antworten vervollständigen. Dieses Ergebnis lässt sich folgendermaßen erklären: Wenn Kandidaten die Lösung nicht genau wissen, so „raten“ sie die Frage nicht nach dem Zufallsprinzip, sondern lassen Teilwissen und Anhaltspunkte mit in die Antwortauswahl einfließen. Wird eine Antwort ausgelassen, so kann nicht zwischen völligem Nichtwissen und vorhandenem Teilwissen unterschieden werden. Beantworten die Kandidaten jedoch alle Fragen, so kann davon ausgegangen werden, dass der Kandidat, der mehr weiß, auch mehr Richtiganworten erzielt (31). Ebenso ist davon auszugehen, dass derjenige, der eine Frage auslässt, nicht einfach gar nichts weiß, sondern dass er sich nicht sicher ist und nichts riskieren möchte. Somit fördert die „don't know option“ den risikofreudigen und testerfahrenen Kandidaten, der weiß, dass er nichts zu verlieren hat, wohingegen der vorsichtige und zurückhaltende Kandidat benachteiligt ist. Harden et al. empfehlen daher, die „don't know option“ in MC-Prüfungen nicht mehr anzubieten.

e) Das „One-best-answer“ Problem

Ein weiterer Kritikpunkt des MC-Testverfahrens ist, dass MC-Fragen implizieren, dass es zu jeder Frage eine „beste“ oder „korrekte“ Antwort gibt. Zweifellos entspricht dies nicht dem klinischen Alltag. Im klinischen Alltag gibt es nicht immer eine „beste“ Lösung aus einer Auswahl von fünf Möglichkeiten (23).

„We seem to be living in a world of multiple choices.“, beschreibt Anbar (2) die Konsequenzen für das alltägliche Denken. Die Folge dieses „Denkens in mehreren Möglichkeiten“ sei, dass der Student oder Arzt bei der Lösung eines Problems im klinischen Alltag ebenfalls dieses Schema vor Augen habe, anstatt selber eine eigene Idee zu entwickeln und innovativ zu werden.

Zusammenfassend lässt sich sagen, dass MC-Prüfungen aufgrund ihrer hohen Objektivität und Reliabilität, ihrer guten Anwendbarkeit und geringen Entwicklungs- und Durchführungskosten einen festen Platz im Prüfungssystem in der medizinischen Ausbildung einnehmen. Nachteilige Eigenschaften wie die sehr differenziert zu betrachtenden Aussagen zur Validität, die nachteiligen Auswirkungen auf das studentische Lernverhalten und der Antwort-Hinweiseffekt sind ernst zu nehmen und zeigen die Grenzen dieses Verfahrens auf. MC-Prüfungen sind als alleiniges Testverfahren nicht geeignet, ärztliche Kompetenz in all ihren Dimensionen zu erfassen, und bedürfen innerhalb eines Prüfungssystems einer notwendigen Ergänzung durch andere Prüfungsmethoden.

1.1.7.2) Vorteile und Nachteile von Offenen Fragen

Zunächst muss erwähnt werden, dass die vorhandene Literatur für Offene Fragen wesentlich weniger umfangreich ist, als dies für die MC-Fragen der Fall ist (25). Diese Tatsache liegt darin begründet, dass MC-Fragen insgesamt viel häufiger zum Einsatz kamen und dementsprechend das Interesse an Studien größer war (25).

Einer der oft genannten Vorzüge von Offenen Fragen gegenüber MC-Fragen ist, dass Offene Fragen offensichtlich besser in der Lage sind, Wissen und Leistungen auf einem bestimmten Ausbildungsniveau zu erfassen. Dies lässt auf eine höhere Validität deuten. Jedoch werden diesem Vorteil die höheren Kosten und der größere zeitliche Aufwand entgegengehalten. Ob diese Behauptungen der Realität entsprechen, wird die folgende Gegenüberstellung zeigen. Dabei orientiert sich die Darstellung an der Auflistung der Vor- und Nachteile, wie sie in Kapitel 1.1.7.1 für die MC-Fragen zu finden ist.

a) Objektivität und Reliabilität

Da die Offenen Fragen per Hand ausgewertet werden, ist die Wahrscheinlichkeit für Auswertfehler höher. Die Literatur ist sich weitgehend auch einig, dass die Objektivität von Offenen Fragen unter derjenigen von MC-Fragen liegt. Die Objektivität der Offenen Fragen wird mit „durchschnittlich“, die der MC-Fragen mit „hoch“ bewertet (67). Die Beurteilerfehler bei MC-Fragen werden als „gering“, die der Offenen Fragen als „durchschnittlich“ eingeschätzt (42). Da diese Bewertungen lediglich relative Einschätzungen darstellen, bedürfen sie einer exakteren Bestimmung. Ausschlaggebend hierbei ist, welche Faktoren sich in welchem Ausmaß auf die Objektivität auswirken.

Hinsichtlich der Reliabilität finden sich unterschiedliche Aussagen. In der Simulationsstudie von Frary (25) finden sich für die Reliabilität von Offenen Fragen leicht höhere Werte als für die MC-Fragen. Dem gegenüber steht das „Gutachten über das Prüfungswesen“ der Planungsgruppe Medizin (67). Dieses beurteilt die Reliabilität von Offenen Fragen geringer als diejenige von MC-Fragen. In der Argumentation dafür, dass dabei die Beurteilung sehr uneinheitlich sein kann, liegt jedoch ein Widerspruch, da die Autoren der Planungsgruppe Medizin sich dabei auf die Arbeit von Frary (25) beziehen (67). Newble (55) vergleicht ebenfalls in seiner Arbeit die Reliabilität von Offenen Fragen und MC-Fragen und findet dabei etwas niedrigere Werte auf Seiten der Offenen Fragen. Feletti (24) macht die Reliabilität von Offenen Fragen von der Qualität der Modell-Antwortbögen abhängig.

b) Validität

Bezüglich der content validity, die von der inhaltlichen Zusammenstellung einer Prüfung abhängt, findet sich kein grundsätzlicher Unterschied zwischen den Offenen Fragen und den MC-Fragen. Die Kontentvalidität kann für beide Fragetypen, bei entsprechender Bemühung vorausgesetzt, als gleich gut angenommen werden.

Die criterion validity weist in der Studie von Rothman (72), in der kurze Essay-Fragen als Vertreter der Offenen Fragen und MC-Fragen verglichen werden, etwa gleich hohe Werte für beide Fragetypen auf. Zu einem ähnlichen Ergebnis kommt auch Nnodim (60). Während die criterion validity für die MC-Fragen und die Essays etwa gleich hoch ist (0,66 bzw. 0,63), bewirkt eine Kombination dieser Verfahren eine Steigerung der Validität (0,70).

Hinsichtlich der construct validity zeigt Nnodim (60) auf, dass Essay Fragen die Gesamtleistung von Prüflingen besser einschätzen können als MC-Fragen. Dies unterstreicht die gute kriterienbezogene Validität von Offenen Fragen, führt aber auch zu der Frage, welche Fähigkeiten mit Offene Fragen erfasst werden. Einen Hinweis darauf, dass mit Offenen Fragen andere Fähigkeiten („skills“) registriert werden als mit MC-Fragen, findet man in der Untersuchung von Norman (63). Rothman (72) kommt zu dem Schluss, dass MC-Fragen spezifisches Faktenwissen erfassen, während kurze Essay-Fragen die Fähigkeit ermitteln, klare Antworten auf Probleme zu geben, die in kurzen Fallstudien dargelegt sind.

Insgesamt beurteilt Frary (25) die Validität von Offenen Fragen als leicht erhöht gegenüber den MC-Fragen. Zu dem gleichen Ergebnis kommt die Planungsgruppe Medizin in ihrem „Gutachten über das Prüfungswesen“ (67).

c) Kosten, Aufwand, Ökonomie

Offene Fragen sind in ihrer Entwicklung, Durchführung und Auswertung aufwendiger, was einen größeren zeitlichen und finanziellen Aufwand impliziert (23, 24, 55, 61, 67).

Eine qualitative Aussage zum zeitlichen Aufwand macht die Arbeit von Stalenhoef-Halling et al. (79). Bei gleicher Reliabilität erfordert eine Prüfung 18 Offene Fragen bzw. 180 MC-Fragen. Für die Konstruktion und Auswertung von 18 Offenen Fragen bedeutet dies einen zeitlichen Aufwand von 225 Stunden, für die 180 MC-Fragen jedoch nur einen Aufwand von 144 Stunden. Dies stellt eine klare Präferenz zugunsten der MC-Fragen dar.

d) Auswirkungen auf das Lernverhalten

In der Arbeit von Stalenhoef-Halling et al. (79) zum Einfluss unterschiedlicher Testverfahren (MC-Fragen und Offene Fragen) auf das Lernverhalten der Studierenden wurden diese sechs Wochen vor und nach dem Test befragt. Vor der Prüfung gaben 56,6% der Kandidaten an, für die Offenen Fragen mehr zu lernen, nach der Prüfung waren es 49%. Bei der Frage nach der Lernstrategie (Faktenwissen versus Verstehen, Auswendiglernen) zeigen die Ergebnisse, dass die Kandidaten nach der Prüfung über die neu eingeführten Offenen Fragen neutraler urteilen.

Nach Yang et al. (91) stellen Offene Fragen einen größeren Arbeitsaufwand für die Studierenden dar, wohingegen Hakstian (29) zu dem Ergebnis kommt, dass die Art der erwarteten Prüfung keinen Einfluss auf das Lernverhalten und die Art der Prüfungsvorbereitung hat.

e) Antworthinweiseffekte und Ratefaktoren, „One-best-answer“-Problem

Da Offene Fragen keine Antworthinweise im Sinne von Auswahl-Antworten enthalten, gibt es auch keine Antworthinweiseffekte. Bei der Beantwortung von Offenen Fragen spielen Ratefaktoren auch eine Rolle, allerdings eine weitaus geringere als bei MC-Fragen. Wenn der Kandidat zur Beantwortung der Offenen Frage zwischen mehreren möglichen Antworten auswählen muss und sich nicht sicher ist, besteht zwar auch eine gewisse Ratewahrscheinlichkeit, jedoch muss er zuvor sinnvolle Antwortmöglichkeiten bereits aktiv generiert haben.

Dass Hinweiseffekt und Ratefaktor zu besseren Leistungen in MC-Fragen führen und das Ergebnis positiv verfälschen ist das Resultat der Studie von Newble (55). Er schließt daraus, dass Offene Fragen das tatsächliche Wissen besser erfassen als MC-Fragen, wohingegen MC-Fragen das Wissen der Kandidaten überschätzen. McCloskey und Holland (49) liefern ebenfalls den Nachweis dafür, dass der Antworthinweiseffekt in MC-Fragen für das Zustandekommen besserer Ergebnisse im Vergleich zu Essay-Fragen entscheidend ist.

Das „One-best-answer“-Problem, einer der wesentlichen Nachteile von MC-Fragen, entfällt bei den Offenen Fragen. Die Auswertung von Offenen Fragen anhand eines Modellantwort-Bogens sieht zwar durchaus eine bestimmte Antwort vor, beinhaltet aber einen weitaus größeren Spielraum zwischen den Extremen „richtig“ und „falsch“, als dies bei MC-Fragen der Fall ist.

f) Akzeptanz

Die Mehrheit der Kandidaten bevorzugt eine Prüfungs-Kombination aus Offenen Fragen und MC-Fragen gegenüber einer Prüfung mit lediglich einer der beiden Komponenten (49, 86). Wallace (86) stellte fest, dass zwischen der Prüfungsleistung und der Präferenz für eine bestimmte Prüfungsmethode keine Korrelation besteht.

Zusammenfassend lässt sich feststellen, dass Reliabilität und Validität von Offenen Fragen als gleich gut oder besser beurteilt wird als diejenige von MC-Fragen. Bezüglich der construct validity scheinen die Offenen Fragen den MC-Fragen sogar überlegen zu sein. Ein höherer zeitlicher und finanzieller Aufwand auf Seiten der Offenen Fragen ist unumstritten, jedoch beinhalten diese keine Ratefaktoren und Hinweiseffekte.

Offene Fragen haben also gegenüber MC-Fragen den Vorteil einer höheren Validität (construct validity) um den Preis eines größeren zeitlichen und finanziellen Aufwandes.

1.2) Zielsetzung der Arbeit

Bei den Überlegungen zur Entstehung dieser Arbeit waren folgende Faktoren von Bedeutung:

Auf Grund der Neuerungen nach der 9. Änderung der Approbationsordnung und deren Umsetzung zum Wintersemester 2003/2004 bestehen derzeit mehrere Modelle zur Verbesserung der medizinischen Ausbildung. An der Universität Heidelberg z.B. erfolgt die studentische Lehre nach dem Modellstudiengang „Heicumed“, während die Universität Regensburg die gesetzlichen Neuerungen durch die 9. Approbationsordnung in das schon bestehende Kurssystem integriert. Jedes dieser Modelle befasst sich auch explizit mit dem Prüfungssystem, das den in Kapitel 1.1.5.2 genannten Gütekriterien genügt, sich effizient gestaltet und den gesellschaftlichen und fachspezifischen Veränderungen angepasst ist, denen die medizinische Ausbildung insgesamt unterworfen ist.

Im Zentrum dieser Arbeit liegt die Beobachtung der Veränderung des Wissensstandes im Fachbereich Orthopädie im Laufe der medizinischen Ausbildung. Dabei stellt sich die Frage, wo, d.h. in welchem Studienabschnitt das orthopädische Wissen seinen Ausgang nimmt, wo es die stärkste Veränderung erfährt und inwiefern dieses Wissen sich stabilisiert.

Die Beurteilung des Wissenszuwachses wird durch die Wissensüberprüfung zu Beginn (Prä-Test) und am Ende (Post-Test) des Orthopädiepraktikums ermöglicht. Somit liefert diese Arbeit den Lehrenden eine Rückmeldung über die Präsenz der vermittelten Lehrinhalte und zeigt im interuniversitären Vergleich die Schwerpunkte der Lehre der beiden Universitäten mit den unterschiedlichen Lehrsystemen auf.

Ebenfalls bleibt zu überprüfen, welchen Einfluss der jeweilige Ausbildungsstand auf die Leistungen der Probanden im Multiple-Choice-Teil des Fragebogens im Vergleich zu den Ergebnissen für die Offenen Fragen hat.

Es galt folgende Hypothesen zu falsifizieren/verifizieren:

1. Im 1. klinischen und 2. klinischen Semester sind die schlechtesten Ergebnisse des Tests zu finden, da sich das Grundlagenwissen aus der Vorklinik nur bedingt auf die im Test spezifischen orthopädischen Fragen transferieren lässt.
2. Im 3. klinischen und 4. klinischen Semester (Prä-Test) nimmt das orthopädische Wissen seinen Anfang, da diese Studierenden bereits die einwöchige Einführung in die Orthopädie besucht haben. Ebenfalls steigt hier das Interesse am Fachbereich Orthopädie (Teil III des Fragebogens), da die Studierenden mit dem Fachgebiet Orthopädie konfrontiert werden.
3. Das orthopädische Wissen erfährt im Laufe des Orthopädie-Praktikums (Vergleich Prä-Test mit Post-Test) die stärkste Veränderung, da die Studierenden für die Abschlussprüfung am Ende des Orthopädie-Praktikums sowohl praktische als auch theoretische Fähigkeiten aufweisen müssen und eine ganztägige Beschäftigung mit dem Fachgebiet Orthopädie stattfindet.
4. Im 4. klinischen und 5. klinischen Semester erfolgt im Vergleich mit den Werten des Post-Tests ein Rückgang des orthopädischen Wissens, das sich dann im 6. klinischen Semester kaum mehr verändert und stagniert.
5. In allen Testgruppen finden sich bessere Werte bei der Beantwortung der MC-Fragen als bei den Offenen Fragen, da sich eine 20%ige Wahrscheinlichkeit für das Ankreuzen der richtigen Antwort ergibt und bei den Offenen Fragen aktives Reproduzieren von Faktenwissen abgefragt wird.
6. Die Zufriedenheit der Studierenden (Teil I des Fragebogens) und das Interesse am Fachgebiet Orthopädie (Teil III des Fragebogens) zeigen im Laufe der medizinischen Ausbildung immer mehr ein ambivalentes Verhalten auf.
7. Der Schwierigkeitsgrad der gestellten Fragen verschiebt sich mit Zunahme des Ausbildungsstandes von der Kategorie „schwierig“ hin zur Kategorie „leicht“.

2.) Material und Methode

2.1) Problemstellung

An dieser Stelle seien die Fragestellungen, die anhand der vorliegenden Arbeit zur Zufriedenheit der Studenten mit der aktuellen Studiensituation unter Berücksichtigung der modifizierten Approbationsordnung, zum Vergleich zweier schriftlicher Prüfungsmethoden, zum Wissensstand im Fachbereich Orthopädie und zum Interesse am Fachbereich Orthopädie analysiert werden sollen, noch einmal zusammenfassend genannt:

Es stellt sich die Frage, wie sich die Zufriedenheit der Studierenden im Fachbereich Medizin in Hinblick auf Inhalte, Struktur und Umsetzung durch die Dozenten unter Berücksichtigung der modifizierten Approbationsordnung äußert. Ebenso Gegenstandspunkt der Befragung ist, welche Änderungen die Studenten an der aktuellen Studiensituation vornehmen würden.

Hinsichtlich der Prüfungsmethoden wird sowohl nach der bevorzugten Prüfungsart der Studierenden gefragt als auch die Anwendung von Multiple-Choice Fragen im Vergleich zu Offenen Fragen diskutiert.

Anhand der Auswertungen aus Teil II des Fragebogens (Wissensstand im Fachbereich Orthopädie) lässt sich der Wissensstand im Fachbereich Orthopädie im Laufe der medizinischen Ausbildung dokumentieren. Durch den Vergleich des Wissensstandes vor und nach der Absolvierung des Orthopädiepraktikums (Prä-Test und Post-Test) ergibt sich eine Aussage über den Wissenszuwachs der Studenten. Da die Durchführung der Befragung an Universitäten mit unterschiedlichen Lehrsystemen durchgeführt wurde, kann ein interuniversitärer Vergleich hinsichtlich des Wissensstandes, des Wissenszuwachses und der Schwerpunkte der Lehre gezogen werden.

Die Einschätzung und Akzeptanz der Befragung ist ebenso ein zu diskutierender Gegenstandspunkt der vorliegenden Arbeit wie auch das Interesse am Fachbereich Orthopädie in den unterschiedlichen Testgruppen.

2.2) Testentwicklung und Fragebogenerstellung

Ausgehend von der oben aufgeführten Problemstellung wurde ein anonymer Fragebogen entwickelt, der in drei Aufgabenteile gegliedert ist und außerdem eine kurze Einleitung zum Sinn des Fragebogens und eine Erläuterung zum Ausfüllen des Fragebogens enthält.

Eine ungekürzte Abbildung des Fragebogens ist im Anhang einzusehen (s. S. 173 ff).

Teil I des Fragebogens besteht aus neun Multiple-Choice Fragen und einer Offenen Frage, Teil II aus zwölf Multiple-Choice Fragen und drei Offenen Fragen, Teil III aus fünf Multiple-Choice Fragen, so dass sich eine Gesamtanzahl von 30 Fragen ergibt.

2.2.1) Teil I des Fragebogens (Zufriedenheit der Studierenden)

Teil I des Fragebogens besteht aus neun Multiple-Choice Fragen (MC-Fragen) und einer offenen Frage. Bei acht der neun MC-Fragen stehen fünf, bei einer MC-Frage drei Antwortmöglichkeiten zur Auswahl, von denen die zutreffende ausgewählt werden soll. Bei allen neun MC-Fragen wurde der Typ A („One of five“) verwendet. Die Typ A-Frage besteht aus einem Fragenstamm und in diesem Fall fünf bzw. drei möglichen Antworten, aus denen die zutreffende auszuwählen ist. Bei den MC-Fragen mit fünf Antwortmöglichkeiten gab es je nach Frageninhalt die Skalierungen „sehr gut“, „gut“, „unentschieden“, „schlecht“, „sehr schlecht“ (Frage 1, 2, 4, 7) bzw. „trifft völlig zu“, „trifft zu“, „unentschieden“, „trifft nicht zu“, „trifft gar nicht zu“ (Frage 5, 6, 8) bzw. „sehr zufrieden/absolut überzeugt“, „zufrieden/überzeugt“, „unentschieden“, „nicht zufrieden/nicht ganz überzeugt“, „gar nicht zufrieden/Zweifel am Studium“ (Frage 9). Bei der MC-Frage mit drei Antwortmöglichkeiten standen die Antwortmöglichkeiten „MC-Fragen“, „Offene Fragen“ und „mündliche Prüfung“ zur Verfügung (Frage 3). Bei der offenen Frage war die Antwort anhand von Stichpunkten zu geben (Frage 10).

Entscheidend für die Zusammenstellung von Teil I des Fragebogens waren die aktuelle Studiensituation nach der 9. Änderung der Approbationsordnung und vor allem die Sicht der Studierenden. Dabei erfolgte die Auswahl anhand der folgenden inhaltlichen Planung:

- 1.) Fragen zur neuen Approbationsordnung (Frage 1, 2)
- 2.) Frage zum Prüfungswesen (Frage 3)
- 3.) Fragen zum Studium hinsichtlich Inhalt, Struktur und Umsetzung durch die Dozenten (Frage 4, 5, 6, 7,8)
- 4.) Fragen zur Zufriedenheit der Studierenden (Frage 9, 10)

Die zehn Fragen aus Teil I des Fragebogens sind im Anhang (s. S. 173 f) einzusehen.

2.2.2) Teil II des Fragebogens (Wissensstand im Fachbereich Orthopädie)

Teil II des Fragebogens besteht aus zwölf MC-Fragen und drei Offenen Fragen. Bei den MC-Fragen handelt es sich um Fragen mit fünf Antwortmöglichkeiten, von denen die zutreffende ausgewählt werden soll. Es wurden neun Fragen vom Typ A („One of five“) und drei Fragen vom Typ K verwendet. Die Typ A-Frage besteht aus einem Fragenstamm und fünf möglichen Antworten, aus denen die richtige auszuwählen ist. Bei der Typ K-Frage folgt auf den Fragenstamm bzw. auf eine Aussage eine Liste mit Ergänzungen, die mit Nummern versehen sind und die in bestimmten, vorgegebenen Kombinationen die Lösungsmöglichkeit darstellen.

Die drei Offenen Fragen stellen sogenannte „short-answer-questions“ dar. Per definitionem sind diese „short-answer-questions“ sehr spezifische Fragen, die in Stichworten innerhalb eines bestimmten Umfangs beantwortet werden sollen. In diesem Fall zielt die Frage auf eine Aufzählung einzelner Komponenten eines Krankheitsbildes oder anatomischen Struktur ab.

Die zwölf MC-Fragen und die drei Offenen Fragen aus Teil II des Fragebogens sind im Anhang (s. S. 175 ff) einzusehen.

Entscheidend für die Zusammenstellung von Teil II des Fragebogens war die inhaltliche Komponente. Dabei erfolgte die Auswahl anhand der folgenden inhaltlichen Planung:

Zum Einsatz kamen Fragen zu typisch orthopädischen Symptomen, zur Statik, zu orthopädischen Zeichen in der Diagnostik, zur orthopädischen Diagnostik, zu spezifisch-orthopädischen Krankheitsbildern und zu anatomischen Begriffen.

Anhand dieser Aufteilung sollte erreicht werden, ein möglichst breites Spektrum orthopädischer Thematik zu erfassen. In Hinblick auf die Auswertung stellt sich die Frage, inwiefern der Ausbildungsstand der jeweils untersuchten Testgruppe die Leistungen innerhalb der einzelnen Fragenkategorien beeinflusst. Die im Fragebogen verwendeten MC-Fragen stammen aus Altklausuren, die vom Institut für Lehre des Orthopädischen Universitätsklinikums Regensburg zur Erlangung des Orthopädie-Scheins in den letzten Jahren verwendet wurden. Somit entspricht der Schwierigkeitsgrad der gestellten Fragen dem Niveau der Fragen im Zweiten Abschnitt der Ärztlichen Prüfung nach neuer AO.

Neben der oben genannten Einteilung der Fragen nach Form und Inhalt ist eine weitere Einteilung möglich, entsprechend des zur Lösung erforderlichen Denkprozesses. Dabei wird die Einteilung kognitiver Prozesse nach McGuire vorgenommen (50). Diese besagt, dass es drei Niveaus kognitiver Leistungen gibt:

- a) Niveau 1: Abruf und Wiedererkennen von Faktenwissen
- b) Niveau 2: Verstehen und Interpretation von Informationen
- c) Niveau 3: Problemlösendes Denken

Eine Einteilung nach dieser Systematik ergibt folgendes Bild für die 15 Fragen:

- 1.) In neun Fragen von 15 wird Faktenwissen abgerufen, entsprechend Niveau 1 [Fragen 13, 14, 15, 16, 17, 21, 22, 23, 25 (60%)].
- 2.) In drei Fragen von 15 sind Verstehen und Interpretation die erforderlichen Prozesse, entsprechend Niveau 2 [Fragen 11, 19, 24 (20%)].
- 3.) Drei von 15 Fragen fordern die komplexe Fähigkeit des Problemlösens („problem solving“) [Fragen 12, 18, 20 (20%)].

2.2.3) Teil III des Fragebogens (Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)

Im letzten Teil des Fragebogens wurden die Probanden anhand von fünf Fragen zu dem Schwierigkeitsgrad der Fragen, zum Interesse am Fachbereich Orthopädie und zur Durchführung einer solchen Evaluation bzw. Feststellung des Fachwissensstandes befragt. Die Einschätzung des Schwierigkeitsgrades der Fragen aus Teil II erfolgte getrennt für die MC-Fragen (Frage 26) und die Offenen Fragen (Frage 27), wobei jeweils fünf Antwortmöglichkeiten vorgegeben waren: „sehr leicht“, „leicht“, „angemessen“, „schwierig“ und „sehr schwierig“.

Zum Interesse am Fachgebiet sollten ebenfalls zwei MC-Fragen Aufschluss geben. Die erste Frage bezieht sich dabei auf das allgemeine Interesse am Fachgebiet Orthopädie, wobei die Antwortmöglichkeiten „sehr großes Interesse vorhanden“, „Interesse vorhanden“, „unentschieden“, „wenig Interesse vorhanden“ und „gar kein Interesse vorhanden“ zur Auswahl standen (Frage 28). Die zweite Frage zielt auf das spezielle Interesse am Fachbereich Orthopädie hinsichtlich des Berufsziels als Facharzt für Orthopädie ab und gab als Antwortmöglichkeiten die Skalierung „streben den Beruf als Orthopäde/in an“, „kann ich mir vorstellen“, „unentschieden“, „kann ich mir nicht vorstellen“ und „kommt für mich sicher nicht in Frage“ vor (Frage 29). Anhand der letzten Frage (Frage 30) konnten die Probanden sich dazu äußern, ob sie die Durchführung einer solchen Evaluation zum Thema Studienordnung und einer Feststellung des Fachwissensstandes befürworten oder nicht. Dazu standen die Antwortmöglichkeiten „ja“ und „nein“ zur Auswahl, wobei bei Ankreuzen der letzteren der Zusatz „warum nicht“ mit vier Zeilen Platz zur Begründung sich anschloss.

Die fünf Fragen aus Teil III des Fragebogens sind im Anhang (s. S. 178) einzusehen.

2.3) Untersuchungsaufbau

Die Untersuchungen fanden im Zeitraum von April 2006 bis August 2006 an den Universitätskliniken Regensburg und Heidelberg statt. Befragt wurden an der Universität Regensburg die Studenten vom 1. klinischen Semester bis zum 6. klinischen Semester, wobei diejenigen Studenten, die im Sommersemester 2006 das Praktikum der Orthopädie belegt hatten (3. und 4. klinisches Semester), zweimal befragt wurden:

einmal zu Beginn des Kurses und einmal am Ende des Kurses, so dass sich ein Prä- und ein Post-Test ergab. An der Universität Heidelberg wurde die Untersuchung nur in einer Gruppe des Orthopädie-Praktikums durchgeführt, die jedoch aus Studierenden unterschiedlicher Semester besteht. Diesen Testgruppen (Prä- und Post-Test) gilt besondere Bedeutung, da sich dabei sowohl die Veränderung des Wissensstandes infolge des Wissenszuwachses erfassen, als auch ein interuniversitärer Vergleich hinsichtlich Wissensstand und Lehre ziehen lässt. Ebenfalls kann auf Grund der unterschiedlichen Studiensysteme in Heidelberg (Heicumed = Heidelberger Curriculum Medicinale) und Regensburg (Kurssystem) ein weiterer Vergleich bezüglich der Lehrmodelle vorgenommen werden. Für den Prä-Test und für den Post-Test wurde der gleiche Fragebogen unverändert wieder verwendet.

2.4) Untersuchungsgut

An der Universität Regensburg wurden die Studenten vom 1. klinischen Semester bis zum 6. klinischen Semester befragt. Bei den Studenten im 3. klinischen Semester und bei einem Teil der Studenten im 4. klinischen Semester handelte es sich um die Teilnehmer des Orthopädie-Praktikums. Diese Gruppe wurde zweimal getestet (Prä- und Post-Test).

Um einen Vergleich zwischen dem Wissensstand vor und nach dem Praktikum Orthopädie und dem daraus resultierenden Wissenszuwachs von Studenten unterschiedlicher Universitäten ziehen zu können, wurde die Befragung auch an der Universität Heidelberg durchgeführt. Das Orthopädie-Praktikum findet dort in der Stiftung Orthopädische Universitätsklinik, Schlierbacher Landstraße 200a, 6900 Heidelberg statt. Somit ist ein interuniversitärer Vergleich hinsichtlich des Wissensstandes der Studenten und der Lehre möglich. An der Universität Heidelberg nahmen Studierende aus dem 2. klinischen bis 6. klinischen Semester am Orthopädie-Praktikum teil. An dieser Universität wurde die Befragung nur bei den Teilnehmern des Orthopädie-Praktikums im Stundenplan des Sommersemesters 2006 durchgeführt, was somit insgesamt drei Prä- und Post-Test-Gruppen ergibt. Besondere Bedeutung fällt dabei auf den Vergleich der Universität Heidelberg mit der Universität Regensburg, da die Universität Heidelberg den Modellstudiengang „Heicumed“ in der Ausbildung der Studenten als Studiensystem verwendet.

2.5) Durchführung der Befragung

Die Befragung der Studenten an der Universität Regensburg erfolgte in geschlossenen Gruppen im Rahmen von Pflichtveranstaltungen oder im Anschluss an schriftliche Klausuren, um eine möglichst hohe Teilnehmeranzahl pro Semester zu erreichen. Die Genehmigungen des jeweiligen Institutsleiters und des jeweiligen Dozenten der betroffenen Pflichtveranstaltung war vorhanden. Zu Beginn der Befragung wurden die Studenten kurz über den Hintergrund der anonymen Untersuchung und über den zur Verfügung stehenden Zeitraum zur Beantwortung der Fragen informiert. Sie wurden aufgefordert, den Fragebogen eigenständig und nach bestem Wissen zu beantworten. Allen befragten Studenten standen 12 Minuten der Unterrichtszeit zur Verfügung. Somit fand also eine Befragung unter klausurähnlichen Bedingungen statt. Die Befragung und Auswertung der Fragebögen erfolgte für alle Gruppen anonym.

Die Studenten des 3. klinischen Semesters und ein Teil des 4. klinischen Semesters der Universität Regensburg waren die Teilnehmer des Orthopädie-Praktikums. Diese erhielten den Fragebogen als Prä-Test jeweils zu Beginn des Orthopädie-Praktikums. Da das Orthopädie-Praktikum in vier Gruppen zeitversetzt absolviert wird, fand der Post-Test am letzten Tag des Praktikums der jeweiligen Gruppe bzw. am folgenden Unterrichtstag in einer anderen Pflichtveranstaltung statt. Die Studierenden dieser beiden Gruppen wurden zu Beginn der Befragung explizit darauf hingewiesen, dass die Fragebögen anonym ausgewertet werden und somit auch keine Wertung bzw. Benotung stattfindet, die sich auf den Erwerb des Orthopädie-scheins bzw. dessen Benotung auswirken könnte.

An der Universität Heidelberg erfolgte die Befragung am ersten Tag zu Beginn des Orthopädie-Praktikums (Prä-Test) und am letzten Tag (Post-Test). Auch an dieser Universität fand die Untersuchung unter den gleichen Voraussetzungen und Rahmenbedingungen statt. Die Studenten erhielten die gleiche Einführung und ebenfalls den gleichen Zeitraum von 12 Minuten zur Beantwortung der Fragen, um einen späteren Vergleich ohne Verzerrungen in der Ausführung der Untersuchung zu gewährleisten.

2.6) Auswertung und Analyse

Die Auswertung des Fragebogens erfolgte getrennt für Teil I, II und III.

2.6.1) Auswertung von Teil I (Zufriedenheit der Studierenden)

Für die zehn Fragen in Teil I des Fragebogens erfolgte eine Ermittlung der relativen Häufigkeiten der Angaben zur neuen Approbationsordnung, zur bevorzugten Prüfungsmethode, zum Inhalt und zur Strukturierung des Studiums, zur Umsetzung durch die Dozenten und zur Zufriedenheit mit der aktuellen Studiensituation. Hier bleibt zu prüfen, ob die subjektiven Einstufungen einen Zusammenhang zum jeweiligen Ausbildungsstand zeigen und wie sich jene Einstufungen vor allem in Hinblick auf die Zufriedenheit mit der aktuellen Studiensituation im interuniversitären Vergleich verhalten. Hinsichtlich der Strukturierung des Studiums und der Umsetzung durch die Dozenten bietet sich eine Vergleichsmöglichkeit zwischen der Universität Regensburg mit dem Kurssystem und der Universität Heidelberg mit dem Heicumed-Modellstudiengang.

2.6.2) Auswertung von Teil II (MC-Fragen)

Das Lösungsschema für die MC-Fragen ist im Anhang (siehe S. 179) einzusehen.

Von den fünf Antwortmöglichkeiten jeder Frage war jeweils nur eine Antwort richtig. Für jede richtige Antwort wurde 1 Punkt vergeben, Abzüge für falsch beantwortete Fragen wurden nicht vollzogen.

Für jede Testperson wurde die Anzahl der richtig beantworteten Fragen absolut (Punktwert) und in Prozentwerten ermittelt. Innerhalb der einzelnen Testgruppen wurden Mittelwerte in absoluten Zahlen (Punktwerte) und Prozentwerten sowie die Standardabweichung berechnet.

Die statistische Analyse erfolgte unter Verwendung des Student'schen t-Tests für paare und unpaare Stichproben. Das Signifikanzniveau wird auf $p < 0,05$ festgelegt.

Ausgewertet wurde dieser Teil des Fragebogens nur für die Kandidaten, die für jede der 12 Fragen eine Lösungsmöglichkeit angegeben hatten. Dieses Vorgehen wurde aus folgenden Gründen gewählt: Für die MC-Fragen war keine „ich-weiß-nicht“-Option vorgegeben. Unabhängig vom Wissensstand war es jeder Testperson möglich, eine Auswahlantwort zu raten, wenn er die korrekte Antwort nicht wusste. Es wurde bei der Durchführung der Untersuchung ausdrücklich darauf hingewiesen, dass alle MC-Fragen zu bearbeiten sind. Um hier die gleichen Voraussetzungen für alle Kandidaten wirksam zu machen, konnten nur vollständig bearbeitete MC-Fragen gewertet werden.

Eine Ausnahme bilden hier die Gruppen 3, 4 und 8. Wegen des Vergleichs der Ergebnisse von Prä-Test und Post-Test wurden hier ausnahmslos alle Fragebögen gewertet, auch wenn diese nur unvollständig ausgefüllt waren. In der Auswertung erscheint daher auch der Anteil der unbeantworteten Fragen für diese drei Testgruppen.

2.6.3) Auswertung von Teil II (Offene Fragen)

Anhand eines im Vorfeld festgelegten Modell-Antwortbogens („model answers“) erfolgte die Korrektur der Offenen Fragen. Dieser enthielt die erwarteten Stichworte und ein Schema zur Punktvergabe. In Frage 23, 24 und 25 waren jeweils maximal 4 Punkte zu erreichen. Falschantworten hatten keinen Punktabzug bzw. keine Punktabzüge zur Folge.

Die statistische Analyse erfolgte unter Verwendung des Student´schen t-Tests für paare und unpaare Stichproben. Das Signifikanzniveau wird auf $p < 0,05$ festgelegt.

Ausgewertet wurde dieser Teil des Fragebogens nur für die Kandidaten, die die Offenen Fragen auch beantwortet hatten. Diejenigen Testpersonen, die diesen Teil vollständig unausgefüllt gelassen hatten, gingen nicht in die Gesamtbewertung mit ein. Dieses betrifft vor allem das 1. klinische Semester und das 2. klinische Semester, da hier nur ein Teil der Testpersonen die Offenen Fragen beantwortete. Dieses Vorgehen erfolgte aus folgenden Gründen:

Die Offenen Fragen wurden in den Gruppen 1 und 2 in mehreren Fragebögen nicht beantwortet. Um einen Vergleich der Ergebnisse dieser Gruppen mit den Ergebnissen der anderen Gruppen überhaupt zu ermöglichen, wurden Fragebögen mit null Punkten für die Offenen Fragen nicht in die Wertung für diesen Teil des Fragebogens aufgenommen.

Korrekturschema für Teil II (Offene Fragen):

Frage 23: Nennen Sie die muskulären Strukturen der Rotatorenmanschette!

Jeweils 1 Punkt wurde vergeben für folgende Stichworte:

- a) musculus supraspinatus
- b) musculus infraspinatus
- c) musculus subscapularis
- d) musculus teres minor

Bei 4 richtigen Antworten erhielt die Testperson die Höchstzahl von 4 Punkten.

In einigen Lehrbüchern wird das „ligamentum coracohumerale“ ebenfalls zu den Strukturen der Rotatorenmanschette gezählt. Für diese Antwort gab es allerdings keinen Punkt, da in der Fragestellung explizit nach den muskulären Strukturen gefragt wird.

Frage 24: Nennen Sie 4 Ursachen, die zu einem Beckenschiefstand führen können!

Jeweils 1 Punkt wurde vergeben für folgende Kernaussagen:

- a) Hüftabduktionskontraktur
- b) Hüftadduktionskontraktur
- c) Kniebeugekontraktur
- d) Spitzfuß

Folgende Stichworte wurden ebenfalls jeweils mit einem Punkt gewertet:

Skoliose, Fraktur, Hüftdysplasie, Beinlängendifferenz (BLD) und Gluteusatrophie.

Bei 4 oder mehr richtigen Antworten erhielt die Testperson die Höchstzahl von 4 Punkten.

Frage 25: Nennen Sie 4 Frühzeichen, die bei einem Perthes gelten!

Jeweils ein Punkt wurde vergeben für folgende Kernaussagen:

- a) Belastungsabhängige Schmerzen in der Hüfte
- b) Schmerzen im Knie
- c) Schmerzhaftes Innenrotationszeichen im Hüftgelenk
- d) Schmerzhaftes Abduktionszeichen im Hüftgelenk

Folgende Stichworte wurden ebenfalls jeweils mit einem Punkt gewertet:

Schonhinken, Gehschwäche, positives Drehmannzeichen.

Bei 4 oder mehr richtigen Antworten erhielt die Testperson die Höchstzahl von 4 Punkten.

Für jede Testperson wurde der Punktwert für jede der drei Fragen sowie die Punktzahl insgesamt absolut und in Prozentwerten ermittelt. Ebenso wurde für die einzelnen Testgruppen der Mittelwert für die Beantwortung jeder der drei Fragen sowie der Mittelwert und die Standardabweichung für die in Teil II (Offene Fragen) erreichte Gesamtpunktzahl (absolut und in Prozentwerten) berechnet.

Zur Erstellung des Modell-Antwortbogens dienten die wichtigsten aktuellen Lehrbücher für Studierende der Medizin (15, 26, 59, 73, 90).

2.6.4) Auswertung von Teil III (Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)

Für die Fragen in Teil III des Fragebogens erfolgte eine Ermittlung der relativen Häufigkeiten. Diese wurden für die Einstufung des Schwierigkeitsgrades der MC-Fragen und der Offenen Fragen, für das Interesse am Fachbereich Orthopädie allgemein und im speziellen als Berufsziel sowie zur Akzeptanz der Befragung errechnet.

Hier bleibt zu überprüfen, ob diese subjektive Schwierigkeitseinstufung der gestellten Fragen und das persönliche Interesse am Fachbereich Orthopädie einen Zusammenhang zum jeweiligen Ausbildungsstand zeigen. Zur Auswertung dieses Teils

gehört auch der interuniversitäre Vergleich hinsichtlich des allgemeinen und speziellen Interesses am Fachbereich Orthopädie.

2.6.5) Auswertung des Vergleichs von MC-Fragen und Offene Fragen

Ein direkter Vergleich zwischen den in Teil II gestellten MC-Fragen und Offenen Fragen ist nicht möglich, da die MC-Fragen und die Offenen Fragen inhaltlich nicht übereinstimmen und keine „matched-pairs“-Technik verwendet wurde. „Matched-pairs“-Technik bedeutet die Konstruktion zweier Fragebögen mit inhaltlich gleichen Fragen, die jeweils zur Hälfte aus MC-Fragen und Offenen Fragen bestehen. Dabei entsprechen die MC-Fragen des einen Fragebogens inhaltlich den Offenen Fragen des zweiten Fragebogens und umgekehrt.

Sehr wohl möglich ist dagegen der intra- und interpersonelle Vergleich des Abschneidens der Testpersonen in der jeweiligen Prüfungsart. Für die einzelnen Gruppen wurde die Differenz der Ergebnisse aus dem Bereich der MC-Fragen und der Offenen Fragen ermittelt.

Gleichfalls wurden die prozentualen Leistungsveränderungen von Gruppe zu Gruppe getrennt für die MC-Fragen und die Offenen Fragen berechnet und miteinander verglichen.

2.6.6) Auswertung des Vergleichs von Prä-Test und Post-Test der Universitäten Regensburg und Heidelberg

Ebenso wie für alle Testgruppen erfolgte auch hier die Auswertung von Teil I, II und III. Zusätzlich fand hier für jede Testgruppe ein Vergleich der Ergebnisse im Prä- und Post-Test statt, sowohl getrennt für Teil I, Teil II (MC-Fragen), Teil II (Offene Fragen) und Teil III als auch für den Gesamtpunktwert aus Teil II (Wissensstand im Fachbereich Orthopädie). Für diese Testgruppen wurden daraus Mittelwerte und Standardabweichungen berechnet. Die Veränderung der Ergebnisse wurde in Prozentwerten ausgedrückt und in den Testgruppen der Universitäten Regensburg und Heidelberg miteinander verglichen.

3.) Ergebnisse

Insgesamt nahmen an dieser Befragung 506 Testpersonen teil. Davon waren 476 Studenten an der Universität Regensburg immatrikuliert, 30 Studenten an der Universität Heidelberg.

Tabelle 1: Gruppeneinteilung für das Untersuchungskollektiv an der Universität Regensburg

Gruppe	Ausbildungsstand	Immatrikulierte Studenten	n = Anzahl der Testpersonen
Gruppe 1	1. klinisches Semester	37	n = 37
Gruppe 2	2. klinisches Semester	123	n = 121
Gruppe 3	3. klinisches Semester	33	n = 33
➤ 3-Prä	Prä-Test	33	n = 33
➤ 3-Post	Post-Test	33	n = 33
Gruppe 4	4. klinisches Semester	42	n = 42
➤ 4-Prä	Prä-Test	42	n = 42
➤ 4-Post	Post-Test	42	n = 42
Gruppe 5	4. klinisches Semester	88	n = 86
Gruppe 6	5. klinisches Semester	42	n = 42
Gruppe 7	6. klinisches Semester	118	n = 115
Summe		483	n = 476

Um eine möglichst hohe Anzahl an Studierenden zu erreichen, wurden die Fragebögen meist im Anschluss an Klausuren oder in Pflichtveranstaltungen verteilt, so dass 98,55% der in den klinischen Semestern an der Universität Regensburg immatrikulierten Studenten an der Befragung teilnahmen. Die Restquote von 1,45% lässt sich durch krankheitsbedingtes Fehlen erklären.

Tabelle 2: Gruppenaufteilung für das Untersuchungskollektiv an der Universität Heidelberg

Gruppe	Ausbildungsstand	n = Anzahl der Testpersonen
Gruppe 8	2. klinisches Semester	n = 15
	3. klinisches Semester	n = 4
	4. klinisches Semester	n = 10
	6. klinisches Semester	n = 1
	➤ 8-Prä	n = 30
➤ 8-Post	n = 30	
Summe		n = 30

An der Universität Heidelberg wurde eine Gruppe des Orthopädie-Praktikums im Sommersemester 2006 zufällig ausgewählt, an der die Befragung durchgeführt wurde. Die Studierenden wurden am ersten Tag zu Beginn des Orthopädie-Praktikums und am letzten Tag des Orthopädie-Praktikums befragt. Auf Grund des Modellstudiengangs Heicumed befinden sich in einer Gruppe Studierende aus unterschiedlichen klinischen Semestern.

Tabelle 3: Gruppenaufteilung für das gesamte Untersuchungskollektiv

Gruppe	Lehreinrichtung	n = Anzahl der Testpersonen
Gruppe 1-7	Universität Regensburg	n = 476
Gruppe 8	Universität Heidelberg	n = 30
Summe		n = 506

3.1) Rücklauf

Die Befragung der Studierenden wurde in den Gruppen 1, 2, 5, 6 und 7 im Anschluss an Klausuren durchgeführt. Alle ausgeteilten Fragebögen konnten zurückerhalten werden, so dass sich für diese Gruppen ein Rücklauf von 100% ergibt.

Die Gruppen 3 und 4 waren die Teilnehmer des Orthopädie-Praktikums der Universität Regensburg. Hier erfolgte der Prä-Test am ersten Tag des Orthopädie-Praktikums. Der Post-Test wurde aufgrund einer Aufteilung in kleinere Untergruppen im Laufe des Praktikums in der ersten Woche nach Absolvierung des Orthopädie-Praktikums im Rahmen einer anderen Pflichtveranstaltung durchgeführt. Für diese beiden Gruppen wurde ebenfalls ein Rücklauf von 100% erzielt.

Die Studierenden der Gruppe 8 waren die Teilnehmer des Orthopädie-Praktikums an der Universität Heidelberg. Hier wurden die Fragebögen am ersten Tag (Prä-Test) und am letzten Tag (Post-Test) des Orthopädie-Praktikums ausgeteilt und nach Beantwortung der Fragen wieder eingesammelt, so dass die Rücklaufquote ebenfalls 100% beträgt.

Somit konnte in allen Gruppen eine Rücklaufquote von 100% erreicht werden.

3.2) Ergebnisse der Fragebogenauswertung für Teil I (Zufriedenheit der Studierenden)

In Teil I des Fragebogens erfolgte eine Ermittlung der relativen Häufigkeiten der Angaben zur neuen Approbationsordnung (Fragen 1, 2), zur bevorzugten Prüfungsmethode (Frage 3), zum Inhalt und zur Strukturierung des Studiums sowie zur Umsetzung durch die Dozenten (Fragen 4, 5, 6, 7, 8) und zur Zufriedenheit mit der aktuellen Studiensituation (Fragen 9, 10).

Im Folgenden werden die Ergebnisse für die Fragen 1 bis 10 im Einzelnen dargestellt:

Frage 1) Einschätzung des Wissensstandes zu den Änderungen durch die neue Approbationsordnung

Tabelle 4: Ergebnisse für Frage 1

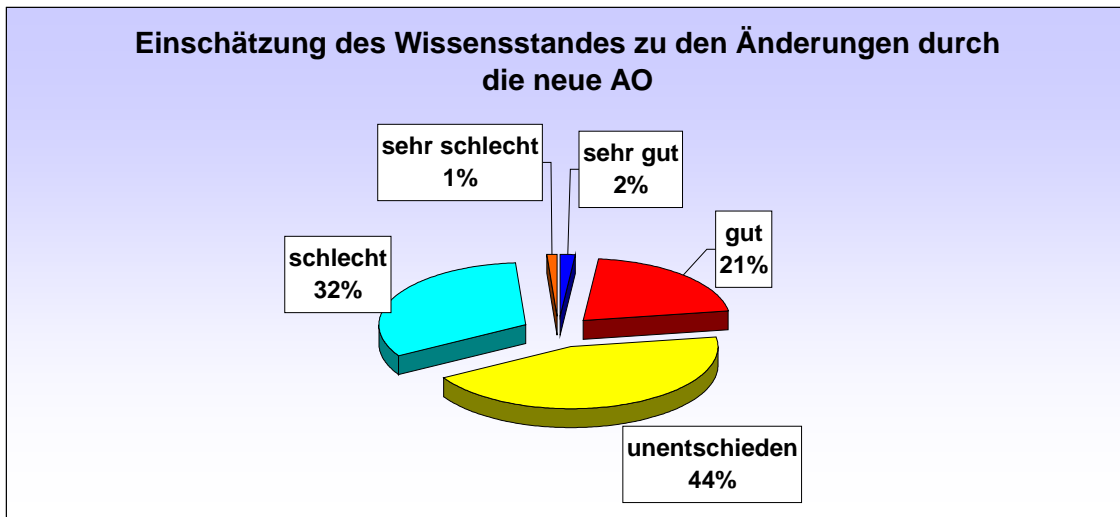
Gruppe	n	Prozent				
		sehr gut	gut	unentschieden	schlecht	sehr schlecht
1	37	5,41	13,50	37,84	37,84	5,41
2	121	4,13	23,97	38,84	33,06	0,00
3-Prä	33	0,00	27,27	42,42	30,30	0,00
3-Post	33	0,00	24,24	45,45	30,30	0,00
4-Prä	42	0,00	21,43	50,00	26,19	2,38
4-Post	42	2,38	21,43	38,10	38,10	0,00
5	86	2,33	19,77	45,35	30,23	2,33
6	42	2,38	14,29	45,24	35,71	2,38
7	115	0,00	11,30	46,96	41,74	0,00
8-Prä	30	0,00	33,33	53,33	13,33	0,00
8-Post	30	0,00	40,00	50,00	10,00	0,00

Insgesamt zeigt sich, dass sich in allen Gruppen der prozentuale Höchstwert in der Kategorie „unentschieden“ befindet, von 37,48% in Gruppe 1 bis 53,33% in Gruppe 8-Prä. Nur in den Gruppen 1 und 4-Post finden sich ebenfalls gleich hohe Werte für die Kategorie „schlecht“.

In den Gruppen 1-7 (Universität Regensburg) sind die Werte für die Kategorie „schlecht“ alle prozentual höher als für die Kategorie „gut“. Gegensätzliches tritt bei den Gruppen 8-Prä und 8-Post (Universität Heidelberg) auf. Hier wird die Kategorie „gut“ höher eingestuft als die Kategorie „schlecht“.

Für die Extremaussagen „sehr gut“ und „sehr schlecht“ finden sich die einzigen einstelligen prozentualen Werte, von 0,00% bis 5,41%.

Diagramm 2: Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Wissensstandes zu den Änderungen durch die neue AO



Frage 2) Anzahl der Staatsexamina („Hammerexamen“ am Ende)

Tabelle 5: Ergebnisse für Frage 2

Gruppe	n	Prozent				
		sehr gut	gut	unentschieden	schlecht	sehr schlecht
1	37	2,70	2,70	35,14	37,84	21,62
2	121	3,31	4,13	30,58	37,19	24,79
3-Prä	33	3,03	0,00	33,33	36,36	27,27
3-Post	33	3,03	6,06	24,24	42,42	24,24
4-Prä	42	0,00	9,52	9,52	52,38	28,57
4-Post	42	0,00	7,14	4,76	54,76	33,33
5	86	0,00	6,98	5,81	55,81	31,40
6	42	0,00	4,76	9,52	54,76	30,95
7	115	0,00	4,35	10,43	53,91	31,30
8-Prä	30	0,00	13,33	30,00	40,00	16,67
8-Post	30	0,00	10,00	30,00	40,00	20,00

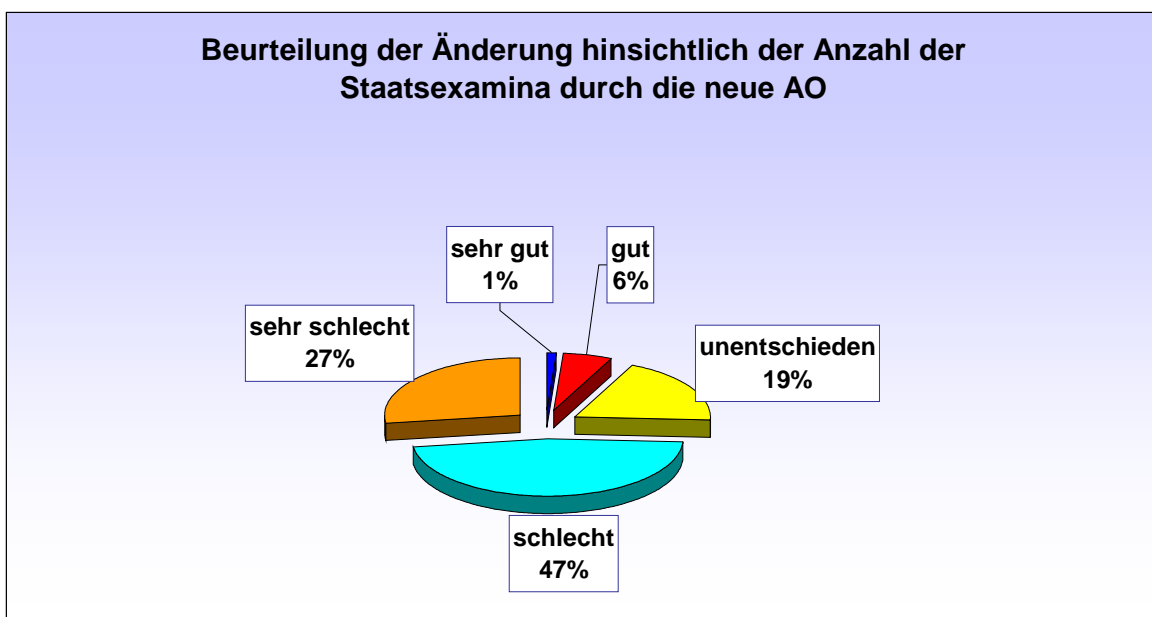
Die Kategorie „schlecht“ erhält in allen Gruppen (Gruppen 1-8) die höchsten Werte, von 36,36% in Gruppe 3-Prä bis 55,81% in Gruppe 5.

In den Gruppen 1-7 (Universität Regensburg) finden sich für die Kategorien „sehr gut“ und „gut“ nur einstellige prozentuale Werte, von 0,00% bis 9,52%. In den Gruppen 8-Prä und 8-Post (Universität Heidelberg) werden in der Kategorie „gut“ zweistellige prozentuale Werte erreicht (13,33% bzw. 10,00%).

In allen Gruppen (Gruppen 1-8) finden sich einstellige Werte für die Antwort „sehr gut“ (0,00% bis 3,31%), während die andere Extremaussage „sehr schlecht“ Werte von 16,67% (Gruppe 8-Prä) bis 33,33% (Gruppe 4-Post) erreicht.

Für die Kategorie „unentschieden“ unterscheiden sich die Werte für die Gruppen 1, 2, 3 und 8 (24,24% bis 35,14%) von den Werten für die Gruppen 4, 5, 6 und 7 (4,76% bis 10,34%).

Diagramm 3: Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung der Änderung hinsichtlich der Anzahl der Staatsexamina durch die neue Approbationsordnung



Frage 3) Bevorzugte Prüfungsmethode

Tabelle 6: Ergebnisse für Frage 3

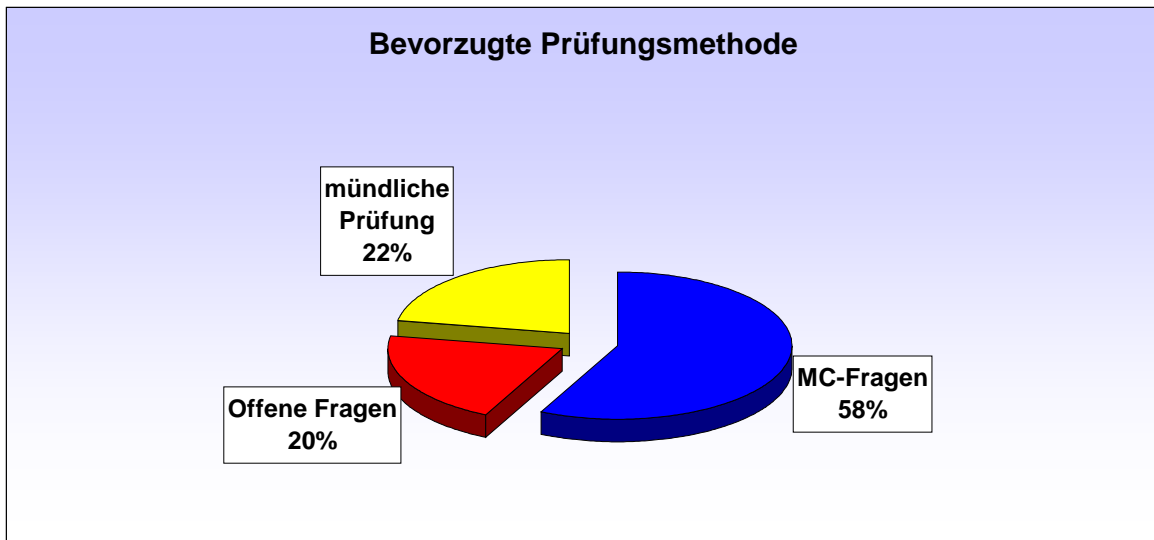
Gruppe	n	Prozent		
		MC-Fragen	Offene Fragen	Mündliche Prüfung
1	37	75,68	18,91	5,41
2	121	59,50	19,83	20,66
3-Prä	33	60,61	12,12	27,27
3-Post	33	45,45	30,30	24,24
4-Prä	42	54,76	21,43	23,81
4-Post	42	45,24	33,33	21,43
5	86	47,67	25,58	26,74
6	42	52,38	23,81	23,81
7	115	59,13	18,26	22,61
8-Prä	30	70,00	6,67	23,33
8-Post	30	76,67	0,00	23,33

Für die Antwort „MC-Fragen“ finden sich in allen Gruppen (Gruppen 1-8) die höchsten prozentualen Werte, die sich im Bereich von 45,24% in Gruppe 3-Post bis hin zu 76,67% in Gruppe 8-Post befinden.

In den Gruppen 1, 2, 5, 6 und 7 liegen die Werte für die Antwort „Offene Fragen“ zwischen 18,26% (Gruppe 7) und 25,58% (Gruppe 5). In dieser Kategorie kommt es im Vergleich der Gruppen 3-Prä zu 3-Post und 4-Prä zu 4-Post zu einem prozentualen Zuwachs (von 12,12% auf 30,30% und von 21,43% auf 33,33%). Gruppe 8-Prä (6,67%) und Gruppe 8-Post (0,00%) verzeichnen für die Offenen Fragen die niedrigsten Werte.

Für die Antwort „Mündliche Prüfung“ finden sich in den Gruppen 2-8 Werte von 20,66% bis 27,27%, nur in Gruppe 1 weicht der Wert von 5,41% erheblich ab.

Diagramm 4: Mittelwerte aller Gruppen (Gruppen 1-8) zur bevorzugten Prüfungsmethode



Frage 4) Vermittlung der Inhalte innerhalb der Praktika

Tabelle 7: Ergebnisse für Frage 4

Gruppe	n	Prozent				
		sehr gut	gut	unentschieden	schlecht	sehr schlecht
1	37	2,70	45,95	43,24	8,11	0,00
2	121	3,31	50,41	38,84	6,61	0,83
3-Prä	33	0,00	57,58	33,33	9,09	0,00
3-Post	33	9,09	57,58	33,33	0,00	0,00
4-Prä	42	14,29	66,67	16,67	2,38	0,00
4-Post	42	9,52	57,14	21,43	11,90	0,00
5	86	6,98	48,84	31,40	10,47	2,33
6	42	4,76	42,86	33,33	16,67	2,38
7	115	1,74	36,52	39,13	20,00	2,61
8-Prä	30	0,00	80,00	20,00	0,00	0,00
8-Post	30	13,33	63,33	23,33	0,00	0,00

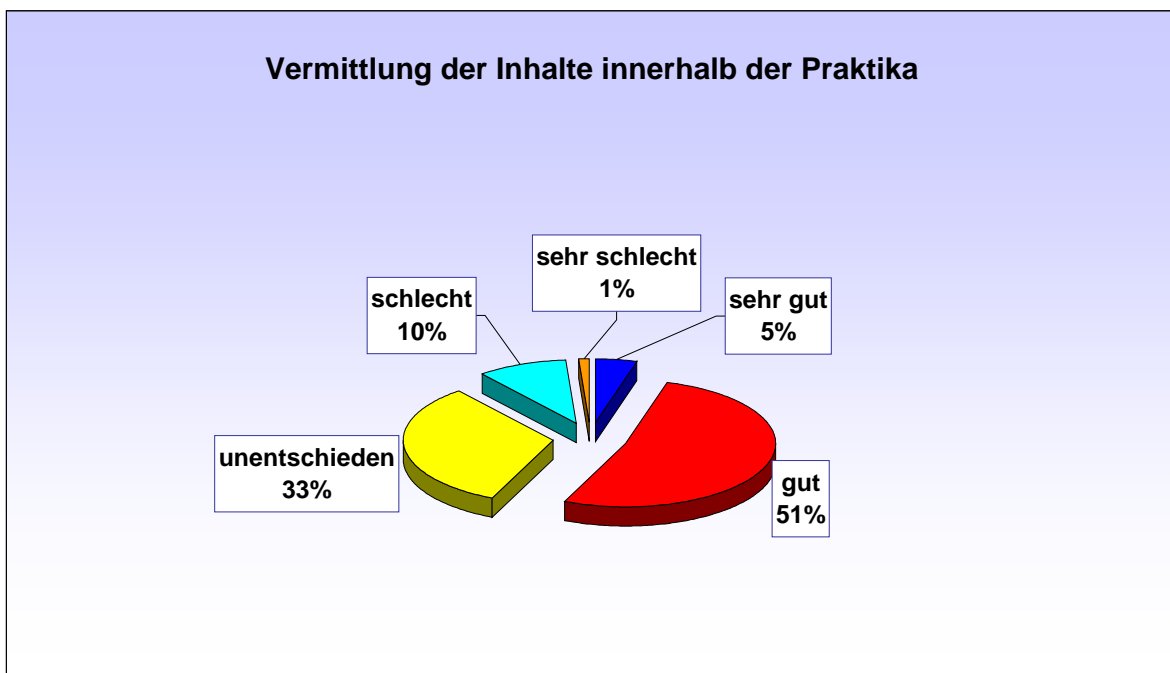
In den Gruppen 1, 2, 3, 4, 5, 6 und 8 erhält die Kategorie „gut“ die höchsten Werte (42,86% bis 80,00%). Die Antwort „unentschieden“ weist in diesen Gruppen die zweithöchsten Werte auf (16,67% bis 43,24%).

In Gruppe 7 liegt der Wert für die Antwort „gut“ nur knapp (36,52%) unter dem für die Kategorie „unentschieden“ (39,13%). Somit ist hier im Vergleich mit den anderen Gruppen der höchste mit dem zweithöchsten Wert ausgetauscht.

In Kategorie „sehr schlecht“ finden sich ausschließlich niedrige einstellig prozentuale Werte (0,00% bis 2,61%).

Eine starke Varianz lässt sich für die Kategorien „sehr gut“ (0,00% - 14,29%), „unentschieden“ (16,67% bis 43,24%) und „schlecht“ (0,00% bis 20,00%) feststellen, in denen die prozentualen Werte stark voneinander abdriften.

Diagramm 5: Mittelwerte aller Gruppen (Gruppen 1-8) zur Vermittlung der Inhalte innerhalb der Praktika



Frage 5) Ausreichende Vorstellung von Patienten in den Vorlesungen

Tabelle 8: Ergebnisse für Frage 5

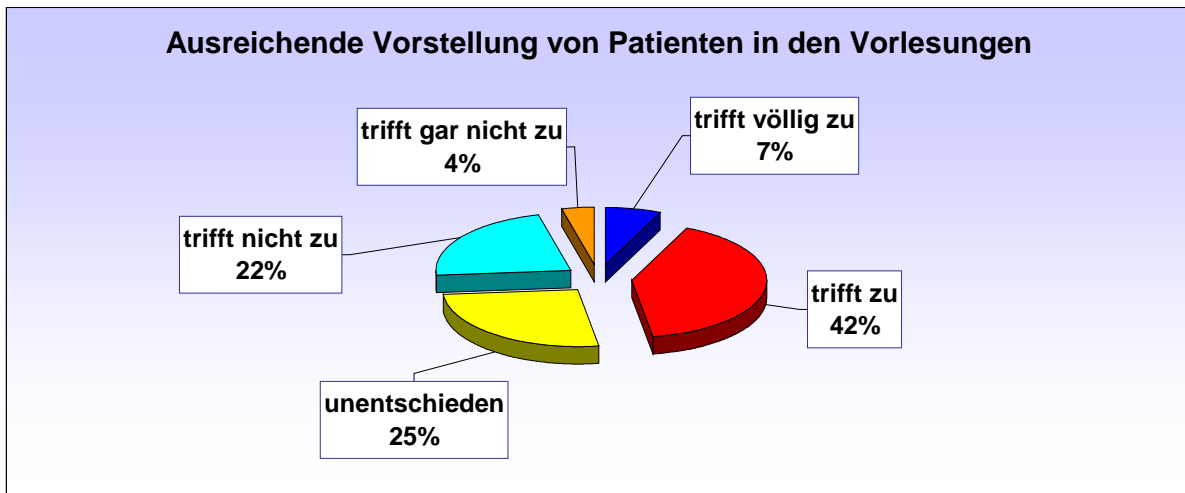
Gruppe	n	Prozent				
		trifft völlig zu	trifft zu	unentschieden	trifft nicht zu	trifft gar nicht zu
1	37	21,63	72,97	2,70	2,70	0,00
2	121	4,96	57,85	24,73	12,40	0,00
3-Prä	33	9,09	54,55	21,21	15,15	0,00
3-Post	33	18,18	45,45	21,21	15,15	0,00
4-Prä	42	0,00	52,38	23,81	19,05	4,76
4-Post	42	7,14	33,33	42,86	14,29	2,38
5	86	4,65	40,70	34,88	17,44	2,33
6	42	4,76	30,95	26,19	30,95	7,14
7	115	2,61	19,13	18,26	47,83	12,17
8-Prä	30	10,00	23,33	40,00	23,33	3,33
8-Post	30	6,67	50,00	26,67	13,33	3,33

In der Kategorie „trifft gar nicht zu“ treten in den Gruppen 1, 2, 3, 4, 5, 6, und 8 nur einstellig prozentuale Werte auf (0,00% bis 7,14%). Eine Ausnahme stellt hier die Gruppe 7 dar, in der sich mit 12,17% in dieser Kategorie ein höherer Wert befindet.

In den Gruppen 4-Post (42,86%) und 8-Prä (40,00%) beinhaltet die Kategorie „unentschieden“ den Höchstwert, in Gruppe 7 die Kategorie „trifft nicht zu“ (47,83%). In den anderen Gruppen (Gruppen 1, 2, 3-Prä, 3-Post, 4-Prä, 5, 6 und 8-Post) vertritt die Kategorie „trifft zu“ den Wert mit dem höchsten Prozentsatz (30,95% bis 72,97%).

Die Prozentwerte für die Antwort „trifft völlig zu“ variieren von 2,61% in Gruppe 7 bis hin zu 21,63% in Gruppe 1.

Diagramm 6: Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung einer ausreichenden Vorstellung von Patienten in den Vorlesungen



Frage 6) Ausreichende Vorstellung von Patienten in den Praktika

Tabelle 9: Ergebnisse für Frage 6

Gruppe	n	Prozent				
		trifft völlig zu	trifft zu	unentschieden	trifft nicht zu	trifft gar nicht zu
1	37	5,41	59,46	18,91	13,51	2,70
2	121	4,13	33,06	28,10	34,71	0,00
3-Prä	33	9,09	30,30	21,21	39,39	0,00
3-Post	33	27,27	51,52	12,12	9,09	0,00
4-Prä	42	4,76	59,52	26,19	9,52	0,00
4-Post	42	28,57	47,62	16,67	7,14	0,00
5	86	17,44	48,84	16,28	15,12	2,33
6	42	14,29	50,00	16,67	16,67	2,38
7	115	11,30	53,91	15,65	16,52	2,61
8-Prä	30	20,00	56,67	16,67	6,67	0,00
8-Post	30	20,00	63,33	16,67	0,00	0,00

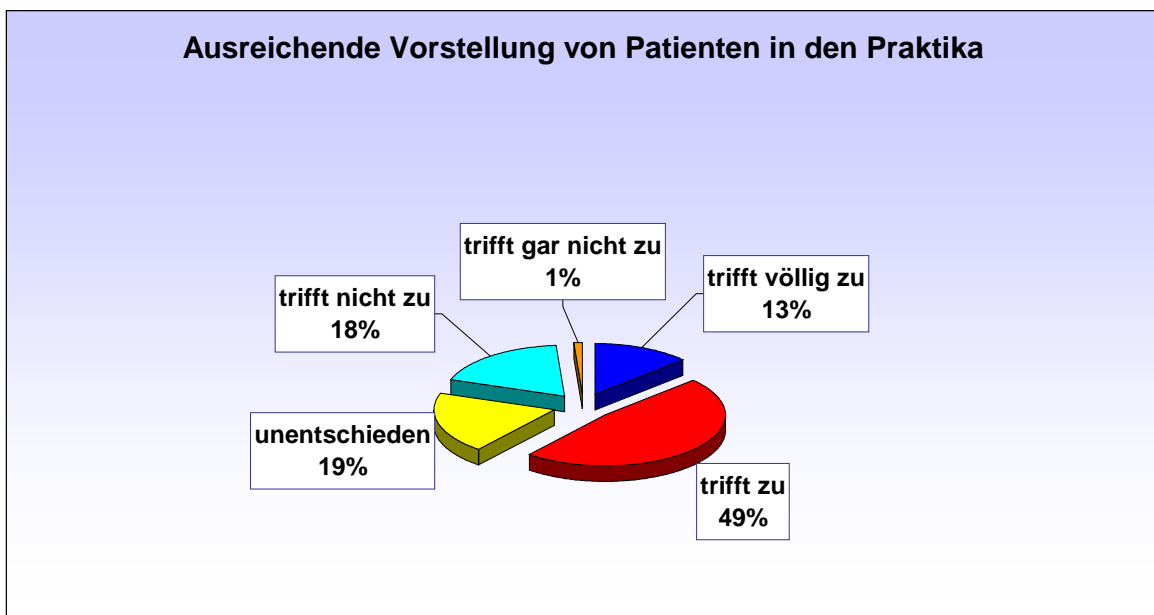
In den Gruppen 2 und 3-Prä wird der Höchstwert durch die Kategorie „trifft nicht zu“ vertreten (34,71% bzw. 39,39%), in allen anderen Gruppen finden sich die höchsten Werte für die Kategorie „trifft zu“ (47,62% bis 63,33%).

Für die Aussage „trifft gar nicht zu“ treten in allen Gruppen die niedrigsten Werte auf (0,00% bis 2,70%).

In der Kategorie „trifft völlig zu“ (4,13% bis 28,57%) variieren die Werte stark voneinander. Dabei fallen für diese Antwort in den Gruppen 3-Post und 4-Post mit 27,27% und 28,57% die Werte am höchsten aus und weisen im Vergleich zu den Werten der zugehörigen Gruppen 3-Prä bzw. 4-Prä einen prozentualen Zuwachs von 18,18% bzw. 23,81% auf.

Für die Kategorie „trifft nicht zu“ (0,00% bis 39,39%) lässt sich ebenfalls eine starke Varianz der Prozentwerte feststellen, wobei hier die Gruppen 2 und 3-Prä mit 34,71% und 39,39% die Topwerte verkörpern.

Diagramm 7: Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung einer ausreichenden Vorstellung von Patienten in den Praktika



Frage 7) Strukturierung des Studiums in Bezug auf die Blockpraktika

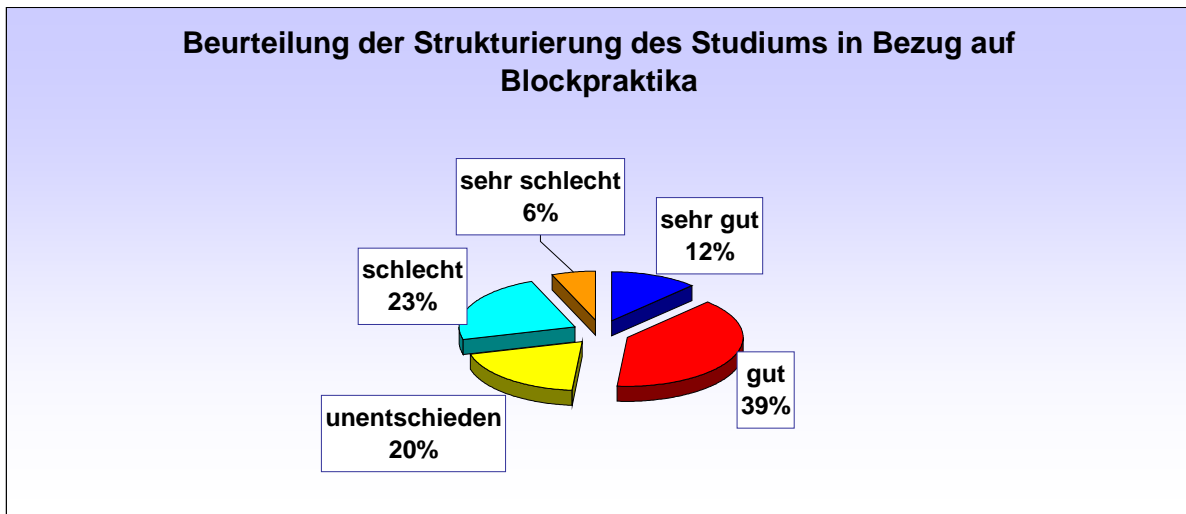
Tabelle 10: Ergebnisse für Frage 7

Gruppe	n	Prozent				
		sehr gut	gut	unentschieden	schlecht	sehr schlecht
1	37	27,03	29,73	21,62	19,51	8,11
2	121	15,70	47,11	23,14	12,40	1,65
3-Prä	33	12,12	54,55	24,24	9,09	0,00
3-Post	33	24,24	36,36	18,18	21,21	0,00
4-Prä	42	14,29	28,57	28,57	21,43	7,14
4-Post	42	0,00	26,19	23,81	42,86	7,14
5	86	9,30	30,23	16,28	34,88	9,30
6	42	7,14	33,33	19,05	30,95	9,52
7	115	6,96	33,91	12,17	33,91	13,04
8-Prä	30	13,33	66,67	16,67	3,33	0,00
8-Post	30	10,00	56,67	30,00	3,33	0,00

Allgemein lässt sich feststellen, dass sich die prozentuale Mehrheit für die Antworten „sehr gut“, „gut“ und „unentschieden“ in den Gruppen 1, 2 und 3 mit steigendem Ausbildungsstand in Richtung der Antworten „unentschieden“, „schlecht“ und „sehr schlecht“ verlagert.

In der Kategorie „sehr gut“ weist Gruppe 1 mit 27,03% den höchsten prozentualen Wert auf, während Gruppe 4-Post mit 0,00% den Minimalwert für diese Kategorie repräsentiert. Die Antwort „gut“ erhält in allen Gruppen vergleichsweise hohe Werte, die sich von 26,19% in Gruppe 4-Post bis 54,55% in Gruppe 3-Prä erstrecken. In den Gruppen der Universität Heidelberg erreicht diese Kategorie noch höhere prozentuale Werte (66,67% bzw. 56,67%). In der Kategorie „schlecht“ treten Werte von 3,33% (Gruppen 8-Prä und 8-Post) bis 42,86% (Gruppe 4-Post) auf. Im Vergleich der Gruppen 3-Prä zu 3-Post und 4-Prä zu 4-Post steigen die Werte um mindestens das Doppelte an (9,09% auf 21,21% bzw. 21,43% auf 42,86%).

Diagramm 8: Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung der Strukturierung des Studiums in Bezug auf die Blockpraktika



Frage 8) Ausreichender praktischer Unterricht von 4 Wochen Chirurgie, 5 Wochen Innere Medizin und 2 Wochen Orthopädie

Tabelle 11: Ergebnisse für Frage 8

Gruppe	n	Prozent				
		trifft völlig zu	trifft zu	unentschieden	trifft nicht zu	trifft gar nicht zu
1	37	2,70	8,11	40,54	32,43	16,22
2	121	0,00	4,13	23,14	47,11	25,62
3-Prä	33	0,00	6,06	24,24	48,48	21,21
3-Post	33	0,00	12,12	21,21	45,45	21,21
4-Prä	42	2,38	14,29	7,14	57,14	19,05
4-Post	42	0,00	9,52	14,29	50,00	26,19
5	86	1,16	6,98	11,63	38,37	41,86
6	42	0,00	7,14	9,52	28,57	54,76
7	115	0,87	5,22	5,22	26,09	62,61
8-Prä	30	0,00	6,67	76,67	13,33	3,33
8-Post	30	0,00	13,33	76,67	10,00	0,00

Die Antwort „trifft völlig zu“ erreicht in allen Gruppen die mit Abstand niedrigsten Werte, von 0,00% bis 2,70%.

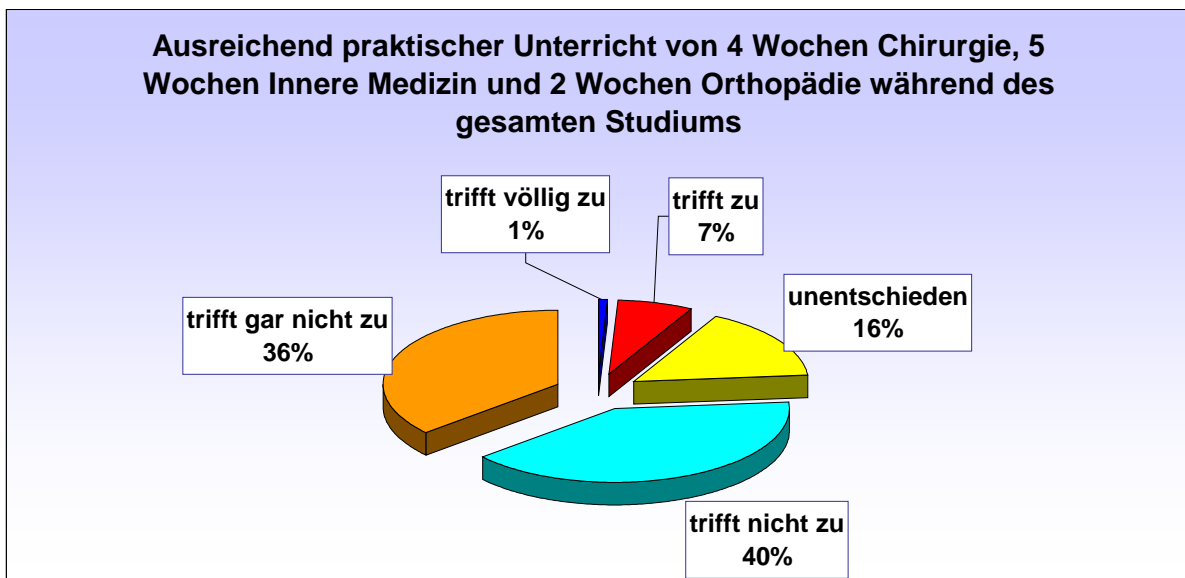
Für die Kategorie „trifft zu“ finden sich Werte von 4,13% (Gruppe 2) bis 14,29% (Gruppe 4-Prä).

In der Kategorie „unentschieden“ variieren die Werte. Der niedrigste prozentuale Wert erscheint in der Auswertung für Gruppe 7 (5,22%), der höchste Wert für die Gruppen 8-Prä und 8-Post mit jeweils 76,67%.

Für die Antworten „trifft nicht zu“ und „trifft gar nicht zu“ erscheinen in den Gruppen 1-7 durchweg zweistellige prozentuale Werte. Dabei überwiegen in den Gruppen 1 bis 4-Post die Werte der Kategorie „trifft nicht zu“ (32,43% bis 57,14%), während in den Gruppen 5, 6 und 7 die Topwerte in der Kategorie „trifft gar nicht zu“ vertreten sind (41,86% bis 62,61%).

Bei den Gruppen der Universität Heidelberg (Gruppen 8-Prä und 8-Post) kristallisiert sich der absolute Höchstwert für die Antwort „unentschieden“ heraus (76,67% bzw. 76,67%), während sich die restlichen Prozente auf die anderen Kategorien verteilen.

Diagramm 9: Mittelwerte der Gruppen 1-6 (Universität Regensburg) zur Beurteilung von ausreichend praktischem Unterricht



Frage 9) Zufriedenheit als Medizinstudent mit aktueller Situation

Tabelle 12: Ergebnisse für Frage 9

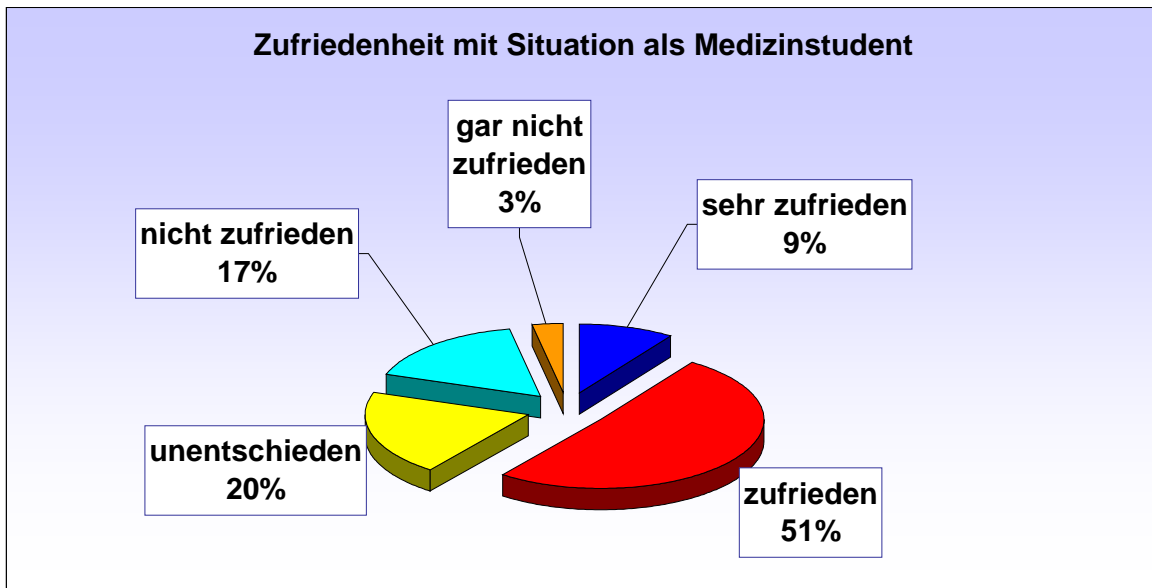
Gruppe	n	Prozent				
		sehr zufrieden	zufrieden	unentschieden	nicht zufrieden	gar nicht zufrieden
1	37	16,22	62,16	13,51	5,41	2,70
2	121	14,88	52,07	19,83	11,57	1,65
3-Prä	33	18,18	51,52	21,21	9,09	0,00
3-Post	33	15,15	57,58	12,12	15,15	0,00
4-Prä	42	4,76	69,05	11,90	14,29	0,00
4-Post	42	4,76	64,29	16,67	14,29	0,00
5	86	4,65	50,00	23,26	18,60	3,49
6	42	4,76	35,71	26,19	23,81	9,52
7	115	3,48	25,22	31,30	32,17	7,83
8-Prä	30	13,33	76,67	6,67	3,33	0,00
8-Post	30	3,33	83,33	13,33	0,00	0,00

Bis auf Gruppe 7 werden die Topwerte durch die Kategorie „zufrieden“ vertreten. In den Gruppen 1 bis 5 reichen die Werte von 50,00% (Gruppe 5) bis 64,29% (Gruppe 4-Post). In den Gruppen mit höherem Ausbildungsstand (Gruppen 6 und 7) erfolgt eine Abnahme der Werte für die Kategorie „zufrieden“, während die Werte für die Kategorien „unentschieden“ (26,19% bzw. 31,30%) und „nicht zufrieden“ (23,81% bzw. 32,17%) ansteigen und die Antwort „nicht zufrieden“ in Gruppe 7 sogar den Höchstwert darstellt.

Die Werte der Kategorie „sehr zufrieden“ erreichen in den Gruppen 1 bis 3-Post zweistellige Werte (15,15% bis 18,18%), fallen daraufhin mit steigendem Ausbildungsstand in den Gruppen 4-Prä (4,76%) bis Gruppe 7 (3,48%) ab.

Die Antwort „gar nicht zufrieden“ übertrifft nur in den beiden Gruppen 6 (9,52%) und 7 (7,83%) die Antwort „sehr zufrieden“, sonst stellen die Werte für diese Kategorie mit 0,00% bis 3,49% die niedrigsten Prozentzahlen dar.

Diagramm 10: Mittelwerte aller Gruppen (Gruppen 1-8) zur Zufriedenheit mit Situation als Medizinstudent



Frage 10) Änderungspunkte zur aktuellen Studiensituation

In Frage 10 konnten die Teilnehmer der Untersuchung Ihre Änderungsvorschläge bezüglich der aktuellen Studiensituation anhand von Stichpunkten abgeben.

Da in allen Gruppen sehr ähnliche Stichpunkte aufgeführt wurden, sind in der folgenden Tabelle die am häufigsten genannten Antworten aufgelistet und mit den jeweiligen dazugehörigen Prozentwerten versehen. Wurde in einer Gruppe eine Antwortmöglichkeit der Tabelle nicht erwähnt, so wurde das betreffende Feld der Tabelle mit „ -- " versehen.

Es ist zu erwähnen, dass diese Frage nicht in jedem Fragebogen ausgefüllt wurde und auch der Umfang der Beantwortung von nur einem Stichwort bis hin zu einem sechszeiligen Text variierte.

In Gruppe 1 haben 54,05% der Untersuchungsteilnehmer die Frage 10 bearbeitet, 78,51% waren es in Gruppe 2, 84,84% in Gruppe 3-Prä, 81,81% in Gruppe 3-Post, 80,95% in Gruppe 4-Prä, 83,33% in Gruppe 4-Post, 63,95% in Gruppe 5, 57,14% in Gruppe 6, 76,52% in Gruppe 7, 80,00% in Gruppe 8-Prä und 66,67% in Gruppe 8-Post.

Tabelle 13: Ergebnisse für Frage 10

Gruppe	n	Prozent								
		kein Hammer-examen	keine Block-Praktika	wichtige Fächer länger	mehr Praxis	mehr Zeit zum Lernen	keine Anwesenheitspflicht	nur Kurzzeit-gedächtnis	Vergütung PJ	keine Benotung
1	37	5,41	13,51	--	8,11	5,41	8,11	--	--	10,81
2	121	17,36	15,70	29,75	22,31	9,92	1,65	--	--	4,13
3-Prä	33	18,18	12,12	24,24	18,18	9,09	9,09	--	--	--
3-Post	33	15,15	9,09	24,24	18,18	3,03	6,06	--	9,09	--
4-Prä	42	21,43	16,67	30,95	19,05	2,38	--	--	11,90	--
4-Post	42	19,05	14,29	26,19	21,43	4,76	--	--	9,52	--
5	86	23,26	11,63	34,88	23,26	6,98	3,49	4,65	11,63	--
6	42	23,81	14,29	35,71	23,81	4,76	--	9,52	14,29	--
7	115	33,91	12,17	44,35	30,43	13,04	3,48	13,91	18,26	--
8-Prä	30	23,33	--	--	26,67	6,67	10,00	--	--	--
8-Post	30	20,00	--	10,00	20,00	6,67	20,00	--	--	--

In den Gruppen 2 bis 7 wird am häufigsten der Änderungspunkt „wichtige Fächer länger“ gewünscht. Die Werte reichen hier von 24,24% in den Gruppen 3-Prä und 3-Post bis 44,35% in Gruppe 7. Oft findet sich bei dieser Antwort noch der Vermerk auf den Fachbereich „Innere Medizin“, das als „wichtiges Fach“ länger doziert werden soll.

Für den von den Untersuchungsteilnehmern angegebenen Änderungspunkt „kein Hammerexamen“ steigen die Werte mit dem Ausbildungsstand an. So geben in Gruppe 1 nur 5,40% dieses Stichwort an, während 33,91% der Studierenden in Gruppe 7 diesen Standpunkt vertreten.

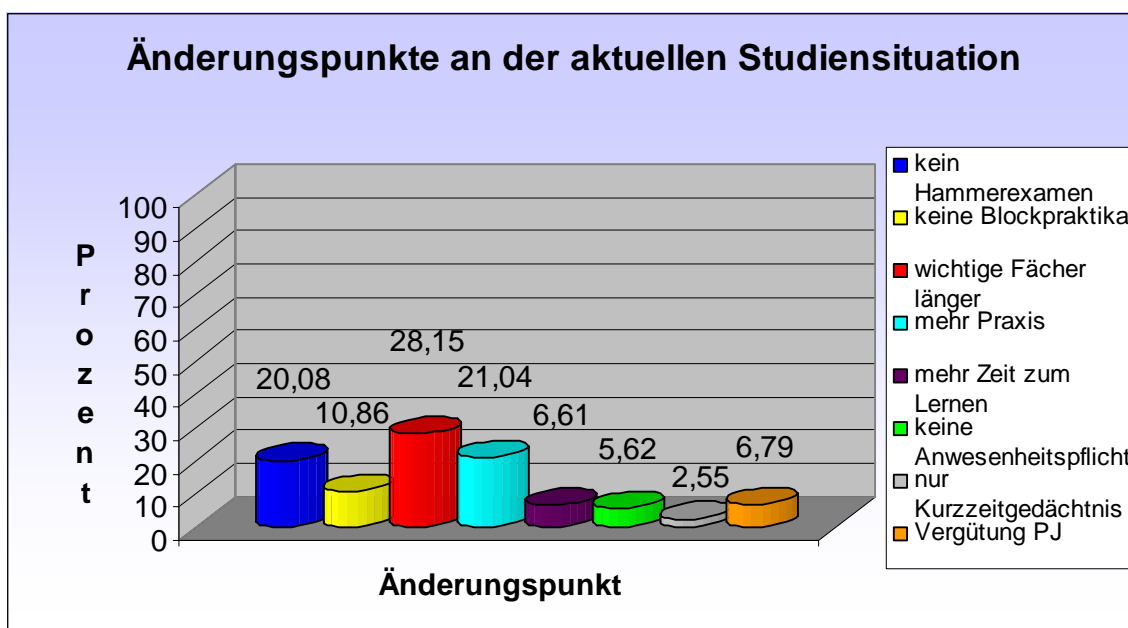
Hohe Prozentsätze in allen Gruppen finden sich auch für das Stichwort „mehr Praxis“. Hier kommt ebenfalls ein Anstieg der Werte von Gruppe 1 (8,11%) zu Gruppe 7 (30,43%) zum Vorschein.

Durchschnittlich hohe Prozentsätze erhält auch die Aussage „keine Blockpraktika“ (9,09% bis 16,67%).

Der Wunsch nach einer „Vergütung des Praktischen Jahres“ beginnt in Gruppe 3-Post (9,09%) und wird bis hin zu Gruppe 7 (18,26%) immer öfter angegeben.

In den Gruppen 5, 6 und 7 wird der Änderungspunkt geäußert, dass das Lehrsystem „nur auf das Kurzzeitgedächtnis“ abzielt (4,65%, 9,52% bzw. 13,91%)

Diagramm 11: Mittelwerte aller Gruppen (Gruppen 1-8) zu Änderungspunkten an der aktuellen Studiensituation



3.3) Ergebnisse der Fragebogenauswertung für Teil II (MC-Fragen zum Wissensstand im Fachbereich Orthopädie)

Es wurde für jede Gruppe der Mittelwert der richtig beantworteten Fragen absolut und in Prozentwerten sowie die Standardabweichung berechnet.

Da zu Beginn der Untersuchung in den einzelnen Gruppen ausdrücklich darauf hingewiesen wurde, auch bei völligem Unwissen auf Grund der 20%igen Ratewahrscheinlichkeit der richtigen Antwort alle MC-Fragen zu bearbeiten, konnten alle Fragebögen in diesem Teil vollständig ausgefüllt zurückerhalten und somit in die Auswertung einbezogen werden.

**Tabelle 14: Ergebnisse Fragenteil II (MC-Fragen)
Richtigantworten und Standardabweichung (absolut und in Prozent)**

Gruppe	n	Mittelwert		Standardabweichung	
		Punktwert (von 12)	in Prozent (12=100%)	Punktwert	in Prozent
1	37	3,68	30,63	5,46	14,77
2	121	3,96	32,99	24,46	20,21
3-Prä	33	7,21	60,10	7,51	24,61
3-Post	33	9,24	77,02	4,88	14,81
4-Prä	42	7,45	62,10	12,49	29,74
4-Post	42	9,24	76,98	8,84	21,06
5	86	8,48	70,64	16,98	19,75
6	42	7,64	63,69	8,83	21,03
7	115	7,30	60,87	24,82	21,58
8-Prä	30	6,40	53,33	5,37	17,92
8-Post	30	8,77	73,06	5,43	18,12

In Gruppe 1 und Gruppe 2 treten die niedrigsten prozentualen Werte für die Richtigantworten der MC-Fragen auf (30,63% bzw. 32,99%).

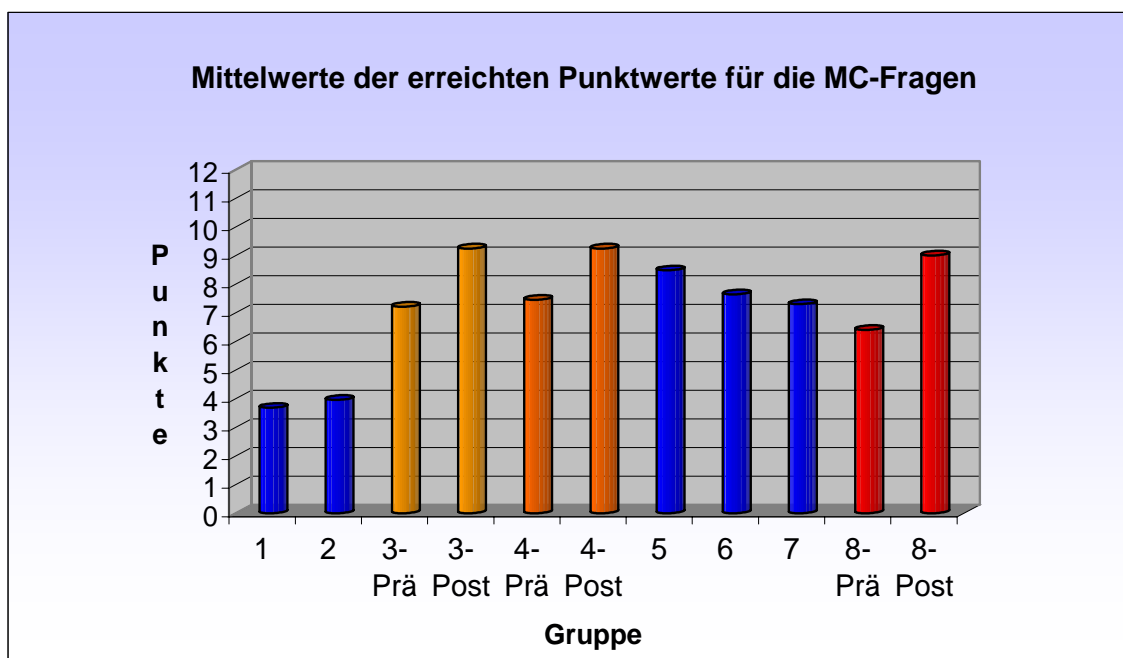
Gruppe 3 erreicht zu Beginn des Orthopädie-Praktikums 60,10% (3-Prä) an Richtigantworten und weist am Ende des Orthopädie-Praktikums einen Zuwachs auf 77,02% (3-Post) auf. Gruppe 4 erfährt ebenfalls im Laufe des Orthopädie-Praktikums einen prozentualen Zuwachs an Richtigantworten (62,10% in Gruppe 4-Prä und 76,98% in Gruppe 4-Post).

In den Gruppen 5 (70,64%), 6 (63,69%) und 7 (60,87%) geht der prozentuale Wert an Richtigantworten mit steigendem Ausbildungsstand zurück.

In Gruppe 8 (Universität Heidelberg) erfolgt im Vergleich der Gruppe vor dem Orthopädie-Praktikum (8-Prä) mit nach dem Orthopädie-Praktikum (8-Post) ein prozentualer Zuwachs an Richtigantworten. So erreicht Gruppe 8-Prä einen prozentualen Wert von 53,33% an Richtigantworten, wohingegen Gruppe 8-Post einen Wert von 73,06% aufweisen kann.

Die zugehörigen Standardabweichungen (absolut und in Prozent) sind Tab. 14 (siehe Seite 84) zu entnehmen.

Diagramm 12: Mittelwerte der erreichten Punktwerte der einzelnen Gruppen für die MC-Fragen



Die statistische Analyse erfolgte unter Verwendung des Student'schen t-Tests für paare und unpaare Stichproben. Es werden immer zwei Gruppen miteinander verglichen. Das Signifikanzniveau wird auf $p < 0,05$ festgelegt. Dies bedeutet, dass bei einem Wert von $p < 0,05$ ein statistisch signifikanter Unterschied zwischen den beiden Gruppen vorliegt, bei $p > 0,05$ ist der Unterschied statistisch nicht signifikant. Beim Vergleich der Gruppen 3-Prä mit 3-Post, 4-Prä mit 4-Post und 8-Prä mit 8-Post handelt es sich um paare Stichproben, bei den anderen Gruppen um unpaare Stichproben.

Tabelle 15: Statistische Analyse: p-Werte für die MC-Fragen

Gruppen	FR 11	FR 12	FR 13	FR 14	FR 15	FR 16	FR 17	FR 18	FR 19	FR 20	FR 21	FR 22
	p – Werte											
1 2	0,66	0,65	0,37	0,89	0,78	0,25	0,95	0,35	0,12	0,51	0,06	0,41
2 3-Prä	0,49	0,00	0,08	0,04	0,00	0,00	0,59	0,01	0,85	0,01	0,00	0,00
3-Prä 3-Post	0,00	0,04	0,00	0,01	0,02	0,08	0,00	0,08	0,04	0,02	1,00	0,32
4-Prä 4-Post	0,00	0,16	0,00	0,16	0,01	0,16	0,00	1,00	0,04	0,00	0,16	0,32
4-Post 5	0,05	0,00	0,63	0,85	0,08	0,21	0,49	0,80	0,75	0,44	0,56	0,20
5 6	0,21	0,04	0,44	0,85	0,23	0,97	0,27	0,72	0,70	0,46	0,52	0,28
6 7	0,47	0,65	0,50	0,82	0,36	0,62	0,81	0,79	0,63	0,52	0,90	0,92
8-Prä 8-Post	0,02	0,01	0,00	1,00	0,00	0,00	0,00	1,00	0,08	0,04	0,01	0,02

Bei der statistischen Analyse Im Vergleich von Gruppe 1 mit Gruppe 2 (unpaare Stichproben) ergeben sich ausschließlich p-Werte $> 0,05$ (0,06 bis 0,95).

Beim Vergleich von Gruppe 2 (2.klinisches Semester) mit Gruppe 3-Prä (3.klinisches Semester nach Absolvierung der einwöchigen Einführung in die Orthopädie und vor Absolvierung des Orthopädie-Praktikums) handelt es sich um unpaare Stichproben. Es treten für die Fragen 11, 13, 17 und 19 p-Werte $> 0,05$ auf (0,08 bis 0,85). Bei den Fragen 12, 14, 15, 16, 18, 20, 21 und 22 befinden sich die p-Werte unterhalb des festgelegten Signifikanzniveaus von 0,05 (0,00 bis 0,04).

Die Analyse von Gruppe 3-Prä mit Gruppe 3-Post beinhaltet paare Stichproben. Dabei zeigen sich für die Fragen 11, 12, 13, 14, 15, 17, 19 und 20 p-Werte unterhalb des Signifikanzniveaus von 0,05 (0,00 bis 0,04). Bei den Fragen 16, 18, 21 und 22 erscheinen p-Werte $> 0,05$ (0,08 bis 1,00).

Gruppe 4-Prä und Gruppe 4-Post sind paare Strichproben. Beim Vergleich dieser beiden Gruppen werden für die Fragen 12, 14, 16, 18, 21 und 22 Werte oberhalb des Signifikanzniveaus erzielt (0,16 bis 1,00). Für die Fragen 11, 13, 15, 17, 19 und 20 finden sich p-Werte $< 0,05$ (0,00 bis 0,04).

Bei der Analyse von Gruppe 4-Post mit Gruppe 5 (unpaare Stichproben) beträgt der p-Wert nur bei Frage 12 mit 0,00 einen Wert unterhalb des Signifikanzniveaus. Bei den restlichen Fragen ist der p-Wert im Bereich von 0,05 bis 0,85.

Beim Vergleich der Gruppe 5 mit Gruppe 6 (unpaare Stichproben) findet sich nur bei Frage 12 ein p-Wert unterhalb des Signifikanzniveaus (0,04), beim Vergleich der Gruppe 6 mit Gruppe 7 ausschließlich p-Werte $> 0,05$ (0,36 bis 0,92).

Die Gruppen 8-Prä und 8-Post der Universität Heidelberg repräsentieren paare Stichproben. Es ergeben sich im Vergleich der beiden Gruppen für die Fragen 14, 18 und 19 p-Werte $> 0,05$ (0,08 bis 1,0), für die restlichen Fragen fallen die p-Werte in den Bereich 0,00 bis 0,04.

3.4) Ergebnisse der Fragebogenauswertung für Teil II (Offene Fragen zum Wissensstand im Fachbereich Orthopädie)

Für die Auswertung der Offenen Fragen wurden für jede Gruppe die Mittelwerte der in den einzelnen Fragen erreichten Punktwerte ermittelt. Infolge dieser Werte wurde die im Mittel erreichte Gesamtpunktzahl absolut und in Prozent sowie die Standardabweichung berechnet.

In den Gruppen 1 und 2 wurde in mehreren Fragebögen (18,92% bzw. 18,18%) keine der drei Offenen Fragen beantwortet. Um einen Vergleich der Ergebnisse dieser Gruppen mit den Ergebnissen der anderen Gruppen überhaupt zu ermöglichen, wurden Fragebögen mit null Punkten für die Offenen Fragen nicht in die Wertung für diesen Teil des Fragebogens aufgenommen. In Gruppe 1 gehen somit nur 30 von 37 Fragebögen in die Wertung ein, in Gruppe 2 nur 99 von 121.

**Tabelle 16: Ergebnisse Fragenteil II (Offene Fragen)
Erreichte Punktzahl und Standardabweichung
(absolut und in Prozent)**

Gruppe	n	Mittelwert		Standardabweichung	
		Punktwert (von 12)	in Prozent (12=100%)	Punktwert	in Prozent
1	30	4,73	39,44	1,71	42,76
2	99	4,93	41,16	1,63	40,84
3-Prä	33	3,94	32,83	1,24	30,96
3-Post	33	6,72	56,06	1,41	35,23
4-Prä	42	5,69	47,42	1,43	35,78
4-Post	42	7,12	59,33	1,40	35,22
5	86	6,02	50,10	1,24	31,22
6	42	5,81	48,42	1,18	29,72
7	115	5,61	46,81	1,10	27,63
8-Prä	30	4,73	39,44	1,31	32,95
8-Post	30	6,86	57,22	1,33	33,34

Gruppe 1 und 2 erreichen einen prozentualen Wert von 39,44% bzw. 41,16% an Richtigantworten für die Offenen Fragen und übertreffen somit als einzige Gruppen die Ergebnisse für die MC-Fragen (30,63% in Gruppe 1 bzw. 32,99% in Gruppe 2).

In Gruppe 3-Prä findet sich der niedrigste Wert aller Gruppen mit 32,83%. Im Laufe des Orthopädie-Praktikums erfolgt ein prozentualer Zuwachs, so dass Gruppe 3-Post einen Wert von 56,06% an Richtigantworten aufweisen kann.

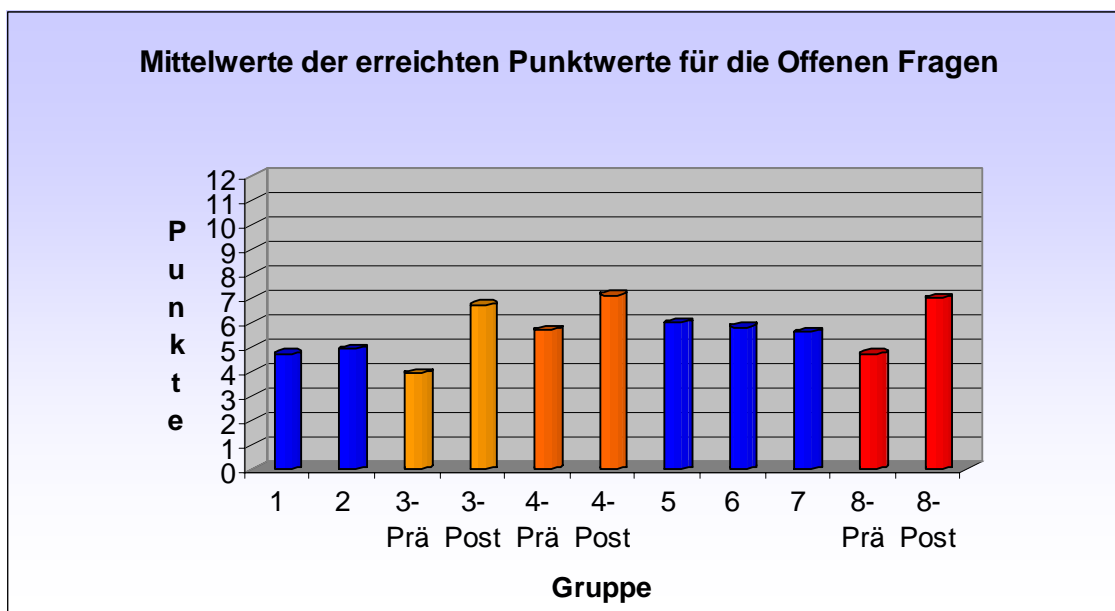
In Gruppe 4 erfolgt ebenfalls ein Zuwachs während des Orthopädie-Praktikums. Hier erreicht Gruppe 4-Prä einen Wert von 47,42%, wohingegen Gruppe 4-Post 59,33% erreicht.

Mit steigendem Ausbildungsstand erfolgt ein Absinken der prozentualen Werte. Gruppe 5 verzeichnet einen Wert von 50,10%, Gruppe 6 von 48,42% und Gruppe 7 von 46,81%.

Gruppe 8 (Universität Heidelberg) kommt im Prä-Test (Gruppe 8-Prä) auf 39,44% und im Post-Test (Gruppe 8-Post) auf 57,22% an Richtigantworten.

Die zugehörigen Standardabweichungen (absolut und in Prozent) sind Tab. 16 (siehe Seite 88) zu entnehmen.

Diagramm 13: Mittelwerte der erreichten Punktwerte der einzelnen Gruppen für die Offenen Fragen



Die statistische Analyse erfolgte unter Verwendung des Student'schen t-Tests für paare und unpaare Stichproben. Es werden immer zwei Gruppen miteinander verglichen. Das Signifikanzniveau wird auf $p < 0,05$ festgelegt. Dies bedeutet, dass bei einem Wert von $p < 0,05$ ein statistisch signifikanter Unterschied zwischen den beiden Gruppen vorliegt, bei $p > 0,05$ ist der Unterschied statistisch nicht signifikant. Beim Vergleich der Gruppen 3-Prä mit 3-Post, 4-Prä mit 4-Post und 8-Prä mit 8-Post handelt es sich um paare Stichproben, bei den anderen Gruppen um unpaare Stichproben.

Tabelle 17: Statistische Analyse: p-Werte für die Offenen Fragen

Gruppe	paar / unpaar	Frage 23	Frage 24	Frage 25
		p-Werte		
1 2	unpaar	0,75	0,36	0,54
2 3-Prä	unpaar	0,00	0,37	0,92
3-Prä 3-Post	paar	0,00	0,00	0,00
4-Prä 4-Post	paar	0,00	0,00	0,00
4-Post 5	unpaar	0,01	0,01	0,35
5 6	unpaar	0,47	0,97	0,85
6 7	unpaar	0,40	0,95	0,99
8-Prä 8-Post	Paar	0,00	0,00	0,00

Beim Vergleich von Gruppe 1 mit Gruppe 2 (unpaare Stichproben) finden für die drei Offenen Fragen jeweils p-Werte, die über dem Signifikanzniveau von 0,05 liegen (0,36 bis 0,75).

Die Analyse von Gruppe 2 (2.klinisches Semester) mit Gruppe 3-Prä (3.klinisches Semester nach Absolvierung der einwöchigen Einführung in die Orthopädie und vor Absolvierung des Orthopädie-Praktikums) beinhaltet unpaare Stichproben. Die Analyse weist für Frage 23 mit einem p-Wert von 0,00 einen statistisch signifikanten Unterschied auf. Für die Fragen 24 und 25 beträgt der p-Wert 0,37 bzw. 0,92 und ist damit höher als das Signifikanzniveau.

Beim Vergleich der Gruppe 3-Prä mit Gruppe 3-Post handelt es sich um paare Stichproben. Die p-Werte für die Fragen 23, 24 und 25 sind im Bereich von $1,01 \cdot 10^{-5}$ bis $1,34 \cdot 10^{-10}$ angesiedelt, so dass sich für alle drei Fragen gerundet p-Werte von jeweils 0,00 ergeben.

Bei der Analyse von Gruppe 4-Prä mit 4-Post wird ebenfalls der Student'sche t-Test für paare Stichproben durchgeführt. Hier befinden sich die p-Werte im Bereich von $6,05 \cdot 10^{-5}$ bis $4,16 \cdot 10^{-8}$, so dass für die Fragen 23, 24 und 25 p-Werte von jeweils 0,00 angegeben werden.

Beim Vergleich der Gruppe 4-Post mit der Gruppe 5 erfolgt die Analyse für unpaare Stichproben. Bei Frage 23 und Frage 24 tritt jeweils ein p-Wert von 0,01 auf, der somit unterhalb dem Signifikanzniveau von 0,05 rangiert. Frage 25 weist mit 0,35 einen Wert größer als $p > 0,05$ auf.

Vergleicht man Gruppe 5 mit Gruppe 6, so wird die statistische Analyse für unpaare Stichproben durchgeführt. Die p-Werte befinden sich alle oberhalb des Signifikanzniveaus von 0,05 und erreichen Werte von 0,47 (Frage 23), 0,85 (Frage 25) und 0,97 (Frage 24).

Bei der Analyse von Gruppe 6 mit Gruppe 7 handelt es sich um unpaare Stichproben. Für alle drei Fragen erscheinen p-Werte $> 0,05$, die von 0,40 (Frage 23) bis 0,99 (Frage 25) reichen.

Paare Stichproben beinhaltet der Vergleich von Gruppe 8-Prä mit Gruppe 8-Post (Universität Heidelberg). Die p-Werte liegen im Bereich von $4,08 \cdot 10^{-5}$ bis $1,68 \cdot 10^{-8}$, so dass sich für alle drei Fragen gerundet p-Werte von 0,00 ergeben.

Im Folgenden werden die Ergebnisse für die drei Offenen Fragen im Einzelnen dargestellt:

Frage 23: Muskuläre Strukturen der Rotatorenmanschette

Frage 24: 4 Ursachen, die zu einem Beckenschiefstand führen können

Frage 25: 4 Frühzeichen, die bei einem Morbus Perthes gelten

**Tabelle 18: Ergebnisse für die Fragen 23, 24, 25 (Teil II, Offene Fragen)
Erreichte Punktzahl (absolut und in Prozent)**

Gruppe	n	Frage 23		Frage 24		Frage 25	
		Punktwert (von 4)	in Prozent	Punktwert (von 4)	in Prozent	Punktwert (von 4)	in Prozent
1	30	3,47	86,67	1,13	28,33	0,13	3,33
2	99	3,41	85,35	1,33	33,33	0,19	4,80
3-Prä	33	2,64	65,91	1,12	28,03	0,18	4,55
3-Post	33	3,64	90,91	2,27	56,82	0,82	20,45
4-Prä	42	3,14	78,57	2,21	55,36	0,33	8,33
4-Post	42	3,64	91,07	2,62	65,48	0,86	21,43
5	86	3,19	79,65	2,13	53,20	0,70	17,44
6	42	3,02	75,60	2,12	52,98	0,67	16,67
7	115	2,84	71,09	2,10	52,61	0,67	16,74
8-Prä	30	2,93	73,33	1,50	37,50	0,30	7,50
8-Post	30	3,60	90,00	2,33	58,33	0,93	23,33

Frage 23 weist die höchsten prozentualen Werte auf. Dabei reichen die Werte von 65,91% in Gruppe 3-Prä bis 91,91% in Gruppe 4-Post.

Gruppe 3-Prä weist den niedrigsten prozentualen Wert auf (65,91%) und verbessert sich im Laufe des Orthopädie-Praktikums auf 90,91% (Gruppe 3-Post). Gruppe 4 erfährt

ebenfalls einen prozentualen Zuwachs während des Orthopädie-Praktikums und verbessert sich von 78,57% (Gruppe 4-Prä) auf 91,09% (Gruppe 4-Post).

In den Gruppen 1 und 2 finden sich hohe Werte (86,67% bzw. 83,35%), während in den Gruppen 5 (79,65%), 6 (75,60%) und 7 (71,09%) ein prozentualer Abfall mit steigendem Ausbildungsstand erfolgt.

Gruppe 8 (Universität Heidelberg) kommt zu Beginn des Orthopädie-Praktikums auf 73,33% (Gruppe 8-Prä), am Ende des Orthopädie-Praktikums auf 90,00% (Gruppe 8-Post).

Bei Frage 24 erreicht der höchste Wert (65,48% in Gruppe 4-Post) nicht den niedrigsten Wert aus Frage 23 (65,91% in Gruppe 3-Prä). Bei dieser Frage kommen Gruppe 1 auf 28,33% und Gruppe 2 auf 33,33%.

In Gruppe 3-Prä findet sich wie bei Frage 23 der niedrigste Wert (28,03%). Im Laufe des Orthopädie-Praktikums erfolgt ein prozentualer Zuwachs, so dass Gruppe 3-Post einen Wert von 56,82% aufweisen kann.

Gruppe 4 verbessert sich von 55,36% (Gruppe 4-Prä) auf 65,48% (Gruppe 4-Post) und weist damit den Höchstwert auf.

Die Gruppen 5, 6 und 7 fallen im Vergleich zu den Postgruppen (3-Post und 4-Post) prozentual ab und erreichen Werte von 53,20%, 52,98% und 52,61%.

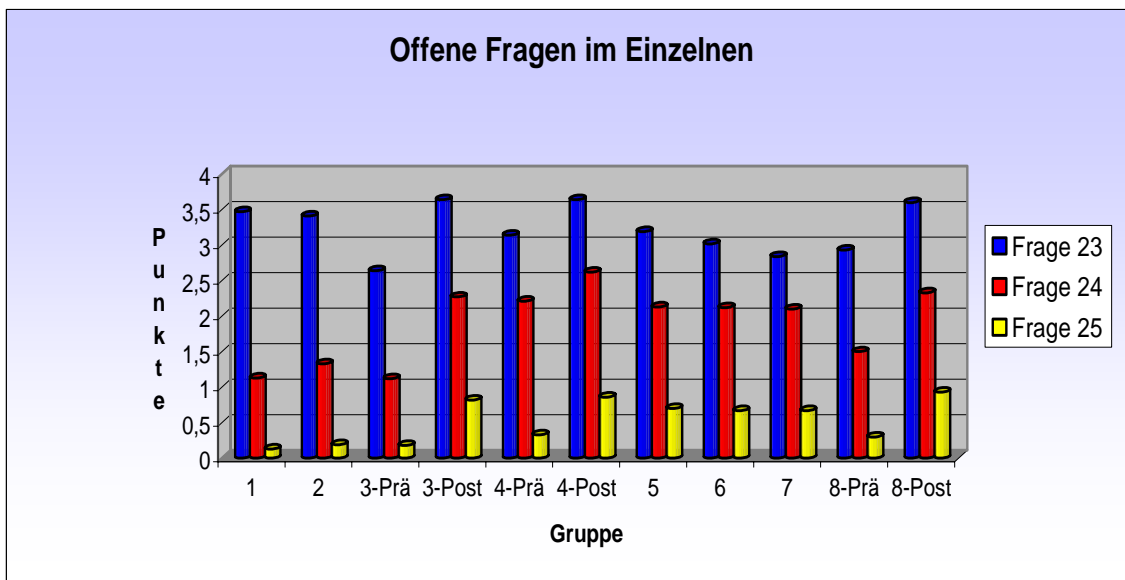
Gruppe 8 (Universität Heidelberg) weist zu Beginn des Orthopädie-Praktikums 37,50% (Gruppe 8-Prä) an Richtigantworten auf. Nach Beendigung des Orthopädie-Praktikums lässt sich ein prozentualer Zuwachs auf 58,33% (Gruppe 8-Post) feststellen.

Hinsichtlich der Offenen Fragen schneiden alle Gruppen bei Frage 25 am schlechtesten ab. So erreichen hier die Gruppen 1 und 2 Werte von 3,33% bzw. 4,80%.

Gruppe 3 verbessert sich von 4,55% (Gruppe 3-Prä) auf 20,45% (Gruppe 3-Post), Gruppe 4 von 8,33% (Gruppe 4-Prä) auf 21,43% (Gruppe 4-Post).

In den Gruppen 5, 6 und 7 finden sich Werte von 17,44%, 16,67% bzw. 16,74%. Gruppe 8 (Universität Heidelberg) verbessert seinen Ausgangswert von 7,50% (Gruppe 8-Prä) auf einen prozentualen Wert von 23,33% (Gruppe 8-Post), was den Höchstwert aller Gruppen für Frage 25 bedeutet.

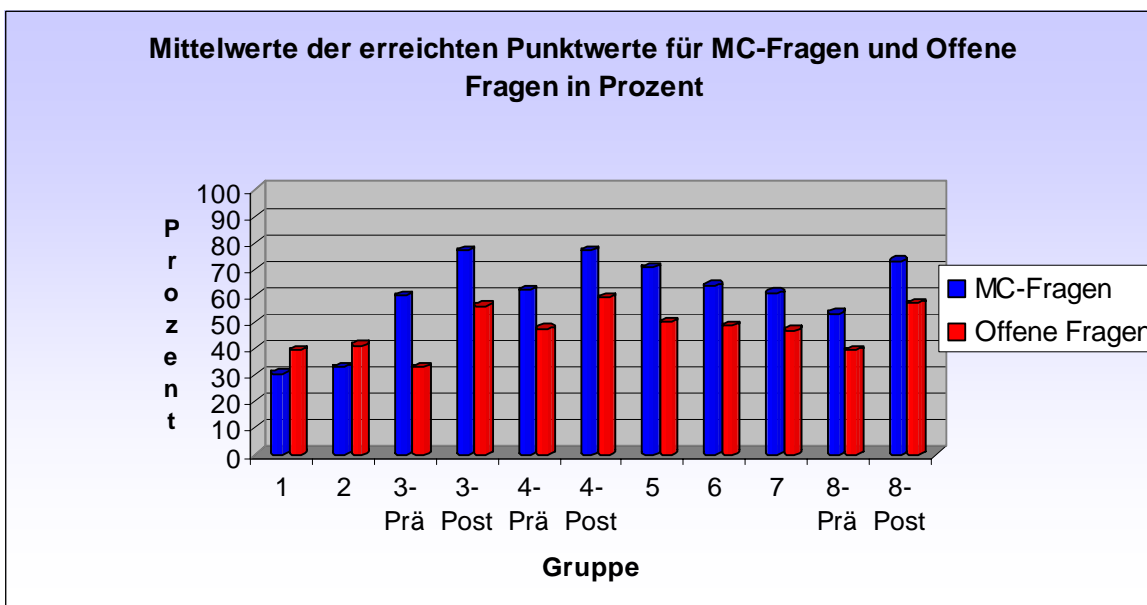
Diagramm 14: Mittelwerte der erreichten Punkte der einzelnen Gruppen für die Offenen Fragen im Einzelnen



An dieser Stelle sollen die Ergebnisse für die MC-Fragen und die Offenen Fragen in einer Grafik zusammengefasst dargestellt werden:

Nur die Gruppen 1 und 2 erzielen höhere prozentuale Werte für die Offenen Fragen (39,44% bzw. 41,16%) als für die Multiple-Choice-Fragen (30,63% bzw. 32,99%).

Diagramm 15: Mittelwerte der erreichten Punktzahlen für MC-Fragen und Offene Fragen in Prozent



3.5) Ergebnisse der Fragebogenauswertung für Teil III

(Einschätzung des Schwierigkeitsgrades der gestellten Fragen und Interesse am Fachbereich Orthopädie)

In Teil III des Fragebogens erfolgte eine Ermittlung der relativen Häufigkeiten der Angaben zur Einschätzung des Schwierigkeitsgrades der gestellten Fragen (Fragen 26, 27), zum persönlichen Interesse am Fachbereich Orthopädie (Fragen 28, 29) und zur Befürwortung oder Ablehnung einer solchen Untersuchung (Frage 30).

Im Folgenden werden die Ergebnisse für die Fragen im Einzelnen dargestellt:

Frage 26) Schwierigkeitsgrad der gestellten MC-Fragen

Tabelle 19: Ergebnisse für Frage 26 (Teil III)

Gruppe	n	Prozent				
		sehr leicht	leicht	angemessen	schwierig	sehr schwierig
1	37	0,00	0,00	16,22	45,95	37,84
2	121	0,00	3,31	8,26	62,81	25,62
3-Prä	33	0,00	3,03	12,12	66,67	18,18
3-Post	33	3,03	9,09	78,79	9,09	0,00
4-Prä	42	0,00	0,00	71,43	28,57	0,00
4-Post	42	0,00	7,14	76,19	16,67	0,00
5	86	0,00	4,65	50,00	43,02	2,32
6	42	0,00	4,76	38,10	52,38	4,76
7	115	0,00	0,00	33,04	63,48	3,48
8-Prä	30	0,00	0,00	36,67	50,00	13,33
8-Post	30	0,00	6,67	53,33	33,33	6,67

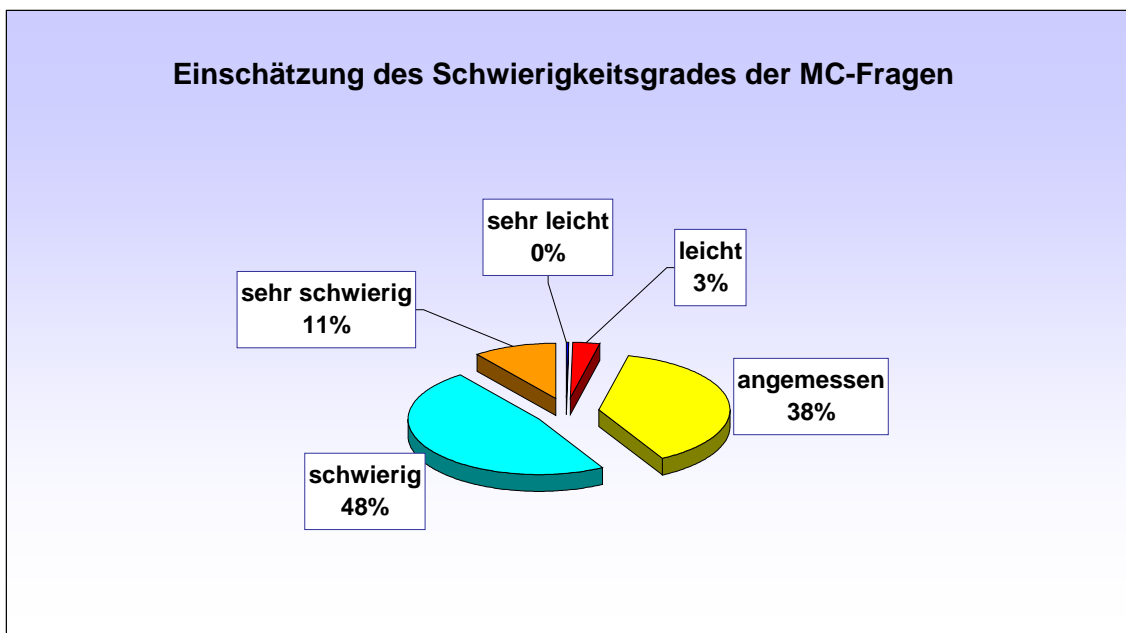
Die Kategorie „sehr leicht“ ist nur in Gruppe 3-Post (3,03%) prozentual vertreten, alle anderen Gruppen weisen für diese Antwort 0,00% auf. Für die Antwort „leicht“ erscheinen nur einstellige Werte, die von 0,00% bis 9,09% (Gruppe 3-Post) reichen.

Die Kategorie „angemessen“ stellt in den Gruppen 3-Post (78,79%), 4-Prä (71,43%), 4-Post (76,19%), in Gruppe 5 (50,00%) und in Gruppe 8-Post (53,33%) den Höchstwert dar. In den Gruppen 1 (16,22%) und 2 (8,26%) finden sich für diese Antwort die niedrigsten Werte, in den Gruppen 6 (38,10%) und 7 (33,04%) fallen die Werte mit dem steigenden Ausbildungsstand gegenüber den Werten aus den Post-Gruppen wieder ab.

Für die Gruppen 1, 2, 3-Prä, 6, 7 und 8-Prä verkörpert die Kategorie „schwierig“ mit Werten von 45,95% (Gruppe 1) bis 66,67% (Gruppe 3-Prä) den Höchstwert.

Die Kategorie „sehr schwierig“ erreicht in Gruppe 1 den Höchstwert (37,84%), fällt in den Gruppen 3-Post, 4-Prä und 4-Post auf 0,00% ab und stabilisiert sich in den Gruppen 5 (2,32%), 6 (4,76%) und 7 (3,48%).

Diagramm 16: Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Schwierigkeitsgrades der gestellten MC-Fragen



Frage 27) Schwierigkeitsgrad der gestellten Offenen Fragen

Tabelle 20: Ergebnisse für Frage 27 (Teil III)

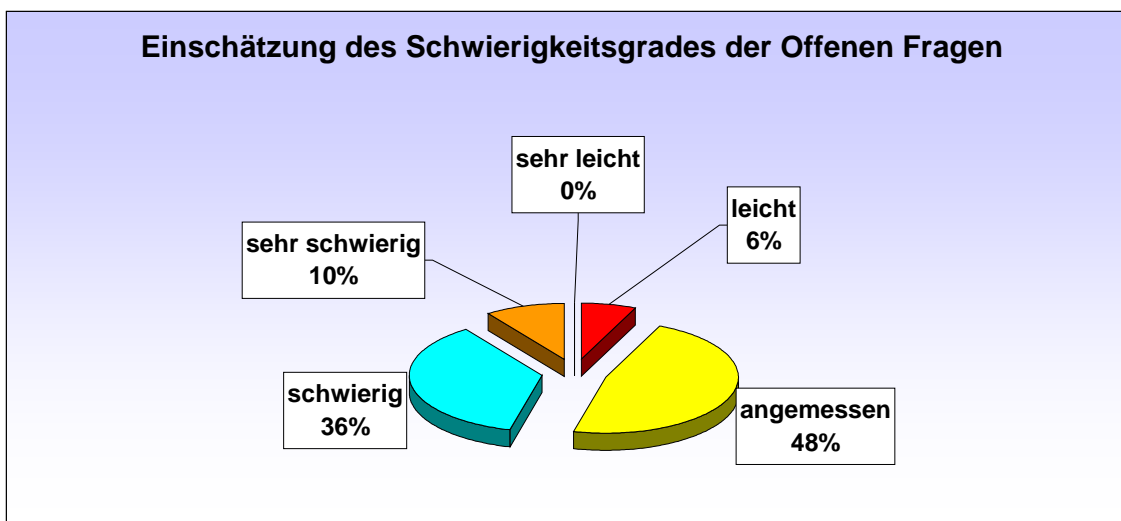
Gruppe	n	Prozent				
		sehr leicht	leicht	angemessen	schwierig	sehr schwierig
1	37	0,00	2,70	27,03	45,95	24,32
2	121	0,00	9,09	21,49	55,37	14,05
3-Prä	33	0,00	12,12	30,30	48,48	9,09
3-Post	33	0,00	12,12	75,76	12,12	0,00
4-Prä	42	0,00	2,38	57,14	38,10	2,38
4-Post	42	0,00	7,14	76,19	16,67	0,00
5	86	0,00	6,98	63,95	22,09	6,98
6	42	0,00	2,38	54,76	33,33	9,52
7	115	0,00	2,61	47,83	36,52	13,04
8-Prä	30	0,00	0,00	40,00	43,33	16,67
8-Post	30	0,00	3,33	63,33	33,33	0,00

Die Antwort „sehr leicht“ erhält in allen Gruppen 0,00%. In den Gruppen 3-Prä und 3-Post wird die Kategorie „leicht“ jeweils mit 12,12% versehen, in allen anderen Gruppen ergeben sich für diese Antwort nur einstellige Werte, von 0,00% in Gruppe 8-Prä bis 9,09% in Gruppe 2.

In den Gruppen 1 (45,95%), 2 (55,37%) und 3-Prä (48,48%) liegt der Höchstwert noch in der Kategorie „schwierig“, verschiebt sich in den Gruppen 3-Post (75,76%), 4-Prä (57,14%), 4-Post (76,19%), 5 (63,95%), 6 (54,76%) und 7 (47,83%) dann aber hin zu Kategorie „angemessen“.

Die Werte für die Kategorie „sehr schwierig“ nehmen von Gruppe 1 (24,32%) über Gruppe 2 (14,05%) und über Gruppe 3-Prä (9,09%) bis zu 0,00% in den Post-Gruppen ab (Gruppe 3-Post und 4-Post), bevor die Werte wieder über Gruppe 5 (6,98%) und Gruppe 6 (9,52%) bis 13,04% (Gruppe 7) ansteigen.

Diagramm 17: Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Schwierigkeitsgrades der gestellten Offenen Fragen



Frage 28) Interesse am Fachbereich Orthopädie

Tabelle 21: Ergebnisse für Frage 28 (Teil III)

Gruppe	n	Prozent				
		sehr großes Interesse	Interesse	unentschieden	wenig Interesse	gar kein Interesse
1	37	8,11	35,14	35,14	18,92	2,70
2	121	13,22	38,84	14,88	25,62	7,44
3-Prä	33	15,15	45,45	15,15	15,15	9,09
3-Post	33	15,15	36,36	24,24	21,21	3,03
4-Prä	42	9,52	52,38	11,90	26,19	0,00
4-Post	42	26,19	40,48	16,67	16,67	0,00
5	86	18,60	34,88	17,44	24,42	4,65
6	42	7,14	33,33	14,29	38,10	7,14
7	115	5,22	32,17	18,26	38,26	6,09
8-Prä	30	10,00	33,33	23,33	30,00	3,00
8-Post	30	6,67	30,00	33,33	26,67	3,33

Für die Kategorie „gar kein Interesse“ treten durchweg nur einstellige prozentuale Werte auf, von 0,00% in Gruppe 4-Prä und 4-Post bis 9,09% in Gruppe 3-Prä.

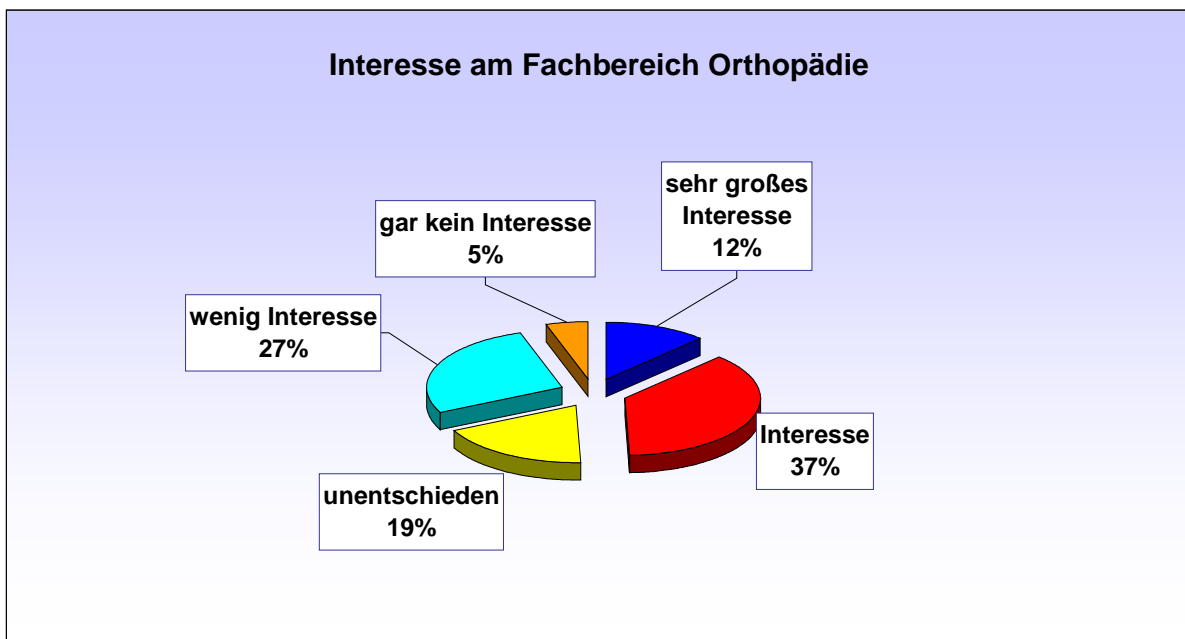
Für die Antwort „sehr großes Interesse“ steigern sich die Werte von Gruppe 1 (8,11%) über Gruppe 2 (13,22%) bis zum Höchstwert in der Gruppe nach dem Orthopädie-Praktikum von 26,19% (Gruppe 4-Post), bevor die Werte über Gruppe 5 (18,60%) und Gruppe 6 (7,14%) bis auf 5,22% in Gruppe 7 wieder absinken.

Für die Gruppen 1-5 befinden sich die Höchstwerte in der Kategorie „Interesse“. Dabei reichen die Werte von 34,88% (Gruppe 5) bis 52,38% (Gruppe 4-Post). Die Gruppen 6 (38,10%) und 7 (38,36%) weisen ihre höchsten prozentualen Werte für die Antwort „wenig Interesse“ auf.

Für Gruppe 8-Prä bedeuten 33,33% in der Kategorie „Interesse“ den Topwert, für Gruppe 8-Post stellen ebenfalls 33,33% in der Kategorie „wenig Interesse“ die Mehrzahl an Antworten dar.

Die Werte für die Kategorie „unentschieden“ richten sich je nach der Verteilung für die anderen Kategorien in jeder Gruppe und erreichen so einen Minimalwert von 11,90% in Gruppe 4-Prä und einen Maximalwert von 35,14% in Gruppe 1.

Diagramm 18: Mittelwerte aller Gruppen (Gruppen 1-8) zum Interesse am Fachbereich Orthopädie



Frage 29) Facharzt für Orthopädie / Unfallchirurgie als Berufswunsch

Tabelle 22: Ergebnisse für Frage 29 (Teil III)

Gruppe	n	Prozent				
		strebe Beruf an	vorstellbar	unentschieden	nicht vorstellbar	sicher nicht
1	37	5,41	29,73	18,92	35,14	10,81
2	121	3,31	36,36	21,49	20,66	18,18
3-Prä	33	6,06	36,36	15,15	21,21	21,21
3-Post	33	9,09	24,24	12,12	33,33	21,21
4-Prä	42	0,00	21,43	28,57	33,33	16,67
4-Post	42	9,52	14,29	42,86	16,67	16,67
5	86	6,98	20,93	31,40	15,12	25,58
6	42	4,76	21,43	19,05	14,29	40,48
7	115	3,48	27,83	13,91	11,30	43,48
8-Prä	30	0,00	20,00	26,66	40,00	13,33
8-Post	30	0,00	30,00	6,67	50,00	13,33

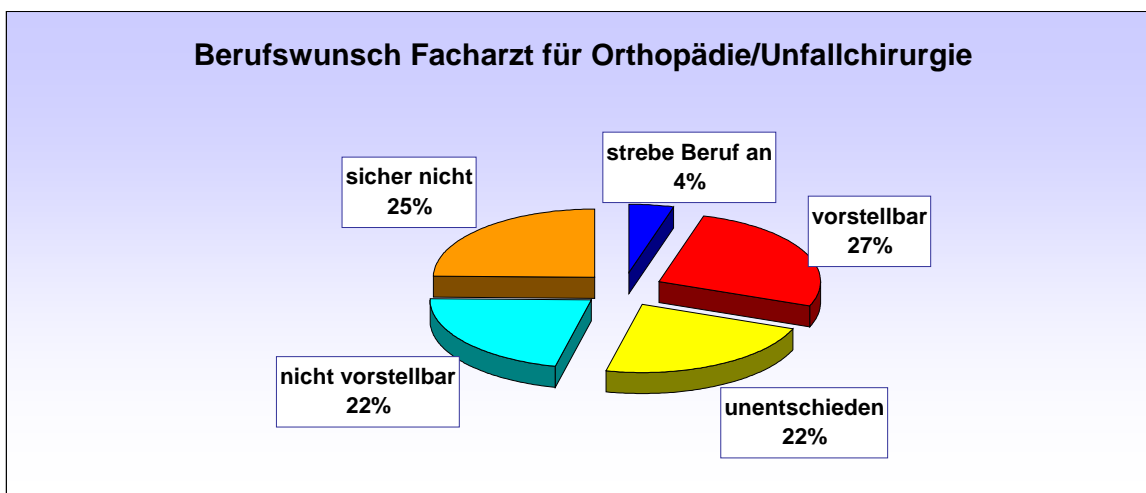
Für die Kategorie „strebe Beruf an“ ergeben sich in allen Gruppen einstellige Werte. Dabei erreicht diese Antwort ihre Höchstwerte von 9,09% (Gruppe 3-Post) und 9,52% (Gruppe 4-Post) in den Gruppen nach dem Orthopädiepraktikum.

Die Antwort „sicher nicht“ findet ihre höchsten prozentualen Werte in den Gruppen mit dem höchsten Ausbildungsstand. Gruppe 6 kommt dabei auf 40,48% und Gruppe 7 auf 43,48%.

In der Kategorie „unentschieden“ findet sich der absolute Höchstwert in der Gruppe 4-Post (42,86%), während sich der niedrigste Wert mit 12,12% in Gruppe 3-Post befindet.

Mit 40,00% (Gruppe 8-Prä) bzw. 50,00% (Gruppe 8-Post) vertritt Gruppe 8 die höchsten prozentualen Werte in der Kategorie „nicht vorstellbar“.

Diagramm 19: Mittelwerte aller Gruppen (Gruppen 1-8) zum Berufswunsch Facharzt für Orthopädie / Unfallchirurgie



Frage 30) Befürwortung einer Evaluation zum Thema Studienordnung und Fachwissenstand

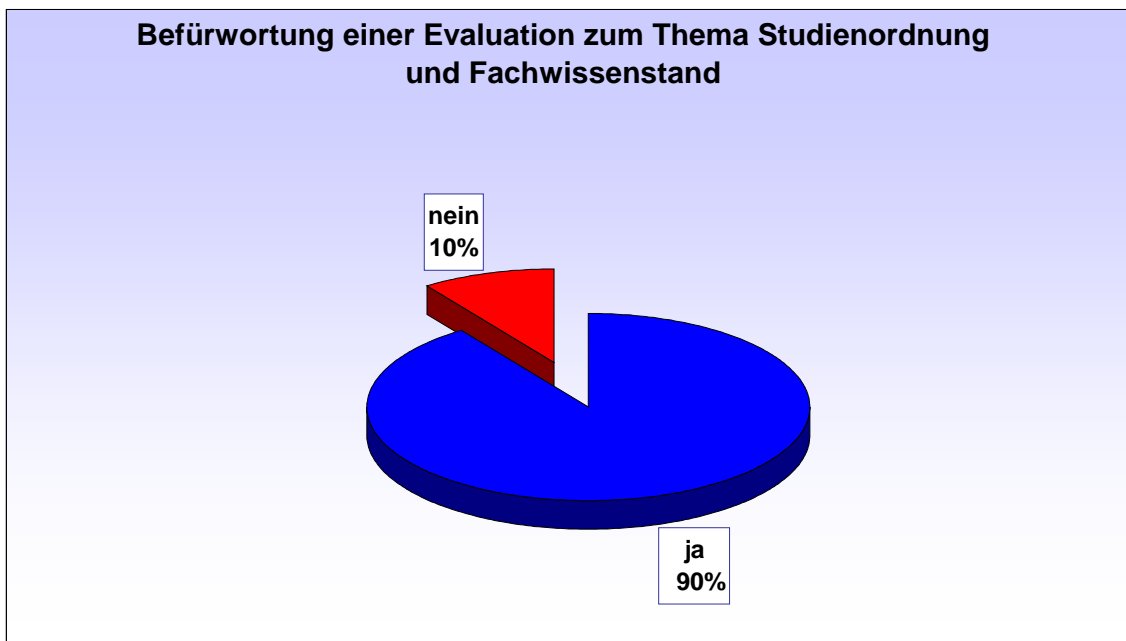
Tabelle 23: Ergebnisse für Frage 30 (Teil III)

Gruppe	n	Prozent	
		ja	nein
1	37	72,97	27,03
2	121	85,95	14,05
3-Prä	33	90,91	9,09
3-Post	33	93,94	6,06
4-Prä	42	90,48	9,52
4-Post	42	88,10	11,90
5	86	90,70	9,30
6	42	90,48	9,52
7	115	96,52	3,48
8-Prä	30	96,67	3,33
8-Post	30	96,67	3,33

Die höchsten Werte für die Kategorie „nein“ befinden sich an der Universität Regensburg in den Semestern mit dem niedrigsten Ausbildungsstand. In Gruppe 1 erreicht die Ablehnung einer Evaluation zum Thema Studienordnung und Fachwissenstand den absoluten Höchstwert von 27,03%. Gruppe 2 weist für diese Kategorie einen prozentualen Wert von 14,05%. Die niedrigsten Werte finden sich in den Gruppen 8-Prä und 8-Post (Universität Heidelberg) mit jeweils 3,33%.

Daraus resultierend finden sich für die Kategorie „ja“ die niedrigsten Werte in den Gruppen 1 (72,87%) und 2 (85,95%). In den Gruppen 3-7 erreichen die Werte 88,10% (Gruppe 4-Post) bis 96,52 (Gruppe 7). Die Spitzenwerte werden durch die Gruppen 8-Prä und 8-Post vertreten (96,67%).

Diagramm 20: Mittelwerte aller Gruppen (Gruppen 1-8) zur Befürwortung einer solchen Evaluation zum Thema Studienordnung und Fachwissenstand



3.6) Vergleich der Ergebnisse von Prä- und Post-Test im Praktikum Orthopädie

In diesem Kapitel sollen die die Ergebnisse der Gruppen 3, 4 und 8 vergleichend dargestellt werden. Diese drei Gruppen können als die Kerngruppen der Untersuchung angesehen werden, da die Erhebung in jeder Gruppe zweimal durchgeführt wurde, d.h. zu Beginn und am Ende des Orthopädie-Praktikums.

Tabelle 24: Vergleich der Ergebnisse von Prä- und Post-Test (Gruppen 3, 4, 8)

Gruppe	n	MC-Fragen		Offene Fragen	
		Punktwert (von 12)	in Prozent (12=100%)	Punktwert (von 12)	in Prozent (12=100%)
3-Prä	33	7,21	60,10	3,94	32,83
3-Post	33	9,24	77,02	6,72	56,06
Ergebnisänderung		2,03	<u>16,92</u>	2,78	<u>23,23</u>
4-Prä	42	7,45	62,10	5,69	47,42
4-Post	42	9,24	76,98	7,12	59,33
Ergebnisänderung		1,79	<u>14,88</u>	1,43	<u>11,91</u>
8-Prä	30	6,40	53,33	4,73	39,44
8-Post	30	8,77	73,06	6,86	57,22
Ergebnisänderung		2,37	<u>19,73</u>	2,13	<u>17,78</u>

Die Ergebnisse lassen sich folgendermaßen zusammenfassen:

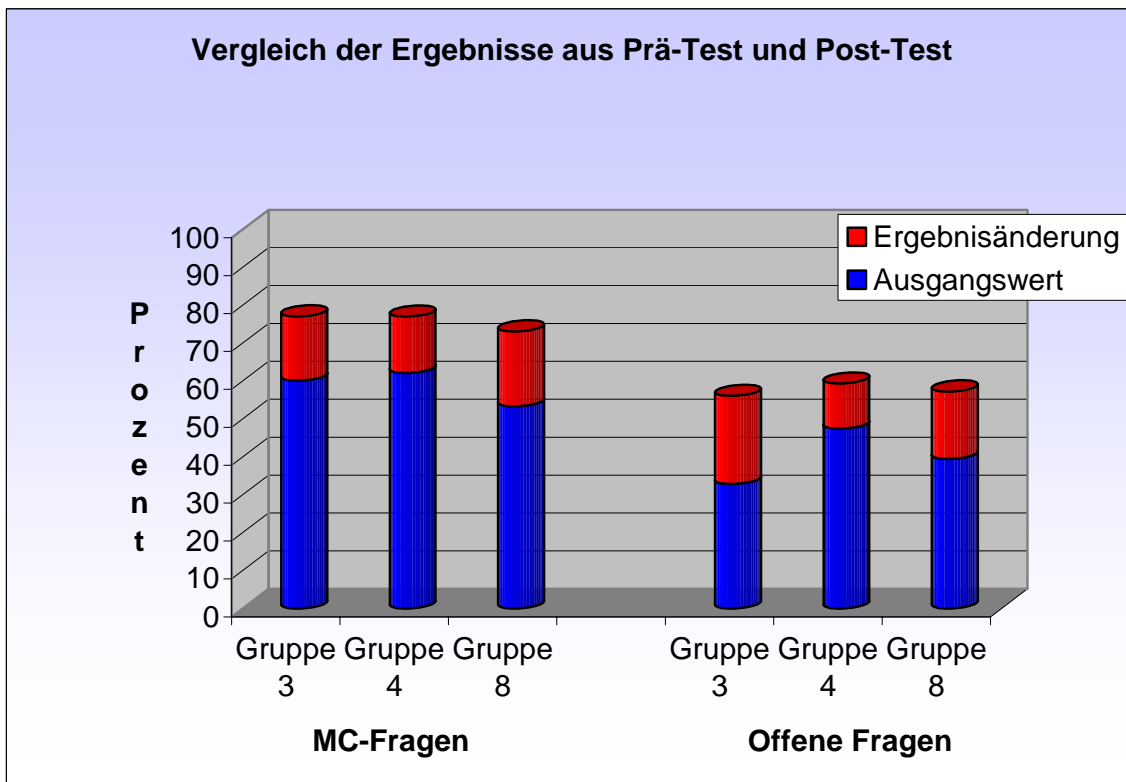
- 1) Die Ergebnisse für die MC-Fragen liegen deutlich über den Werten für die Offenen Fragen
- 2) Gruppe 3 und Gruppe 4 zeigen bei den MC-Fragen ähnliche Ausgangswerte (60,10% bzw. 62,10%) und auch eine ähnliche Ergebnissteigerung (16,92% bzw. 14,88%). Gruppe 8 weist einen niedrigeren Ausgangswert der MC-Fragen auf (53,33%) und erzielt im Verlauf des Orthopädie-Praktikums die höchste Ergebnissteigerung von 19,73%.

3) Gruppe 3 (32,38%) schneidet bei den Offenen Fragen im Prä-Test erheblich schlechter ab als Gruppe 4 (47,42%). Durch das Erreichen eines annähernden Post-Test-Wertes von Gruppe 3 (56,06%) an Gruppe 4 (59,33) kann Gruppe 3 bei den Offenen Fragen die größte positive Ergebnisänderung (23,23%) aufweisen.

Gruppe 8 startet mit 39,44% im Prä-Test der Offenen Fragen. Resultierend durch eine Ergebnissteigerung um 17,78% erreicht diese Gruppe einen Wert von 57,22% im Post-Test.

4) Die größte positive Veränderung vom Prä-Test zum Post-Test zeigt Gruppe 3 bei den Offenen Fragen (23,23%), die kleinste positive Veränderung findet sich ebenfalls bei den Offenen Fragen für Gruppe 4 (11,91%).

Diagramm 21: Vergleich der Ergebnisse von Prä-Test und Post-Test für die Gruppen 3, 4 und 8



3.7) Ergebnisse für die Multiple-Choice-Fragen im Einzelnen und differenzierte Betrachtung aller Fragen nach inhaltlichen Kriterien

Die Auswahl der MC-Fragen erfolgte nach bestimmten inhaltlichen Kriterien (siehe Kapitel 2.2.2, S. 55 f), wodurch sich eine Einteilung in fünf verschiedene Gruppen ergibt (a-e).

Somit wird eine differenziertere Betrachtung der Ergebnisse für die Fragen zum Wissensstand im Fachbereich Orthopädie möglich.

- a) Fragen zu typisch orthopädischen Symptomen
(Frage 11 „Hammerfinger“, Frage 13 „Hallux rigidus“,
Frage 16 „pseudoradikuläre Beschwerden“, Frage 17 „Scheiben-Meniskus“)

Tabelle 25: Ergebnisse für die Fragen 11, 13, 16 und 17

Gruppe	n	Frage 11	Frage 13	Frage 16	Frage 17
		Richtigantworten in %			
1	37	18,92	21,62	24,32	43,24
2	121	15,70	14,88	33,88	44,63
3-Prä	33	21,21	30,30	84,85	39,40
3-Post	33	69,70	57,58	93,94	69,70
4-Prä	42	16,67	26,19	92,86	45,24
4-Post	42	85,71	50,00	97,62	69,05
5	86	70,93	45,35	93,02	62,79
6	42	59,52	38,10	92,86	52,38
7	115	53,04	32,17	90,43	47,83
8-Prä	30	60,00	26,67	60,00	40,00
8-Post	30	76,67	66,67	86,67	83,33

Es zeigt sich, dass Gruppe 1 nur bei Frage 17 (43,24%) einen Wert aufweisen kann, der deutlich über der Ratewahrscheinlichkeit von 20% Richtigantworten liegt.

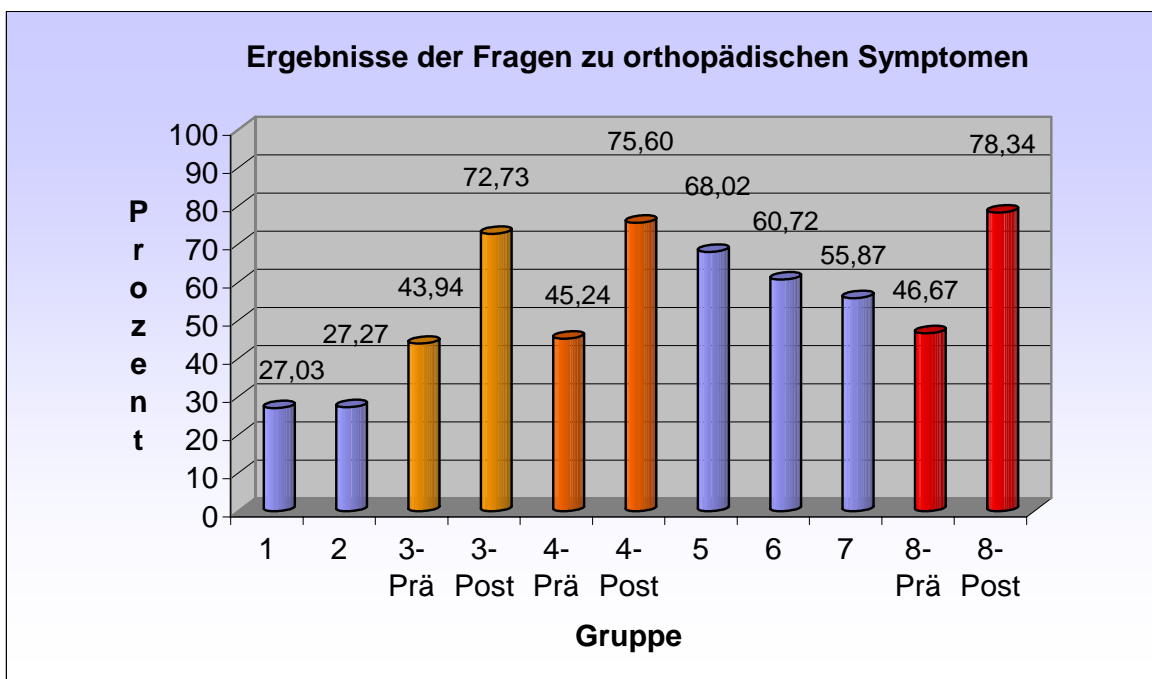
Die niedrigsten prozentualen Werte für die Fragen 11 und 13 werden von Gruppe 2 erzielt (15,70% bzw. 14,88%), die mit Gruppe 1 zusammen die schlechtesten Mittelwerte für die vier Fragen zu typisch orthopädischen Symptomen abliefern (27,03% in Gruppe 1 bzw. 27,27% in Gruppe 2).

Im Vergleich der Prä-Gruppen weist Gruppe 8-Prä mit 60,00% für Frage 11 einen wesentlich höheren Wert auf als die Gruppen 3-Prä (21,21%) und 4-Prä 16,67%). Gegensätzliches tritt bei Frage 16 auf. Hier fällt Gruppe 8-Prä mit 60,00% gegenüber den 84,85% von Gruppe 3-Prä und den 92,86% von Gruppe 4-Prä erheblich ab.

Hinsichtlich dieser vier Fragen kommen die schlechtesten Ergebnisse für Frage 13 zustande (14,88% bis 57,58%), die höchsten prozentualen Werte werden für Frage 16 erreicht (24,32% bis 97,62%).

Für die Gruppen 5 (68,02%), 6 (60,72%), und 7 (55,87%) erfolgt ein Rückgang der prozentualen Richtigantworten mit steigendem Ausbildungsstand.

Diagramm 22: Mittelwerte der Ergebnisse für die Fragen zu typisch orthopädischen Symptomen



- b) Fragen zur Statik
(Frage 12 „funktionelle Beinverkürzung“, Frage 24 „Beckenschiefstand“)

Tabelle 26: Ergebnisse für die Fragen 12 und 24

Gruppe	n	Frage 12	Frage 24
		Richtigantworten in %	erreichte Punktzahl in %
1	37 / 30	18,92	28,33
2	121 / 99	22,31	33,33
3-Prä	33	69,70	28,03
3-Post	33	81,82	56,82
4-Prä	42	83,33	55,36
4-Post	42	88,10	65,48
5	86	76,44	53,20
6	42	47,62	52,98
7	115	43,48	52,61
8-Prä	30	26,67	37,50
8-Post	30	46,67	58,33

Gruppe 1 und Gruppe 2 erreichen mit 23,63% bzw. 27,82% die niedrigsten Mittelwerte für die Fragen zur Statik.

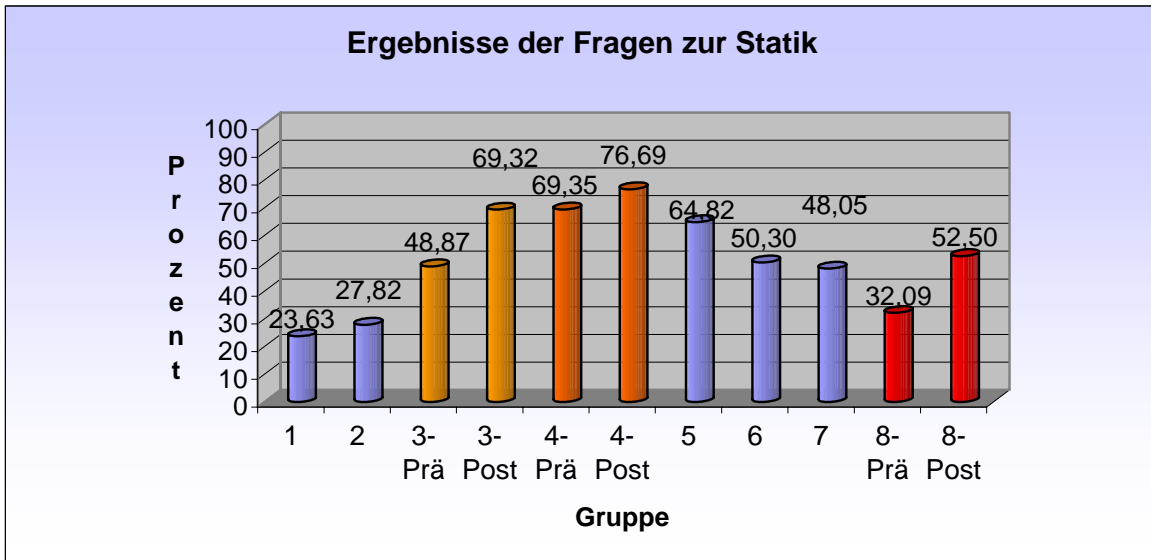
Den niedrigsten Wert für Frage 24 weist die Gruppe 3-Prä auf (28,03%), den höchsten Wert die Gruppe 4-Post (65,48%).

Im Vergleich der Prä-Gruppen schneidet die Gruppe 8-Prä bei Frage 12 (26,67%) wesentlich schlechter ab als die Gruppen 3-Prä (69,70%) und 4-Prä (83,33%).

Während Gruppe 3 und Gruppe 4 für Frage 12 viel höhere prozentuale Werte verzeichnen können als für Frage 24, so tritt bei Gruppe 8 (Universität Heidelberg) genau das Gegenteil auf.

Im Mittelwert der Fragen zur Statik kann Gruppe 4 im Prä-Test einen bereits höheren Wert (69,35%) vorweisen als Gruppe 3 im Post-Test (69,32%).

Diagramm 23: Mittelwerte der Ergebnisse für die Fragen zur Statik



- c) Fragen zur orthopädischen Diagnostik / orthopädische Zeichen in der Diagnostik (Frage 14 „Alters-Osteoporose“, Frage 15 „schmerzhafter Bogen“, Frage 18 „klinische Untersuchungen“, Frage 19 „Streckhemmung im Knie“)

Tabelle 27: Ergebnisse für die Fragen 14, 15, 18 und 19

Gruppe	n	Frage 14	Frage 15	Frage 18	Frage 19
Richtigantworten in %					
1	37	8,11	48,65	43,24	56,76
2	121	7,44	51,24	52,07	71,07
3-Prä	33	24,24	81,82	75,76	72,73
3-Post	33	45,45	96,97	84,85	84,85
4-Prä	42	23,81	85,71	80,95	78,57
4-Post	42	28,57	100,00	80,95	88,20
5	86	30,32	96,51	79,07	86,05
6	42	28,57	90,48	76,19	83,33
7	115	30,43	85,22	78,26	80,00
8-Prä	30	33,33	66,67	63,33	80,00
8-Post	30	33,33	93,33	63,33	90,00

Gruppe 1 (8,11%) und Gruppe 2 (7,44%) weisen mit Abstand die niedrigsten Werte für Frage 14 auf, erreichen aber für die Fragen 15, 18 und 19 Werte die deutlich über der 20%igen Ratewahrscheinlichkeit liegen (43,24% bis 71,07%).

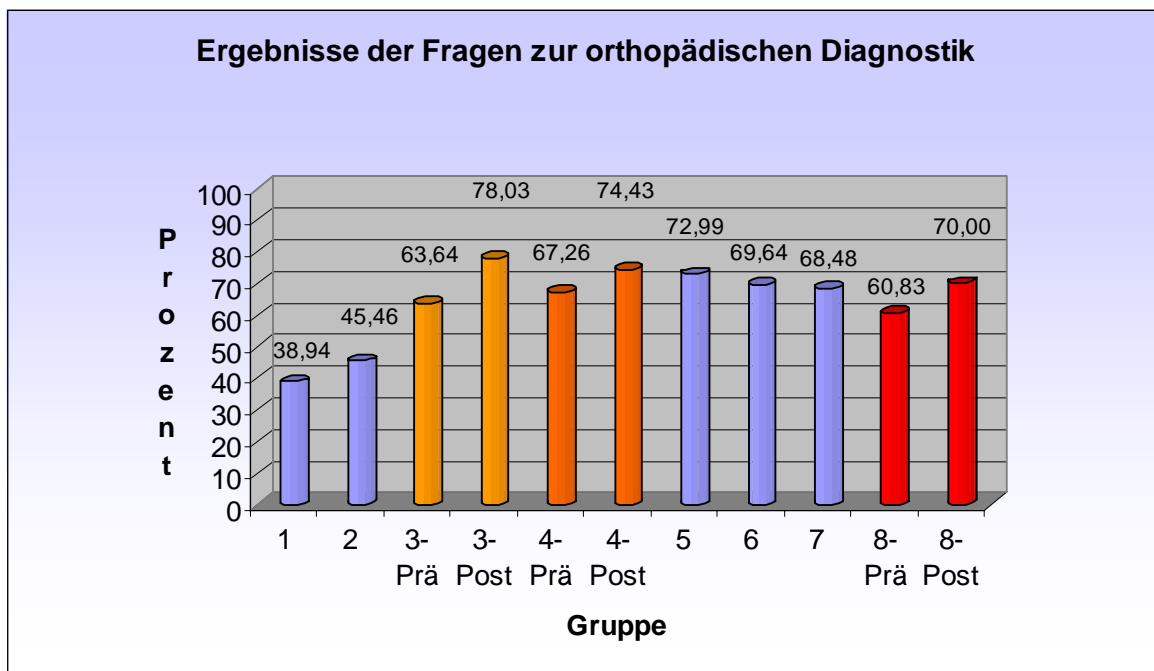
Frage 14 stellt die Frage mit den schlechtesten Ergebnissen aller MC-Fragen dar. Hier reichen die Werte von 7,44% in Gruppe 2 bis 45,45% in Gruppe 3-Post.

Gute Ergebnisse in allen Gruppen kommen für Frage 15 zustande. Der niedrigste Wert für diese Frage wird von Gruppe 1 mit 48,65% gestellt, in Gruppe 4-Post wird der Maximalwert von 100% erreicht.

Für die Mittelwerte dieser Fragenkategorie erscheinen für die Gruppen 1 (38,94%) und 2 (45,46%) höhere Werte als bei den Fragen zu typisch orthopädischen Symptomen (27,03% bzw. 27,27%) und zur Statik (23,63% bzw. 27,82%).

Im Vergleich mit Fragen zu den typisch orthopädischen Symptomen und zur Statik fallen hier die Mittelwerte der Ergebnisse für die Gruppen 5 (72,99%), 6 (69,64%) und 7 (68,48%) nur geringfügig von den Mittelwerten der Gruppen 3-Post (78,03%) und 4-Post (74,43%) ab.

Diagramm 24: Mittelwerte der Ergebnisse für die Fragen zur orthopädischen Diagnostik / zu orthopädischen Zeichen in der Diagnostik



- d) Fragen zu spezifisch-orthopädischen Krankheitsbildern
 (Frage 20 „Morbus Perthes“, Frage 21 „funktionelle Skoliose“,
 Frage 22 „Morbus Scheuermann“, „Frage 25 Morbus Perthes“)

Tabelle 28: Ergebnisse für die Fragen 20, 21, 22 und 25

Gruppe	n	Frage 20	Frage 21	Frage 22	Frage 25
		Richtigantworten in %			erreichte Punktzahl in %
1	37	24,32	21,62	37,84	3,33
2	121/99	29,75	7,44	45,45	4,80
3-Prä	33	57,58	84,85	78,79	4,55
3-Post	33	72,73	84,85	81,82	20,45
4-Prä	42	35,71	88,10	88,10	8,33
4-Post	42	61,90	83,33	90,48	21,43
5	86	54,65	79,07	82,56	17,44
6	42	47,62	73,81	73,81	16,67
7	115	41,74	74,78	73,04	16,74
8-Prä	30	73,33	50,00	60,00	7,50
8-Post	30	86,67	73,33	76,67	23,33

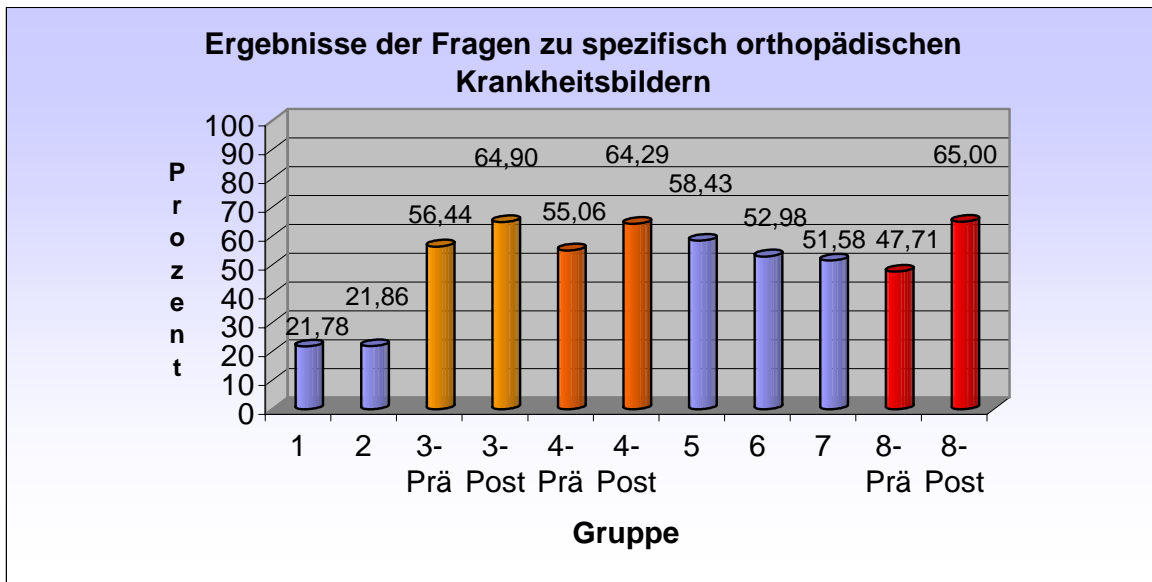
Für Frage 25 treten die schlechtesten Ergebnisse aller gestellten Fragen auf. Die Werte reichen hier von 3,33% in Gruppe 1 bis 23,33% in Gruppe 8-Post.

Mit einem Ergebnis von 7,44% in Frage 21 verzeichnet Gruppe 2 den mit Abstand niedrigsten Wert der MC-Fragen zu spezifisch orthopädischen Krankheitsbildern.

Im Vergleich der Prä-Gruppen weist Gruppe 8-Prä für Frage 20 einen wesentlich höheren Wert (73,33%) auf als die Gruppen 3-Prä (57,58%) und 4-Prä (35,71%).

Hinsichtlich der Mittelwerte dieser vier Fragen finden sich mit 21,78% für Gruppe 1 und 21,86% für Gruppe 2 die schlechtesten Ergebnisse bei den Fragen zu den spezifisch orthopädischen Krankheitsbildern.

Diagramm 25: Mittelwerte der Ergebnisse für die Fragen zu spezifisch orthopädischen Krankheitsbildern



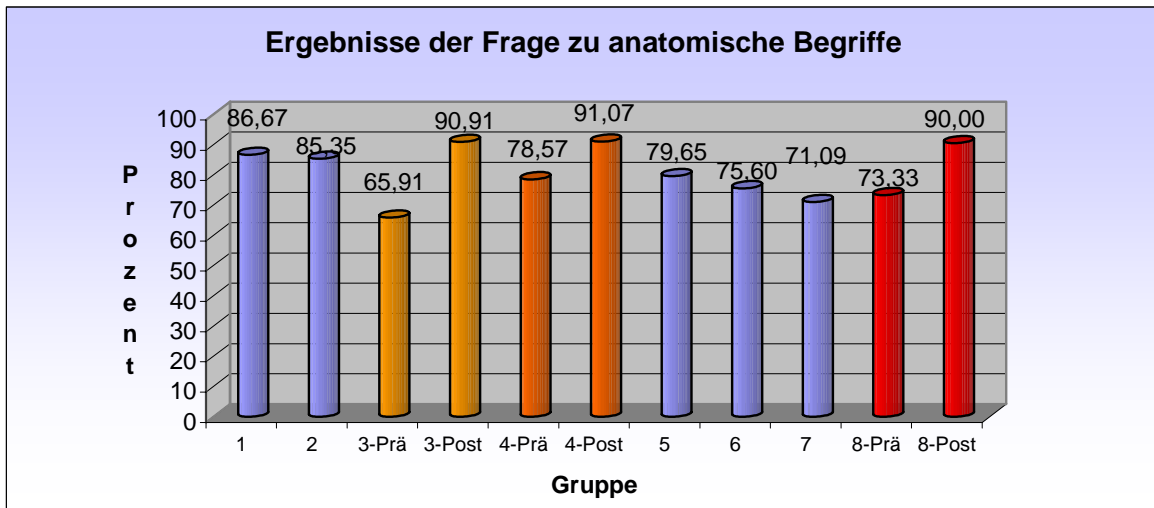
e) Frage zu anatomischen Begriffen (Frage 23 „Rotatorenmanschette“)

Tabelle 29: Ergebnisse für die Frage 23

Gruppe	n	Frage 23
		erreichte Punktzahl in %
1	30	86,67
2	99	85,35
3-Prä	33	65,91
3-Post	33	90,91
4-Prä	42	78,57
4-Post	42	91,07
5	86	79,65
6	42	75,60
7	115	71,09
8-Prä	30	73,33
8-Post	30	90,00

Frage 23 stellt die am besten beantwortete Frage der Offenen Fragen dar. Der niedrigste Wert liegt in Gruppe 3-Prä (65,91%), der höchste in Gruppe 4-Post (91,07%).

Diagramm 26: Mittelwerte der Ergebnisse für die Frage zu anatomischen Begriffen



An dieser Stelle seien die Ergebnisse nach inhaltlichen Kriterien nochmals zusammenfassend dargestellt:

Diagramm 27: Mittelwerte aller Gruppen (Gruppen 1-8) der Ergebnisse zum Wissensstand nach inhaltlichen Kriterien

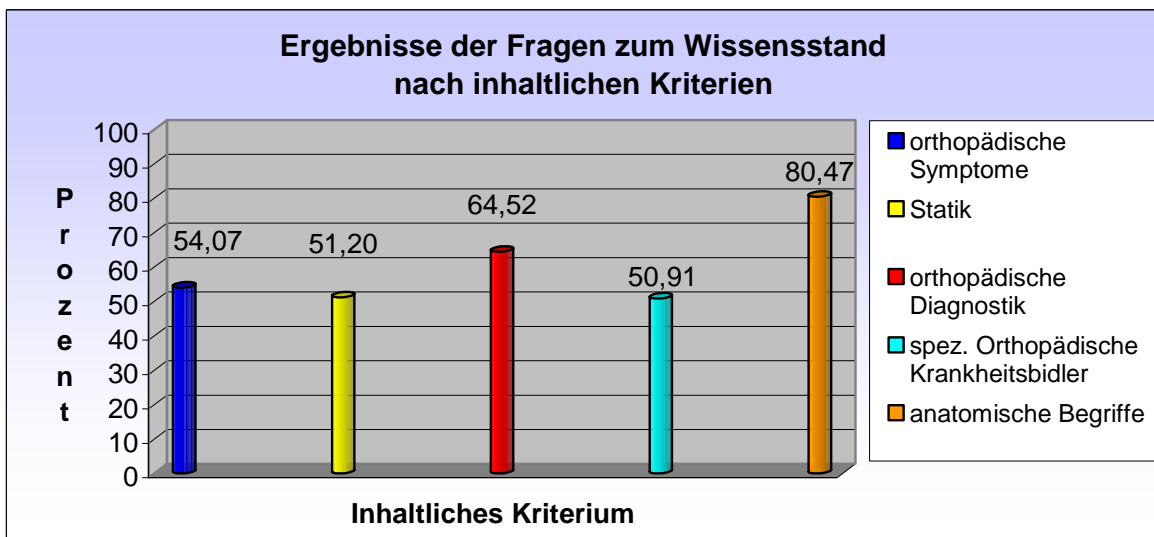


Tabelle 30: Ergebnisse für die MC-Fragen im Einzelnen

Gruppe	n	FR 11	FR 12	FR 13	FR 14	FR 15	FR 16
		Richtigantworten in %					
1	37	18,92	18,92	21,62	8,11	48,65	24,32
2	121	15,70	22,31	14,88	7,44	51,24	33,88
3-Prä	33	21,21	69,70	30,30	24,24	81,82	84,85
3-Post	33	69,70	81,82	57,58	45,45	96,97	93,94
4-Prä	42	16,67	83,33	26,19	23,81	85,71	92,86
4-Post	42	85,71	88,10	50,00	28,57	100,00	97,62
5	86	70,93	67,44	45,35	30,32	96,51	93,02
6	42	59,52	47,62	38,10	28,57	90,48	92,86
7	115	53,04	43,48	32,17	30,43	85,22	90,43
8-Prä	30	60,00	26,67	26,67	33,33	66,67	60,00
8-Post	30	76,67	46,67	66,67	33,33	93,33	86,67
Gruppe	n	FR 17	FR 18	FR 19	FR 20	FR 21	FR 22
		Richtigantworten in %					
1	37	43,24	43,24	56,76	24,32	21,62	37,84
2	121	44,63	52,07	71,07	29,75	7,44	45,45
3-Prä	33	39,40	75,76	72,73	57,58	84,85	78,79
3-Post	33	69,70	84,85	84,85	72,73	84,85	81,82
4-Prä	42	45,24	80,95	78,57	35,71	88,10	88,10
4-Post	42	69,05	80,95	88,10	61,90	83,33	90,48
5	86	62,79	79,07	86,05	54,65	79,07	82,56
6	42	52,38	76,19	83,33	47,62	73,81	73,81
7	115	47,83	78,26	80,00	41,74	74,78	73,04
8-Prä	30	40,00	63,33	80,00	73,33	50,00	60,00
8-Post	30	83,33	63,33	90,00	86,67	73,33	76,67

4.) Diskussion

4.1) Zum Stand der Literatur

Die Literaturrecherche zur Anfertigung der vorliegenden Arbeit zeigt auf, dass zum Themengebiet „Multiple-Choice Fragen und Offene Fragen – Anwendbarkeit und Vergleich“ bisher fast ausschließlich im anglo-amerikanischen Raum geforscht wurde. Ebenso fällt auf, dass die meisten dieser Studien in den Jahren 1975 bis 1990 durchgeführt wurden.

Einzig das Buch „Developing and Validating Multiple-Choice Test Items“ von Thomas M. Haladyna (28) aus dem Jahr 2004 repräsentiert ein themabezogenes Werk aus dem 21. Jahrhundert, das zwar einen sehr detaillierten Einblick in die Welt der MC-Fragen liefert, jedoch sich nicht eingehend dem Vergleich mit anderen Prüfungsarten widmet.

Eine ähnliche Arbeit wurde 1997 von Andrea Elmer angefertigt (20). Die Dissertation mit dem Thema „Analyse des fachorthopädischen Wissensstandes in verschiedenen Ausbildungsabschnitten und Vergleich zweier schriftlicher Prüfungsmethoden“ behandelte wie die vorliegende Arbeit die Themengebiete Prüfungswesen und Prüfungsmethoden in der medizinischen Ausbildung. Im praktischen Teil der Dissertation erfolgte eine Befragung an Studenten der Ruhr-Universität Bochum und an 19 Ärzten aus Kliniken und Praxen im Raum Bochum anhand von 16 Multiple-Choice-Fragen und 4 Offenen Fragen. Die Auswertung dieser Arbeit zeigt folgende Ergebnisse (20):

- 1) Der mit den verwendeten Messmethoden (MC-Fragen, Offene Fragen) ermittelte fachorthopädische Wissensstand entspricht dem jeweiligen Ausbildungsniveau.
- 2) Die größte positive Veränderung erfährt dieses Wissen im Laufe des Orthopädie-Praktikums, so dass angenommen werden kann, dass den Studierenden hier ein erheblicher Teil ihres Fachwissens vermittelt wird.
- 3) Im Anschluss an die Ausbildung stabilisiert sich dieses Wissen im Wesentlichen.
- 4) Offene Fragen stellen grundsätzlich eine sinnvolle Ergänzung zu MC-Prüfungen dar.

Eine Studie zum direkten Vergleich von MC-Fragen und Offenen Fragen wurde 1979 von Newble, Baxter und Elmslie (55) durchgeführt, indem sie mit der „matches-pairs“-Technik zwei äquivalente und inhaltlich identische Ausführungen des Fragebogens anwendeten. Dieser Fragebogen bestand jeweils zur Hälfte aus MC-Fragen und Offenen Fragen. Die Studie zeigte folgendes Resultat:

Die Ergebnisse waren in allen Gruppen für die MC-Fragen besser als für die Offenen Fragen. Newble et al. folgern daraus, dass MC-Fragen das anwendbare, klinische Wissen der Studierenden erheblich überschätzen, vor allem dasjenige der weniger kompetenten Studierenden. Die Autoren fügen an: „No important written or clinical examination should consist of only one type of testing method“ (55, S. 268).

In der Studie von Schwartz et al. (76) wurden anhand von acht verschiedenen Testmethoden (zwei MC-Tests, eine „modified-essay examination“ (MEQ), eine „objective structured clinical examination“ (OSCE) und eine „standardized-patient-examination“) die Leistungen von Studierenden im Anschluss an eine problemorientierte Chirurgie-Ausbildung ermittelt. Die Studierenden wurden ebenfalls von ihren Fakultäts-Tutoren, von externen Prüfern und von ihren Kommilitonen beurteilt. Zur Feststellung des Wissenszuwachses erfolgte ein Prä-Test und ein Post-Test einer Chirurgie-Prüfung des National Board of Medical Examiners (NBME). Die Auswertung der Studie ergab folgende Ergebnisse:

- 1) Es erfolgt ein hochsignifikanter Wissenszuwachs im Laufe des Chirurgie-Praktikums.
- 2) Studentisches Wissen und studentische Fähigkeiten können von Kommilitonen adäquat beurteilt werden.
- 3) Die Leistungssteigerung der Studierenden weist nicht nur auf einen Erwerb von Wissen hin, sondern auch von seiner Anwendbarkeit (clinical performance).
- 4) Die Gesamt-Reliabilität der acht Messmethoden beträgt 0,81. Beim Verzicht auf eine der acht Messmethoden sinkt der Wert für die Reliabilität auf unter 0,80.

Auch Schwartz kommt wie Newble et al. (55) zu folgender Kernaussage: „We advocate using various evaluation methods precisely because each measures a different aspect of student performance“ (76, S. 150).

Mit seiner Arbeit "Ein Beitrag zur Erfolgsmessung der medizinischen Ausbildung im Fach Orthopädie" legte Klocke (41) eine größtenteils auf Selbsteinschätzung basierende und somit auch kritisch zu beurteilende Studie vor. Die Arbeit soll eine Auskunft über die Präsenz einiger medizinischer Sachverhalte der Orthopädie bei den Studenten und den Ärzten geben (41). In der Studie wurden 1146 Ärzte und Studenten aus dem Raum Düsseldorf /Neuss zu 16 prägnanten orthopädischen Begriffen befragt, die je nach Kenntnisstand von „noch nie gehört“ bis „weiß die Antwort, bin sicher“ eingestuft werden sollten. Klocke folgerte aus den Ergebnissen: „Das medizinische Wissen im Fach Orthopädie stieg bis zum Ende des Studiums kontinuierlich an, fällt dann aber etwas ab, sofern man nicht im Fach tätig ist“ (41). Somit fand Klocke eine positive Korrelation zwischen dem Wissensstand und dem jeweiligen Ausbildungsstand der Kandidaten.

Ähnliche Studien wurden im deutschsprachigen Raum vor allem im Rahmen der Evaluation verschiedener Unterrichtsmodelle durchgeführt. So lässt sich hier beispielsweise die Evaluation für das Reformmodell des Praktikums der Chirurgie an der Ludwig-Maximilians-Universität München anfügen. Dabei führte man lehrzielorientierte Tests in der Ergebnisevaluation durch und es fand eine Messwertwiederholung am identischen Kollektiv sowohl vor, während als auch nach Instruktion im Sinne eines Prä-Post-Ansatzes statt (21). Im Vordergrund stand jedoch die Evaluation des reformierten Praktikums und weniger der Vergleich verschiedener Testmethoden.

Die Übersichtsarbeiten „Bericht über die Beurteilungsmethoden in der medizinischen Grundausbildung“ (42) der Kommission der Europäischen Gemeinschaften und die „Untersuchung ausländischer Prüfungssysteme im Studiengang Medizin im Zusammenhang mit der Diskussion zur neuen Approbationsordnung für Ärzte“ (67) der Planungsgruppe Medizin im Auftrag des Bundesministeriums für Gesundheit aus den Jahren 1987 und 1995 stellen die einzigen themabezogenen Werke zur Evaluation und dem Vergleich von Prüfungsmethoden in den letzten Jahren dar.

4.2) Prüfungswesen in der Diskussion

Die Diskussionen über das Prüfungswesen in der medizinischen Ausbildung werden immer Bestand haben. Von Seiten der Gesetzgeber wurden Arbeiten in Auftrag gegeben, die sich mit der medizinischen Ausbildung insgesamt oder im speziellen mit dem Prüfungswesen befassen (67, 89). Hingewiesen sei auf die Problematik der Kopplung von Lehre und Prüfung (19, 30), der Definition von Ausbildungszielen (19, 89), der strukturellen und materiellen Voraussetzungen zur Entwicklung eines entsprechenden Prüfungsverfahrens und der Auswirkungen von Prüfungsverfahren auf das Lernverhalten (19, 66, 79, 89).

Ein bisher in Deutschland vernachlässigtes Aufgabenfeld stellt die Evaluationsforschung dar, die in letzter Zeit jedoch immer mehr Bedeutung zugewiesen bekommt.

In den Kapiteln 1.1.5 bis 1.1.7 (siehe S. 24 ff) wurden die unterschiedlichen Prüfungsmethoden vorgestellt und hinsichtlich der Testgütekriterien miteinander verglichen. Es bleibt festzuhalten, dass keine Methode in der Lage ist, ärztliche Kompetenz in all ihren Dimensionen vollständig zu erfassen. Unter Berücksichtigung der oben genannten Arbeiten und Studien lässt sich folgern, dass ein Prüfungssystem, das den Testgütekriterien genügen will, eine Kombination aus mehreren Prüfungsmethoden darstellen muss: „Die beste Methode besteht darin, nicht ein einziges, sondern verschiedene Prüfungsverfahren vorzusehen“ (42, S.15). Die Arbeiten, die sich mit dem Vergleich verschiedener Prüfungsverfahren beschäftigen (19, 46, 55, 60, 67, 72), kommen zu dem gleichen Ergebnis, dass in einer Prüfung mehrere Prüfungsarten parallel verwendet werden sollten.

Abschließend lässt sich festhalten, dass die vorliegende Arbeit nur die Problematik des Prüfungswesens aufzeigen kann. Bisher erarbeitete neue Prüfungsmodelle, wie das vom Netzwerk aus dem „Positionspapier zur Reform des Prüfungswesens im Medizinstudiengang“ (19) 1994 vorgestellte Modell oder das ebenfalls 1994 vom Murrhardter Kreis entworfene Modell (53), müssen als Anfang einer grundlegenden und einheitlichen Reform des Prüfungssystems gesehen werden und verdeutlichen die Problematik, ein allseits zufrieden stellendes Prüfungskonzept zu entwickeln.

4.3) Vorbemerkungen

Die Durchführung einer solchen Befragung zur Lehre und zur Wissensvermittlung ist unter anderem von der Bereitschaft und Mitarbeit der Untersuchungsteilnehmer abhängig. Ein nur sporadisch beantworteter Fragebogen oder eine Beantwortung der Fragen ohne zugrunde liegende ernsthafte Beschäftigung mit der jeweiligen Fragestellung sind bei der Durchführung einer studentischen Befragung nicht auszuschießen.

Deswegen wurde vor Beginn der Befragung in jeder Gruppe deutlich darauf hingewiesen, dass nur ein ernsthaft ausgefüllter Fragebogen für eine Studie sinnvoll ist, sich jeder Untersuchungsteilnehmer 12 Minuten konzentriert zur Beantwortung der Fragen Zeit nehmen kann und die Untersuchung vor allem auch für die Studierenden einen Sinn aufweist.

Hinsichtlich der Anzahl der gewerteten Fragebögen fällt auf, dass nur in Gruppe 1 und Gruppe 2 ein Teil der Fragebögen nicht in die Wertung aufgenommen werden konnte.

Für die MC-Fragen wurden nur die Fragebögen ausgewertet, in denen für jede der 12 MC-Fragen eine Lösungsmöglichkeit angegeben wurde. Dieses Vorgehen wurde aus folgenden Gründen gewählt: Für die MC-Fragen war keine „ich-weiß-nicht“-Option vorgegeben. Unabhängig vom Wissensstand war es jeder Testperson möglich, eine Auswahlantwort zu raten, wenn er die korrekte Antwort nicht wusste. Um hier die gleichen Voraussetzungen für alle Kandidaten wirksam zu machen, konnten nur vollständig bearbeitete MC-Fragen gewertet werden. Es wurde bei der Durchführung der Untersuchung ausdrücklich darauf hingewiesen, dass alle MC-Fragen zu bearbeiten sind. So wurden in allen Gruppen (Gruppen 1-8) die MC-Fragen vollständig beantwortet und Teil II (MC-Fragen) geht vollständig in die Wertung mit ein.

Die Offenen Fragen wurden nur für diejenigen Kandidaten ausgewertet, die die Offenen Fragen auch beantwortet hatten. Diejenigen Testpersonen, die diesen Teil vollständig unausgefüllt gelassen hatten, gingen nicht in die Gesamtbewertung mit ein. Dieses Vorgehen erfolgte aus folgenden Gründen:

In den Gruppen 1 und 2 wurde in mehreren Fragebögen keine der drei Offenen Fragen beantwortet. Um einen Vergleich der Ergebnisse dieser Gruppen mit den Ergebnissen der anderen Gruppen überhaupt zu ermöglichen, wurden Fragebögen mit null Punkten für die Offenen Fragen nicht in die Wertung für diesen Teil des Fragebogens aufgenommen. In Gruppe 1 wurden in 18,92% der Fragebögen die Offenen Fragen überhaupt nicht bearbeitet. Somit gehen in Gruppe 1 nur 30 von 37 Fragebögen in die Wertung ein. In Gruppe 2 wurden 18,18% der Fragebögen ohne die Bearbeitung der Offenen Fragen abgegeben, so dass in Gruppe 2 nur 99 von 121 Fragebögen gewertet wurden.

Dieses Ergebnis deutet darauf hin, dass die Studierenden im 1. klinischen und 2. klinischen Semester zum Teil mit dem Fragebogen überfordert waren, was auch in den Kommentaren deutlich wurde. Es ist nachzuvollziehen, dass ein überwiegend aus unbekanntem Inhalt bestehender Test die Motivation, sich eingehend mit den Fragen auseinanderzusetzen, erheblich reduziert.

Es muss ebenfalls darauf hingewiesen werden, dass die Durchführung der Befragung in den Gruppen 1, 2, 5, 6 und 7 meist im Anschluss an eine Klausur oder an ein doppelstündiges Pflichtpraktikum stattfand, so dass die Aufmerksamkeit und Konzentration der Studierenden als strapaziert angesehen werden muss.

In mehreren Fällen sahen die Befragten der Gruppen 1 und 2 keinen oder nur wenig Sinn in der Befragung zu diesem Zeitpunkt ihrer Ausbildung, was teilweise auch in den Kommentaren in Frage 30 (Befürwortung einer Evaluation zum Thema Studienordnung und Fachwissenstand) zum Ausdruck kam. Dennoch ist eine Erhebung in diesen Semestern wesentlich im Sinne einer Erhebung des sukzessiven Wissenszuwachses.

4.4) Diskussion der Ergebnisse

Um eine sinnvolle und korrekte Diskussion der Ergebnisse zu garantieren, muss bei der Betrachtung der Ergebnisse folgendes beachtet werden:

Bei der vorliegenden Arbeit handelt es sich nicht um eine Langzeitstudie, sondern um punktuelle Untersuchungen nicht-identischer Testgruppen auf unterschiedlichem Ausbildungsniveau. Da für alle Testgruppen ein identischer Fragebogen verwendet wurde, lassen sich die Ergebnisse miteinander vergleichen. Die Ergebnisse dürfen jedoch nicht in einer fortlaufenden Reihe gedacht werden, da die Testgruppen keinen identischen Charakter besitzen. Aus diesem Grund müssen in der folgenden Diskussion der Ergebnisse die Begriffe „Steigerung“, „Verbesserung“, „Rückgang“ und „Absinken“ in diesem Sinne verstanden werden.

4.4.1) Diskussion der Ergebnisse zu Teil I

Frage 1) Einschätzung des Wissensstandes zu den Änderungen durch die neue Approbationsordnung

Wie bereits in Kapitel 3.2 dargestellt wurde, liegen in allen Gruppen die prozentualen Höchstwerte (37,84% bis 53,33%) in der Kategorie „unentschieden“. Dies ist offenkundig ein Hinweis darauf, dass die Studierenden weder durch die Universitäten detailliert über die Änderungen durch die neue Approbationsordnung informiert wurden, noch durch Eigeninitiative sich ein detailliertes Wissen darüber angeeignet haben. Auffällige Unterschiede treten im interuniversitären Vergleich auf. Da für jede Gruppe der Universität Regensburg (Gruppen 1-7) der Wert für die Antwort „schlecht“ (26,19% bis 41,74%) prozentual höher ist als für die Antwort „gut“ (11,30% bis 27,27%), aber in den Gruppen 8-Prä und 8-Post (Universität Heidelberg) die Werte für die Antwort „gut“ (33,33% bzw. 40,00%) diejenigen für die Antwort „schlecht“ (13,33% bzw. 10,00%) deutlich übersteigen, lässt sich den Studierenden der Universität Heidelberg ein besseres Wissen für die Änderungen über die neue AO zuschreiben. Ursache für diese Tendenz könnte die Tatsache sein, dass den Studierenden der Universität Heidelberg im Zuge der Umstellung des Lehrsystems auf den Modellstudiengang Heicumed gleichzeitig auch ausreichend Information über die Änderungen durch die neue AO zugekommen ist.

Frage 2) Anzahl der Staatsexamina („Hammerexamen“ am Ende)

In der Auswertung zeigt sich, dass die Antwort „schlecht“ von jeder Gruppe am häufigsten gewählt wurde (36,36% bis 55,81%). In den Gruppen mit dem höheren Ausbildungsstand verschieben sich im Vergleich zu den Gruppen mit niedrigerem Ausbildungsstand die zweithöchsten Werte deutlich von der Kategorie „unentschieden“ hin zu „sehr schlecht“. Diese Tatsache liegt darin begründet, dass diese Gruppen mit dem von den Medien deklarierten „Hammerexamen“ in naher Zukunft konfrontiert sind und vor allem ursprünglich mit der Erwartung eines dreigeteilten Staatsexamens den klinischen Teil des Studiums angetreten haben. Da die Änderung der Anzahl der Staatsexamina von den Studierenden als „Zusammenlegung von drei Staatsexamina in ein Examen mit enormen Lernaufwand und wahrscheinlichem Freisemester“ verstanden wird, erklärt dies die sehr niedrigen Werte für die Kategorien „sehr gut“ (0,00% bis 3,31%) und „gut“ (0,00% bis 13,33%). Diese Änderung repräsentiert den wohl am häufigsten diskutierten und am meisten publik gewordenen Änderungspunkt der neuen AO.

Frage 3) Bevorzugte Prüfungsmethode

Hinsichtlich der bevorzugten Prüfungsmethode erreicht die Antwort „MC-Fragen“ in allen Gruppen die mit Abstand höchsten prozentualen Werte (45,24% bis 76,76%). Hier liegt die Erklärung nahe, dass auf Grund der überwiegenden Anwendung von MC-Fragen in den klinischen Semestern und vor allem in der Vorklinik die Studierenden mit der Prüfungsmethode Multiple-Choice Fragen am Vertrautesten sind und die 20%ige Ratewahrscheinlichkeit der richtigen Antwort zu schätzen gelernt haben. Dass die Prüfungsmethode „Offene Fragen“ in den Gruppen 8-Prä (6,67%) und 8-Post (0,00%) der Universität Heidelberg die niedrigsten prozentualen Werte aufweist liegt die Tatsache zugrunde, dass im Modellstudiengang Heicumед das Prüfungssystem vermehrt auf mündliche Prüfungen als Ergänzung zu den MC-Fragen ausgelegt ist. Bei den Gruppen 3 und 4 mit dem Orthopädie-Praktikum im Stundenplan fällt auf, dass die Werte für die Offenen Fragen im Vergleich vom Prä-Test zum Post-Test ansteigen (von 12,12% auf 30,30% bzw. von 21,43% auf 33,33%), was auf Kosten der Werte für die MC-Fragen geschieht. Hier spielt sicherlich der Stundenplan dieser beiden Semester eine Rolle, da beide Gruppen sich im Laufe des Semesters vermehrt mit Offenen Fragen in Klausuren auseinandersetzen mussten, z.B. im Praktikum Innere Medizin.

Frage 4) Vermittlung der Inhalte innerhalb der Praktika

Die prozentualen Höchstwerte für die Antwort „gut“, teilweise relativ hohe Werte für die Extremantwort „sehr gut“ (13,33% bzw. 14,29%) und Minimalwerte für die Antwort „sehr schlecht“ (0,00% bis 2,61%) sprechen für eine ausreichend gute Vermittlung der Inhalte innerhalb der Praktika. Hier gilt es zu unterscheiden, ob das positive Abschneiden der Praktika auf der tatsächlich guten Vermittlung der Inhalte beruht, oder ob die von den Studierenden allgemein gute Einschätzung der Praktika als willkommene Abwechslung zum rein theoretischen Unterricht und als Heranführen an den klinischen Alltag zu Grunde liegt.

Frage 5) Ausreichende Vorstellung von Patienten in den Vorlesungen

Bei dieser Frage tritt der Trend auf, dass sich die Prozentualen Werte von der Kategorie „trifft zu“ (72,97% in Gruppe 1) im Verlauf des Studiums in Richtung der Kategorie „trifft nicht zu“ (47,83% in Gruppe 7) verschieben. Der gleiche Trend ist ebenfalls bei den Extremantworten festzustellen. Während Kategorie „trifft völlig zu“ im 1. klinischen Semester noch 21,63% erreicht und im 6.klinischen Semester nur noch bei 2,61% liegt, erhöhen sich die Werte in der Kategorie „trifft gar nicht zu“ von anfänglichen 0,00% im 1. klinischen Semester bis auf 12,17% im 6. klinischen Semester.

Diese Beobachtungen lassen unterschiedliche Hypothesen zu:

Zum einen ist es möglich, dass die hohen Werte in den Semestern mit niedrigem Ausbildungsstand und vor allem im 1. klinischen Semester auf der Tatsache beruhen, dass nach Absolvierung des Ersten Abschnitts der Ärztlichen Prüfung und mit Eintritt in die klinischen Semester die Vorstellung mehrerer Patienten von den Studierenden als definitiv ausreichend empfunden wird, da diese zuvor nur wenig Kontakt mit Patienten aufweisen konnten. Andererseits können diese Beobachtungen auch darin begründet sein, dass die Studierenden in den höheren klinischen Semestern mit Ausblick auf das Praktische Jahr und das Berufsleben die Vorstellung der Patienten in den Vorlesungen als nicht mehr ausreichend empfinden, da diese in den Praktika vermehrt mit Patienten Umgang haben und von einem anderen Ausgangsniveau an Patientenvorstellungen ausgehen.

Frage 6) Ausreichende Vorstellung von Patienten in den Praktika

Bei der Auswertung fällt auf, dass nur in Gruppe 2 und Gruppe 3-Prä die Antwort „trifft nicht zu“ am stärksten vertreten ist (34,71% bzw. 39,39%), in allen anderen Gruppen wird die Vorstellung von Patienten in den Praktika als ausreichend eingestuft. Die besseren Werte im Vergleich zu Frage 5 lassen sich durch den vermehrten Kontakt mit Patienten innerhalb der Praktika begründen. Die Ursache für die Einschätzung nicht ausreichender Patientenvorstellungen in den Praktika in Gruppe 2 ist wahrscheinlich mit dem Stundenplan dieser Gruppe vereinbar, da das 2. klinische Semester überwiegend in Grundlagenfächern wie Klinische Chemie/Laboratoriumsdiagnostik, Humangenetik und Pharmakologie/Toxikologie ohne Patientenkontakt unterrichtet wurde.

Die Gruppen 3 und 4 mit dem Orthopädie-Praktikum im Stundenplan weisen im Vergleich von Prä-Test mit Post-Test eine Steigerung der Werte für die Antwort „trifft völlig zu“ auf (von 9,09% auf 27,27% bzw. von 4,76% auf 28,57%). Auch hier findet sich die Ursache im Stundenplan, da nicht wie in den ersten beiden klinischen Semestern vorwiegend Grundlagenfächer doziert werden, sondern im 3. klinischen und 4. klinischen Semester das Praktikum Orthopädie, das Praktikum Innere Medizin, das Praktikum Chirurgie und das Praktikum Psychosomatik mit jeweils vermehrten Patientenvorstellungen im Mittelpunkt stehen.

Frage 7) Strukturierung des Studiums in Bezug auf die Blockpraktika

In den Gruppen 1, 2 und 3 befindet sich die prozentuale Mehrheit noch in den Kategorien „sehr gut“, „gut“ und „unentschieden“, verlagert sich dann aber mit steigendem Ausbildungsstand in Richtung der Antworten „unentschieden“, „schlecht“ und „sehr schlecht“. Das prozentual schlechtere Abschneiden der Blockpraktika in den Semestern mit höherem Ausbildungsstand ist dadurch bedingt, dass diese Gruppen im System der Blockpraktika den Nachteil sehen, dass durch den ständigen Wechsel der Fächer und den Abschlussklausuren nach zweiwöchiger Ausbildung das Lernen nur auf das Kurzzeitgedächtnis ausgelegt ist. Diese Erklärung findet sich auch in den Kommentaren zu Frage 10 bestätigt, in der von den Gruppen 5, 6 und 7 eben dieses Problem als Änderungspunkt am derzeitigen Kurssystem proklamiert wird.

Frage 8) Ausreichender praktischer Unterricht von 4 Wochen Chirurgie, 5 Wochen Innere Medizin und 2 Wochen Orthopädie

Im 1. klinischen Semester ohne Erfahrungswert für diese Fachbereiche findet sich der Höchstwert noch in Kategorie „unentschieden“ (40,54%), verschiebt sich über die Antwort „trifft nicht zu“ hin zu den Topwerten für die Kategorie „trifft gar nicht zu“ in den Gruppen 5, 6 und 7 (41,86% bzw. 54,67% bzw. 62,61%). Diese Verschiebung korreliert sicherlich mit dem Erfahrungswert der Studierenden im Laufe der klinischen Ausbildung. Ebenfalls lässt sich als Grund dafür anführen, dass mit steigendem Ausbildungsstand auch vermehrt „kleinere Fächer“ wie zum Beispiel Klinische Umweltmedizin, Epidemiologie/Medizinische Biometrie, Arbeits- und Sozialmedizin, Physikalische Medizin und Medizin des Alterns, gelehrt werden und in den Semestern mit höherem Ausbildungsstand vermehrt auf die „großen Fächer“ wie Innere Medizin und Chirurgie Wert gelegt wird, die den Großteil des Praktischen Jahres einnehmen.

Die Gruppen 8-Prä und 8-Post der Universität Heidelberg weisen bei dieser Frage mit jeweils 76,67% in der Kategorie „unentschieden“ die Höchstwerte auf, da sich im Modellstudiengang Heicumed die zeitliche Aufteilung der Fachbereiche unterschiedlich gestaltet. So waren bei dieser Frage in diesen Gruppen auch mehrere Randnotizen vermerkt, bei denen auf diese Tatsache hingewiesen wurde.

Frage 9) Zufriedenheit als Medizinstudent mit aktueller Situation

Die prozentuale Mehrheit erfährt im Laufe der Ausbildung eine Verlagerung von den Antworten „sehr zufrieden“ und „zufrieden“ hin zu „unentschieden“ und „nicht zufrieden“. Die große Zufriedenheit in den ersten klinischen Semestern lässt sich mit folgendem Sachverhalt erklären: Nach Absolvierung des Ersten Abschnittes der Ärztlichen Prüfung und somit dem Eintritt in den klinischen Abschnitt des Studiums fühlen sich die Studierenden dem Arztberuf näher, indem sie nun nicht mehr an der Universität sondern am Universitätsklinikum unterrichtet werden und sich im Stundenplan klinische Fächer mit Patientenkontakt befinden. Die steigende Unzufriedenheit in den höheren Semestern ist in der studentischen Routine und dem größeren Erfahrungswert auch an subjektiven Antipathien gegenüber der aktuellen Studiensituation oder bestimmten Fachbereichen begründet. Ebenso kann der aktuelle Ausblick in das Arbeitsleben bei den Studierenden Unmut hervorrufen.

Frage 10) Änderungspunkte zur aktuellen Studiensituation

Von der Möglichkeit, eine subjektive Meinung in Bezug auf Änderungspunkte an der aktuellen Studiensituation abzugeben, wurde im Schnitt von 73,43% aller Untersuchungsteilnehmer Gebrauch gemacht. Am häufigsten wurde der Änderungspunkt „wichtige Fächer länger“ angemerkt. Die mit dem höheren Ausbildungsstand steigenden Werte dafür korrelieren mit dem Erfahrungswert der Studierenden (44,35% im 6. klinischen Semester). Oft findet sich in den höheren Semestern noch der Zusatz „mehr Innere Medizin“, „mehr Chirurgie“ oder „weniger Zeit für kleine Fächer wie Psychosomatik, Arbeits- und Sozialmedizin, Geschichte und Ethik der Medizin“. Dies deutet darauf hin, dass sich die Studierenden vor dem PJ mit verpflichtenden vier Monaten Innere Medizin und vier Monaten Chirurgie im Ausbildungsplan nicht ausreichend auf diese Zeit vorbereitet fühlen, da diese grundlegenden Fächer zu wenig auf Kosten von im Sinne der Studierenden unwichtigen Fächern gelehrt wurden.

Stark vertreten ist der Änderungspunkt „mehr Praxis“. Im 1. klinischen Semester ist der Wert dafür auf Grund des Eintretens in den klinischen Abschnitt noch vergleichsweise niedrig (8,11%), erreicht aber in allen anderen Semestern hohe Werte (18,18% bis 30,43%) als Wunsch nach einem noch praktischer ausgerichteten Studium.

Die Werte für den Änderungspunkt „kein Hammerexamen“ durchlaufen ebenfalls einen Anstieg mit steigendem Ausbildungsstand, was sicherlich mit dem noch verbleibenden Zeitraum bis zur Absolvierung des Zweiten Abschnitts der Ärztlichen Prüfung zusammenhängt. Ebenfalls zu dieser Tendenz trägt die Tatsache bei, dass die Studierenden in den höheren Semestern noch mit der Vorstellung eines dreigeteilten Staatsexamens den klinischen Abschnitt ihres Studiums angetreten sind.

Dass der Änderungswunsch „Vergütung PJ“ ab dem 3. klinischen Semester steigende Werte aufweist, verwundert nicht, denn ab diesem Zeitpunkt setzen sich die ersten Studierenden mit Bewerbungen für das PJ auseinander.

Ab dem 4. klinischen Semester wird der Änderungspunkt „nur Kurzzeitgedächtnis“ hinsichtlich des von den Klausuren bestimmten Lernverhaltens abgegeben. Dies ist offensichtlich darin zu begründen, dass sich vor allem im 6. klinischen Semester die Studierenden bereits mit der Vorbereitung auf das Staatsexamen auseinandersetzen und aufgrund ihres Erfahrungswertes die Nachteile eines Blockpraktikums mit Abschlussklausur mit auf das Kurzzeitgedächtnis ausgerichtetem Lernverhalten realisieren.

4.4.2) Diskussion der Ergebnisse zu Teil II (MC-Fragen)

Die Mittelwerte für Gruppe 1 sind mit 30,63% von allen Gruppen am niedrigsten. Da die Mittelwerte 10,63% über der Ratewahrscheinlichkeit liegen, darf angenommen werden, dass die Studierenden im vorklinischen Studienabschnitt zumindest in einigen Bereichen Kontakt zum Fach Orthopädie haben. Die minimale Verbesserung von Gruppe 2 gegenüber Gruppe 1 um 2,36 Prozentpunkte lässt nicht auf einen partiellen Wissenserwerb orthopädischer Inhalte infolge der Absolvierung klinischer Fächer wie Klinische Chemie/Laboratoriumsdiagnostik, Allgemeine Pharmakologie/Toxikologie, Humangenetik und Klinischer Untersuchungskurs zurückfolgern. Vielmehr ist das 2. klinische Semester hinsichtlich des Wissensstandes im Bereich der MC-Fragen mit dem Niveau des 1. klinischen Semester gleichzusetzen.

Die Gruppe 3-Prä absolviert den Prä-Test in der ersten Unterrichtsstunde des Orthopädie-Praktikums, hat aber bereits die einwöchige Einführung in die Orthopädie zu Beginn des Semesters besucht. Die deutliche Steigerung um 27,11 Prozentpunkte von Gruppe 3-Prä zu Gruppe 2 lässt sich durch die inhaltlichen Überschneidungen der Einführung in die Orthopädie mit dem Orthopädie-Praktikum und durch die Auseinandersetzung mit dem Fachbereich Orthopädie erklären.

Weitere Teilnehmergruppe am Orthopädie-Praktikum war die Gruppe 4. Gruppe 4-Prä weist im Vergleich mit Gruppe 2 die insgesamt größte Veränderung der Ergebnisse um 29,11 Prozentpunkte und im Vergleich mit Gruppe 3-Prä einen höheren Ausgangswert von 2,00 Prozentpunkten zu Beginn des Orthopädie-Praktikums auf. Als Erklärung hierfür lässt sich anführen, dass Gruppe 4-Prä auf Grund des höheren Ausbildungsstandes eine Erweiterung des Lehrplans durch Praktika und Kurse durchlaufen hat, in denen es zu inhaltlichen Überschneidungen mit Teilgebieten aus der Orthopädie kommt. Besonders ist hier das Praktikum der Chirurgie zu nennen. Dieses Ergebnis stimmt bis zu diesem Abschnitt der Ausbildung überein mit den Ergebnissen von Klocke (41), der fand, dass das orthopädische Wissen im Verlauf des Studiums kontinuierlich ansteigt.

Ausgehend vom Leistungsniveau der Gruppen 3-Prä und 4-Prä nach Absolvierung der einwöchigen Einführung in die Orthopädie erfolgt erneut eine deutliche Veränderung der Ergebnisse im Vergleich vom Prä-Test zum Post-Test. So erfährt Gruppe 3 während des Orthopädie-Praktikums eine Verbesserung der Ergebnisse um 16,92 Prozentpunkte, Gruppe 4 um 14,88 Prozentpunkte.

Hierbei muss beachtet werden, dass die Ergebnissteigerung von Gruppe 3-Prä und 4-Prä gegenüber Gruppe 2 nach Absolvierung der einwöchigen Einführung in die Orthopädie (\emptyset 28,11 Prozentpunkte) nicht mit der Ergebnissteigerung von Gruppe 3-Post und 4-Post gegenüber Gruppe 3-Prä und Gruppe 4-Prä nach Absolvierung des Orthopädie-Praktikums (\emptyset 15,90 Prozentpunkte) verglichen werden darf. Die Ergebnissteigerung nach Absolvierung der Einführung in die Orthopädie basiert auf einem Ausgangsniveau mit kaum vorhandenem Wissen im Fachbereich Orthopädie, während der Ergebnissteigerung im Laufe des Orthopädie-Praktikums als Ausgangsniveau ein ausreichend vorhandenes Wissen im Fachbereich Orthopädie zugrunde liegt. Somit darf angenommen werden, dass den Studierenden ein erheblicher Teil ihres orthopädischen Faktenwissens im Laufe des Orthopädie-Praktikums vermittelt wird.

Im Anschluss an das Orthopädie-Praktikum erfährt das orthopädische Faktenwissen im Laufe der Zeit einen Rückgang, der im 6. klinischen Semester wieder das Ausgangsniveau vor Absolvierung des Orthopädie-Praktikums (60,87%) erreicht. Dies ist offensichtlich ein Hinweis darauf, dass sich bei einer nicht vorhandenen Auseinandersetzung mit dem Fachbereich Orthopädie das Faktenwissen auf ein Niveau stabilisiert, das mit dem Wissen bei Auseinandersetzung mit dem Fachbereich Orthopädie definitiv nicht konkurrieren kann. Dieses Ergebnis steht im Widerspruch zu den Resultaten von Klocke (41), der fand, dass das orthopädische Wissen bis zum Ende des Studiums kontinuierlich ansteigt. Jedoch scheinen die in Kapitel 1.2 aufgestellte Hypothese und die in der vorliegenden Arbeit erreichten Ergebnisse als annehmbarer und wahrscheinlicher. Dafür spricht der Sachverhalt, dass bei einem spezifischen orthopädischen Test in einer Gruppe mit aktueller Auseinandersetzung und daraus resultierendem Wissenserwerb bessere Ergebnisse erzielt werden können, als in Gruppen, bei denen der Erwerb des Wissens vor längerer Zeit stattfand, spezifisches Faktenwissen in Vergessenheit geraten ist und auch nicht mit einem Erhalt von orthopädischen Faktenwissen durch die Überlagerung mit anderen Fachgebieten zu rechnen ist.

Im Vergleich der Prä-Gruppen mit den Post-Gruppen kann Gruppe 8 (Universität Heidelberg) bei den MC-Fragen die größte Ergebnissteigerung mit 19,73 Prozentpunkten aufweisen. Es muss hier angefügt werden, dass diese Veränderung auf einem deutlich niedrigerem Ausgangswert (53,33%) basiert und als Abschlusswert nach Absolvierung des Orthopädie-Praktikums (73,06%) nur annähernd die Werte des Post-Tests der Gruppen an der Universität Regensburg (77,02% bzw. 76,98%) erreicht. Es darf bei der Teilnehmergruppe der Universität Heidelberg ebenfalls angenommen werden, dass den Studierenden ein erheblicher Teil ihres orthopädischen Faktenwissens im Laufe des Orthopädie-Praktikums vermittelt wird.

Die Werte für die Standardabweichung weisen mit 14,77% (Gruppe 1) bis 29,74% (Gruppe 4-Prä) relativ hohe Werte auf, insgesamt lässt sich aber keine einheitliche Tendenz erkennen.

Bei der statistischen Analyse wurden die p-Werte berechnet und das Signifikanzniveau auf $p < 0,05$ festgelegt. Im Vergleich der Gruppe 1 mit Gruppe 2 (unpaare Stichproben) bestätigen die ausschließlich über dem Signifikanzniveau liegenden p-Werte die Tatsache, dass im 1. klinischen Semester weder eine orthopädisch relevante Wissensvermittlung noch eine Überschneidung der Lehrinhalte mit den dozierten klinischen Grundlagenfächern stattfindet.

Bei der Analyse von Gruppe 2 mit Gruppe 3-Prä (bereits absolvierte Einführung in die Orthopädie) finden sich in acht der 12 Fragen p-Werte $< 0,05$, was auch insgesamt für einen statistisch signifikanten Unterschied zwischen diesen beiden Gruppen und für einen Wissenszuwachs basierend auf der Einführung in die Orthopädie spricht.

Im Vergleich der Prä-Gruppen mit den Post-Gruppen treten bei Gruppe 3 für acht der 12 Fragen und bei Gruppe 4 für sechs der 12 Fragen p-Werte $< 0,05$ auf. Anhand dieser Ergebnisse lässt sich ein statistisch signifikanter Unterschied im Vergleich der Prä-Gruppen mit den Post-Gruppen feststellen, resultierend auf dem Lehrerfolg durch das Orthopädie-Praktikum.

Im Vergleich der Gruppen 4-Post mit Gruppe 5, Gruppe 5 mit Gruppe 6 und Gruppe 6 mit Gruppe 7 finden sich fast ausschließlich p-Werte $> 0,05$, was auf statistisch nicht signifikante Unterschiede zwischen den Gruppen hinweist. Diese Ergebnisse bestätigen den leichten Rückgang der Ergebnisse mit dem steigenden Ausbildungsniveau.

Bei der Analyse der Teilnehmergruppe des Orthopädie-Praktikums an der Universität Heidelberg (Gruppe 8) finden sich für 9 der 12 Fragen p-Werte $< 0,05$. Diese Werte unterhalb des Signifikanzniveaus deuten auf einen statistisch signifikanten Unterschied von Gruppe 8-Post gegenüber Gruppe 8-Prä hin, was auf einen klaren Lehrerfolg basierend auf dem Orthopädie-Praktikum hinweist.

4.4.3) Diskussion der Ergebnisse zu Teil II (Offene Fragen)

An dieser Stelle sei noch einmal darauf hingewiesen, dass in den Gruppen 1 und 2 die Offenen Fragen nicht in allen Fragebögen beantwortet wurden und somit auch nicht alle Fragebögen in die Wertung eingegangen sind. In Gruppe 1 wurden in 18,92% der Fragebögen die Offenen Fragen nicht bearbeitet, in Gruppe 2 waren es 18,18%. Die Studierenden dieser Gruppen dürften bereits bei der Bearbeitung der MC-Fragen ihre Schwierigkeiten gehabt haben und waren mit den Offenen Fragen teilweise überfordert. Daraus resultieren die hohen Abbrecherquoten, wobei mit Frage 23 (Muskuläre Strukturen der Rotatorenmanschette) eine der Offenen Fragen so gestellt war, dass jeder Studierende nach Absolvierung des Ersten Abschnitts der Ärztlichen Prüfung in der Lage sein muss, zumindest Teilantworten zu geben.

Die Offenen Fragen im Teil II des Fragebogens waren als „short-answer-questions“ gestellt worden. Dabei handelt es sich um Fragen, die auf ein sehr spezifisches Wissen abzielen. Der Unterschied zu „Essays“ oder „Modified-Essay-Questions“ liegt darin, dass in diesen wesentlich umfangreichere und komplexere Antworten erwartet werden. Somit lassen sich die „short-answer-questions“ als „mildere“ Prüfungsform bezeichnen, die allerdings durch ihre Spezifität eine gewisse Schwierigkeit beinhaltet.

Der Vergleich der Mittelwerte für die erreichte Gesamtpunktzahl der Offenen Fragen zeigt für die Gruppen 1 und 2 prozentuale Werte von 39,44% bzw. 41,16%. Diese Werte erscheinen im Vergleich mit den Ergebnissen für die MC-Fragen (30,63% bzw. 32,99%) dieser beiden Gruppen relativ hoch und liegen bei einer Differenz von 1,72 Prozentpunkten nur wenig auseinander. Dieses Ergebnis wird in Kapitel 4.4.5 (Vergleich MC-Fragen und Offene Fragen) noch Gegenstand der Diskussion sein.

Gruppe 3-Prä weist mit einem Mittelwert von 32,83% den niedrigsten Wert aller Gruppen für die Offenen Fragen auf. Dieses Ergebnis verwundert, da eigentlich auf Grund der Absolvierung der Einführung in die Orthopädie eine leichte Steigerung gegenüber den Mittelwerten der Gruppen 1 und 2 zu erwarten war. Somit zeigt sich, dass der Wissenszuwachs, der sich in den MC-Fragen (Steigerung um 27,11 Prozentpunkte) deutlich manifestiert hat, im Bereich der Offenen Fragen nicht aktiv reproduzierbar war, sondern sogar ein Rückgang gegenüber den Ergebnissen der Gruppen 1 und 2 auftritt.

Gruppe 4-Prä repräsentiert im Vergleich zu Gruppe 3-Prä das zu erwartende Ergebnis, indem sich mit 47,42% eine leichte Steigerung der Ergebnisse um 6,26 Prozentpunkte gegenüber dem Wert der Gruppe 2 findet. Dieser Wissenszuwachs lässt sich durch das Absolvieren der einwöchigen Einführung in die Orthopädie und durch den Kontakt mit anderen klinischen Fächern und Teilgebieten der Orthopädie erklären. Die geringere prozentuale Steigerung gegenüber den MC-Fragen deutet darauf hin, dass aktives Reproduzieren von Wissen für die Studierenden eine schwierige Aufgabe darstellt.

Die größte Steigerung für die Offenen Fragen liegt zwischen der Prä- und Post-Testung vor. Gruppe 3 kann den größten prozentualen Zuwachs für die Offenen Fragen mit 23,23 Prozentpunkten verzeichnen, Gruppe 4 weist eine Steigerung um 11,91 Prozentpunkte auf. Die erheblich größere Steigerung der Gruppe 3 gegenüber der Gruppe 4 ist vor allem auch durch den niedrigeren Ausgangswert bedingt. Es bleibt somit festzuhalten, dass die Teilnehmer des Orthopädie-Praktikums nicht nur Wissen erworben haben, sondern auch in der Lage waren, dieses Wissen als Antwort auf eine Offene Frage aktiv zu reproduzieren.

Im Anschluss an das Orthopädie-Praktikum erfährt das orthopädische Wissen im Laufe des Studiums einen Rückgang, der sich im Vergleich mit dem Rückgang für die MC-Fragen (Rückgang um 16,11 Prozentpunkte zwischen Gruppe 4-Post und Gruppe 7) etwas leichter darstellt. So fällt der Wert zwischen Gruppe 4-Post und Gruppe 6 um 12,52 Prozentpunkte ab. Dies ist offensichtlich ein Hinweis darauf, dass sich bei einer nicht vorhandenen Auseinandersetzung mit dem Fachbereich Orthopädie das Faktenwissen auf ein Niveau stabilisiert, das mit dem Wissen bei Auseinandersetzung mit dem Fachbereich Orthopädie nicht konkurrieren kann.

Wie bei den Gruppen 3 und 4 findet auch bei der Prä-Post-Testung an der Universität Heidelberg eine Ergebnissteigerung statt, die sich in Gruppe 8 mit einem Wert von 17,78 Prozentpunkten ausdrückt. Somit darf ebenfalls angenommen werden, dass die Teilnehmer des Orthopädie-Praktikums an der Universität Heidelberg nicht nur Wissen erworben haben, sondern auch in der Lage waren, dieses Wissen als Antwort auf eine Offene Frage aktiv zu reproduzieren.

Die Standardabweichung zeigt eine Streuung von 27,63% in Gruppe 7 bis 42,76% in Gruppe 1. Diese hohen Werte sind in der Art und Auswahl der Offenen Fragen begründet, da man entweder die Lösung weiß, oder nicht. Auffällig erscheint, dass die Standardabweichungen in Gruppe 1 und 2 höher ausfallen als in den Gruppen 3-8. Diese allgemeine Tendenz könnte ein Hinweis darauf sein, dass mit Zunahme des Wissens die interpersonellen Wissensunterschiede kleiner werden.

Im Vergleich der Gruppe 1 mit Gruppe 2 bestätigen die ausschließlich über dem Signifikanzniveau liegenden p-Werte die Tatsache, dass im 1. klinischen Semester weder eine orthopädisch relevante Wissensvermittlung noch eine Überschneidung der Lehrinhalte mit den dozierten klinischen Grundlagenfächern stattfindet.

Bei der Analyse von Gruppe 2 mit Gruppe 3-Prä (bereits absolvierte einwöchige Einführung in die Orthopädie) finden sich nur bei Frage 23 ein p-Wert $< 0,05$. Allerdings basiert dieser p-Wert auf dem erheblich schlechteren Abschneiden der Gruppe 3-Prä bei Frage 23. Insgesamt lassen sich bei diesen beiden Gruppen, abgesehen vom auffällig schlechten Ergebnis der Gruppe 3-Prä bei Frage 23, keine statistisch signifikanten Unterschiede erkennen.

Im Vergleich der Prä-Gruppen mit den Post-Gruppen finden sich für Gruppe 3, 4 und 8 ausschließlich p-Werte von $< 0,05$. Anhand dieser Ergebnisse lässt sich ein statistisch signifikanter Unterschied im Vergleich der Prä-Gruppen mit den Post-Gruppen feststellen, resultierend auf dem Lehrerfolg durch das Orthopädie-Praktikum.

Im Vergleich der Gruppe 4-Post mit Gruppe 5 sprechen die P-Werte von jeweils 0,01 für die Fragen 24 und 25 für einen statistisch signifikanten Unterschied zwischen den beiden Gruppen, der in diesem Fall den Wissensrückgang nach Absolvierung des Orthopädie-Praktikums verdeutlicht. Bei der Analyse der Gruppe 5 mit Gruppe 6 und Gruppe 6 mit Gruppe 7 werden nur p-Werte $> 0,05$ erzielt. Diese Ergebnisse bestätigen den leichten Rückgang der Ergebnisse mit dem steigenden Ausbildungsniveau.

Im Folgenden werden die Ergebnisse für die Offenen Fragen im Einzelnen diskutiert:

Frage 23, 24 und 25 stellen so genannte „short-answer-questions“ dar, die in Stichworten innerhalb eines bestimmten Umfangs beantwortet werden sollen.

Dass in **Frage 23 (Muskuläre Strukturen der Rotatorenmanschette)** mit durchschnittlich 80,74% die höchsten prozentualen Werte im Vergleich zu den anderen Offenen Fragen erzielt wurden, verwundert nicht, denn schließlich handelt es sich hierbei um eine Frage aus der Allgemeinen Orthopädie, die auch mit dem Wissen aus anderen vorklinischen (Anatomie) und klinischen Fachbereichen (Chirurgie) zu beantworten war. Dies bestätigt sich dadurch, dass die Studierenden des 1. klinischen und 2. klinischen Semesters durchschnittlich 3,47 bzw. 3,41 Punkte für die Aufzählung der muskulären Strukturen der Rotatorenmanschette nennen können (bei einem Höchstwert von 4 erreichbaren Punkten). Hier ist auffällig, dass in Gruppe 3-Prä der Durchschnitt nur bei 2,64 Punkten liegt, wenn man davon ausgeht, dass dieses Wissen im Anatomie-Kurs in der Vorklinik doziert wurde und ebenso zum ärztlichen Grundwissen zu zählen ist. Es fällt auch ins Auge, dass nach Erreichen der absoluten Höchstwerte in den Post-Gruppen (90,91% bzw. 91,07%) die durchschnittliche Leistung schon im darauf folgenden klinischen Semester stark abfällt (79,64%) und im 6. klinischen Semester schon ein 20%iger Rückgang im Vergleich zum Post-Test zu verzeichnen ist (71,09%). Hier zeigt sich, dass einmalig gelerntes Wissen bei ausbleibender fortführender Anwendung bei einer großen Anzahl von Studierenden offenbar wieder verloren geht.

Frage 24 (4 Ursachen, die zu einem Beckenschiefstand führen) war mit neun möglichen Antworten eine weit gefasste Frage, was jedoch von den Teilnehmern überwiegend nicht genutzt wurde. Wie schon bei Frage 23 handelt es sich ebenfalls um eine Frage aus der Allgemeinen Orthopädie, die auch in anderen klinischen Fachbereichen (Radiologie, Chirurgie) behandelt wird. Das mit durchschnittlich 46,91% der erreichbaren Punkte für diese Frage deutlich schlechtere Abschneiden als bei Frage 23 spiegelt die Tatsache wider, dass im Gegensatz zu Frage 23 nicht nur Abruf und Wiedererkennen von Faktenwissen (Niveau 1), sondern bei Nichtwissen Problemlösendes Denken (Niveau 3) gefordert war.

Es fällt auf, dass vom Großteil der Untersuchungsteilnehmer nicht die spezifisch orthopädischen Kernaussagen „Hüftabduktionskontraktur“, „Hüftadduktionskontraktur“, „Kniebeugekontraktur“ oder „Spitzfuß“ genannt wurden, sondern die Punkte durch die

Aussagen „Fraktur“, „Skoliose“, „Hüftdysplasie“ oder „Beinlängendifferenz“ erzielt wurden. Gerade in den Gruppen 1, 2 und 3-Prä, die mit 28,33%, 33,33% bzw. 28,03% die schlechtesten prozentualen Durchschnittswerte aufzeigen, wären trotz nicht besuchter orthopädischer Lehrveranstaltungen die richtigen Begriffe wie „Fraktur“ und „Beinlängendifferenz“ zu erwarten gewesen. Ob diese schlechten Ergebnisse auf Grund mangelnder Auseinandersetzung mit der Frage und auf Grund fehlenden Problemlösenden Denkens zustande gekommen sind, ist nicht aufzuklären.

Anzeichen für den Erfolg der Lehre durch das Orthopädie-Praktikum liefert der Vergleich der Prä-Gruppen mit den Post-Gruppen. So werden die Höchstwerte für diese Frage nach deutlichen Verbesserungen zum Prä-Test in den Post-Gruppen gefunden (56,82% in Gruppe 3-Post, 65,48% in Gruppe 4-Post bzw. 58,33% in Gruppe 8-Post). Dieses Ergebnis wird durch die statistische Analyse bestätigt (siehe Tabelle 17, S. 90).

Frage 25 (4 Frühzeichen bei einem Morbus Perthes) stellt mit sieben Antwortmöglichkeiten ebenfalls wie Frage 24 eine weit gefasste Frage dar, was von den Untersuchungsteilnehmern auch rege angenommen wurde. So finden sich bei den Antworten zu dieser Frage in einem ausgeglichen Verhältnis sowohl die Kernaussagen wie „Schmerzen in der Hüfte“, „Schmerzen im Knie“, „schmerzhafte Innenrotation“ und „schmerzhafte Abduktion“, als auch die ebenfalls gewerteten Aussagen „Schonhinken“, „Gehschwäche“ und „Drehmann-Zeichen“. Frage 25 repräsentiert im Gegensatz zu den beiden vorherigen Offenen Fragen eine sehr spezifische Frage aus der Speziellen Orthopädie, was den schlechten Durchschnittswert von 11,62% (d.h. im Durchschnitt 0,46 Punkte von 4 möglichen Punkten) erklären kann. Entsprechend wurden die Frühzeichen eines Morbus Perthes im 1. klinischen Semester, im 2. klinischen Semester und in den Prä-Gruppen vor dem Orthopädie-Praktikum nur sehr sporadisch gewusst (3,33% bis 8,33%). Trotz der Ergebnissteigerung um 15,90% (Gruppe 3), 13,10% (Gruppe 4) bzw. 15,83% (Gruppe 8) im Vergleich der Prä-Gruppen mit den Post-Gruppen darf dieses Ergebnis nicht zufrieden stellend sein, da der Morbus Perthes als Krankheitsbild einen Gegenstand im Lehrplan des Orthopädie-Praktikum darstellt und die Ergebnisse der Post-Gruppen mit gerade einmal 20,45% (Gruppe 3-Post), 21,43% (Gruppe 4-Post) bzw. 23,33% (Gruppe 8-Post) sicherlich noch steigerbar sind und auch gesteigert werden müssen.

Zusammenfassend lässt sich für die Auswertung der MC-Fragen und der Offenen Fragen feststellen:

Anhand dieser Untersuchung konnte gezeigt werden, dass der mit den hier verwendeten Messmethoden (MC-Fragen, Offene Fragen) ermittelte fachorthopädische Wissensstand nicht immer mit dem jeweiligen Ausbildungsstand korreliert. Bis zur Absolvierung des Orthopädie-Praktikums entspricht der Wissensstand auch dem Ausbildungsstand, jedoch erfolgt danach ein leichter Rückgang des Wissens, der sich vor Eintritt in das Praktische Jahr (PJ) stabilisiert. Die größte positive Veränderung erfährt das fachorthopädische Wissen nach der Einführung in die Orthopädie und vor allem auch im Laufe des Orthopädie-Praktikums, so dass angenommen werden kann, dass den Studierenden hier ein erheblicher Teil ihres orthopädischen Fachwissens vermittelt wird. Dieses Ergebnis stellt eine Bestätigung für die orthopädische Lehre an der Universität Regensburg und an der Universität Heidelberg dar.

4.4.4) Diskussion der Ergebnisse zu Teil III

Frage 26) Schwierigkeitsgrad der gestellten MC-Fragen

Wie bereits in Kapitel 3.5 dargestellt wurde, verschieben sich die Mehrheiten von den Kategorien „sehr schwierig“ und „schwierig“ in den Gruppen 1, 2 und 3-Prä hin zu der Kategorie „angemessen“ in den Gruppen 3-Post, 4-Prä, 4-Post und 5, bevor sich die Mehrheit in den Gruppen 6 und 7 zurück in Kategorie „schwierig“ verlagert.

Die deutlichen Werte in den Kategorien „sehr schwierig“ und „schwierig“ von 45,95% bzw. 37,84% in Gruppe 1 und die Höchstwerte von 62,81% bzw. 66,67% in der Kategorie „schwierig“ für die Gruppen 2 und 3-Prä sind ein offenkundiger Hinweis darauf, dass die Studierenden überfordert waren.

Sichtbare Unterschiede bestehen in der Einschätzung des Schwierigkeitsgrades zwischen den Gruppen 3-Prä und 4-Prä: Während in Gruppe 3-Prä die MC-Fragen zu 18,18% als „sehr schwierig“ und gerade einmal zu 12,12% als „angemessen“ eingestuft wurden, beträgt der Wert der Kategorie „angemessen“ für Gruppe 4-Prä immerhin 71,43%. Diese Einschätzung korreliert definitiv nicht mit den hier erreichten Mittelwerten

der Richtigantworten, die mit 60,10% für Gruppe 3-Prä und 62,10% für Gruppe 4-Prä relativ dicht zusammen liegen.

In den Gruppen 5, 6 und 7 wandert die Mehrheit von „angemessen“ (50,00% in Gruppe 5) in Richtung „schwierig“ (52,38% in Gruppe 6 bzw. 63,48% in Gruppe 7). Diese Einschätzung folgt dem Verlauf der Mittelwerte der Richtigantworten, die von Gruppe 5 (70,64%) über Gruppe 6 (63,69%) hin zu Gruppe 7 (60,87%) absinken.

Dieses Ergebnis korreliert hinsichtlich des Curriculums auch mit den Ergebnissen zum Wissensstand, bei denen sich ein ähnlicher Verlauf (Anstieg des Wissens gefolgt von Rückgang des Wissens) abzeichnet.

Anhand dieser Beobachtungen lassen sich unterschiedliche Hypothesen aufstellen: Eine Möglichkeit ist, dass der Schwierigkeitsgrad der MC-Fragen weitgehend unabhängig von der eigenen Leistung beurteilt wurde. Andererseits kann es sein, dass die Studierenden ihre Leistung unterbewertet haben. Dass die Einstufung des Schwierigkeitsgrades nicht unbedingt mit der Leistung korreliert, zeigen nicht nur die Werte der Gruppe 3-Prä (18,18% für „sehr schwierig“ und 66,67% für „schwierig“ bei einem Mittelwert von 60,10% Richtigantworten), sondern auch die der Gruppe 8-Post. Diese weist mit 33,33% den mit Abstand größten Anteil der Post-Gruppen für die Antwort „schwierig“ auf, das Ergebnis des Mittelwertes mit 73,06% rangiert aber nur knapp hinter den Werten der Gruppen 3-Post (77,02%) und 4-Post (76,98%).

Frage 27) Schwierigkeitsgrad der gestellten Offenen Fragen

Für die Offenen Fragen zeigt sich bei der Einschätzung des Schwierigkeitsgrades eine Verschiebung der Häufigkeiten von der Kategorie „schwierig“ hin zu „angemessen“. Im Gegensatz zu den MC-Fragen erfolgt hier keine Rückverlagerung zu „schwierig“ in den Gruppen 6 und 7, sondern die Kategorie „angemessen“ bleibt in den Gruppen 3-Post bis 7 mehrheitlich vertreten.

Ebenfalls wie bei den MC-Fragen besteht ein sichtbarer Unterschied bei der Einschätzung zwischen den Gruppen 3-Prä und 4-Prä. So schätzt Gruppe 3-Prä die Offenen Fragen zu 30,30% als „angemessen“ und zu 48,48% als „schwierig“ ein, während bei Gruppe 4-Post sich 57,14% für „angemessen“ und nur 38,10% für „schwierig“ entscheiden. Jedoch korreliert hier im Gegensatz zu den MC-Fragen die Einschätzung auch mit den erreichten Mittelwerten der Richtigantworten, da Gruppe 3-

Prä mit gerade einmal 32,83% das schlechteste Ergebnis aller Gruppen aufweist und Gruppe 4-Prä mit 47,42% das beste Ergebnis der Prä-Gruppen abliefern.

Für die Gruppen 1 und 2 korreliert die Einschätzung mit den Ergebnissen. Beide Gruppen weisen bei der Einschätzung des Schwierigkeitsgrades für die MC-Fragen höhere Werte für die Kategorien „schwierig“ und „sehr schwierig“ auf als für die Einschätzung der Offenen Fragen, was sich auch in den schlechteren Ergebnissen für die MC-Fragen zeigt. In den anderen Gruppen lässt sich keine Tendenz bzw. kein auffälliges Missverhältnis zwischen Einschätzung und Ergebnis feststellen.

Insgesamt zeigt sich, dass die Bearbeitung der MC-Fragen den Untersuchungsteilnehmern mehr Schwierigkeiten bereitet hat als die Beantwortung der Offenen Fragen. So finden sich im Kollektiv aller Gruppen für die MC-Fragen Mittelwerte von 38% für „angemessen“ und 48% für „schwierig“, während auf Seiten der Offenen Fragen 48% für „angemessen“ und nur 36% für „schwierig“ stehen. Diese Einschätzungen stehen im Widerspruch zu den Ergebnissen der Richtiganworten, da abgesehen von Gruppe 1 und 2 in allen anderen Gruppen deutlich bessere Ergebnisse für die MC-Fragen erzielt wurden. Eine mögliche Erklärung für diesen Sachverhalt stellt die Anzahl und die Auswahl der Offenen Fragen dar. Das gute Abschneiden bei Frage 23 in allen Gruppen kann das Vermögen für die Gesamteinschätzung der Offenen Fragen beeinflusst haben. Ebenso treten bei der Bearbeitung von nur drei Offenen Fragen anzahlmäßig weniger Schwierigkeiten und Denkprozesse auf als bei der Auseinandersetzung mit 12 MC-Fragen.

Frage 28) Interesse am Fachbereich Orthopädie

In der Auswertung zeigt sich, dass die Studierenden, die weder die Einführung in die Orthopädie noch das Orthopädie-Praktikum absolviert haben (Gruppen 1 und 2), ihr Interesse für das Fach Orthopädie größtenteils noch nicht einschätzen können. So finden sich in Gruppe 1 jeweils 35,14% für die Antworten „Interesse“ und „unentschieden“, in Gruppe 2 zeigen 38,84% „Interesse“, 14,88% sind noch „unentschieden“ und 25,62% haben „wenig Interesse“.

Nach Besuchen der einwöchigen Einführung in die Orthopädie (Gruppen 3-Prä und 4-Prä) steigt der Anteil derer, die Interesse haben, an (45,45% bzw. 52,38%), erfährt dann aber nach Absolvierung des Orthopädie-Praktikums wieder einen leichten Rückgang

(36,36% bzw. 40,48%) und stabilisiert sich in den Semestern mit höherem Ausbildungsstand (32,17% bis 34,88%).

Im Vergleich der Prä-Gruppen mit den Post-Gruppen fällt in der Kategorie „sehr großes Interesse“ auf, dass in Gruppe 4-Post (26,19%) im Vergleich zu Gruppe 4-Prä (9,52%) das Interesse am Fachbereich Orthopädie bei mehr als einem Viertel der Gruppe stark geweckt wurde. In Gruppe 3 und Gruppe 8 finden sich im Prä-Test und Post-Test konstante Werte für diese Antwort.

Die Mehrheit in Gruppe 6 und Gruppe 7 zeigt „wenig Interesse“ für das Fachgebiet Orthopädie. Der Rückgang des Interesses am Fachgebiet Orthopädie lässt sich damit erklären, dass mit steigendem Ausbildungsstand und somit größerem Erfahrungswert mit den einzelnen Fachgebieten der Medizin die Orientierungsphase der Studierenden hinsichtlich des späteren Berufsziels fortschreitet und eine Entscheidung zu Gunsten anderer Fachgebiete getroffen wird. Diesen Studierenden stehen noch alle Möglichkeiten offen, doch könnte durch bereits absolvierte Famulaturen oder klar abgesteckte Interessensgebiete die Wahl des Schwerpunktfaches bereits stattgefunden haben. Hingegen ist in den Semestern vor und während des Orthopädie-Kursus auf Grund des fehlenden Erfahrungswertes mit den verschiedenen Fachgebieten und der aktuellen Beschäftigung mit dem Fachgebiet Orthopädie das Interesse noch groß.

Frage 28 zielt auf das Allgemeine Interesse am Fachgebiet Orthopädie ab. Das von den Teilnehmern des Orthopädie-Praktikums geäußerte Interesse am Fachbereich Orthopädie repräsentiert sicherlich eine der Vorbedingungen für effektives und sinnvolles Lernen und gibt einen Hinweis auf die intrinsische Motivation der Studierenden. Allerdings muss hier gleichzeitig der Rückkopplungseffekt gesehen werden, da die Auseinandersetzung und Beschäftigung mit einem Fachgebiet wiederum das Interesse an diesem fördert.

Die Untersuchungen zum Interesse am Fachbereich Orthopädie zeigen auf, dass während des Orthopädie-Praktikums und der unmittelbaren Beschäftigung mit dem Fachbereich Orthopädie auch das Interesse dafür am größten ist. Dies spricht für die Umsetzung und Präsentation der Lehrinhalte, das Auftreten der Orthopädie-Dozenten und den Umgang dieser mit den Studierenden. Jedoch muss der Rückgang des Interesses mit steigendem Ausbildungsstand wahrgenommen werden.

Frage 29) Facharzt für Orthopädie/Unfallchirurgie als Berufswunsch

In den Gruppen, die weder die Einführung in die Orthopädie noch das Orthopädie-Praktikum absolviert haben (Gruppen 1 und 2), bescheinigt ein hoher Wert für die Kategorie „vorstellbar“ (29,73% bzw. 36,36%) die Attraktivität des Berufsbildes Orthopäde ohne bisherige Auseinandersetzung mit diesem Fachgebiet.

Die höchsten Werte für Kategorie „strebe Beruf an“ finden sich in den Post-Gruppen nach Absolvieren des Orthopädie-Praktikums (jeweils 9,09% in Gruppe 3-Post und Gruppe 4-Post). In diesen Gruppen treten jedoch niedrigere Werte für die Kategorie „vorstellbar“ auf als in den Prä-Gruppen. Dies ist offensichtlich ein Hinweis darauf, dass infolge der intensiven Auseinandersetzung und Beschäftigung mit dem Fachgebiet Orthopädie bei einigen Studierenden eine Art „Vorauswahl“ hinsichtlich des gewünschten Berufsziels getroffen wurde, sei es pro oder contra.

Bei den Studierenden des 5. klinischen und 6. klinischen Semesters spiegelt sich die These der „Vorentscheidung“ ebenfalls in den Ergebnissen wieder, indem eine Verteilung der Werte für konkrete Aussagen eintritt und die Tendenz zur Mitte („unentschieden“) in den Hintergrund der Ergebnisverteilung tritt. Mit 40,48% bzw. 43,48% wird die Mehrzahl der Studierenden „sicher nicht“ den Beruf als Orthopäde/Unfallchirurg wählen, während sich 21,43% bzw. 27,83% dieses Berufsbild vorstellen können.

Auch während des Orthopädie-Praktikums an der Universität Heidelberg (Gruppe 8) darf von dieser These ausgegangen werden. Im Prä-Test entschieden sich noch 26,66% für die Antwort „unentschieden“, während im Post-Test sich nach der intensiven Auseinandersetzung und Beschäftigung mit dem Fachgebiet Orthopädie bereits eine Verteilung der Ergebnisse hinsichtlich einer Vorentscheidung abzeichnet. So können sich 20,00% den Beruf als Orthopäde/Unfallchirurg vorstellen, 40,00% können dies nicht und nur noch 6,67% sind „unentschieden“.

Frage 29 zielt auf das Spezielle Interesse am Fachgebiet Orthopädie ab. Das von den Teilnehmern des Orthopädie-Praktikums geäußerte Interesse am Berufsbild Orthopäde/Unfallchirurg repräsentiert sicherlich eine der Vorbedingungen für effektives und sinnvolles Lernen und gibt einen Hinweis auf die intrinsische Motivation der Studierenden. Allerdings muss hier gleichzeitig der Rückkopplungseffekt gesehen werden, da die Auseinandersetzung und Beschäftigung mit einem Fachgebiet wiederum das Interesse an diesem fördert.

Frage 30) Befürwortung einer Evaluation zum Thema Studienordnung und Fachwissenstand

Insgesamt haben sich von allen Untersuchungsteilnehmern 90,00% für eine Evaluation zum Thema Studienordnung und Fachwissenstand ausgesprochen, dem gegenüber stehen 10,00% der Studierenden, die dies ablehnen. Die höchsten prozentualen Werte für die Ablehnung einer Evaluation finden sich an der Universität Regensburg in den Semestern mit dem niedrigsten Ausbildungsstand (27,03% bzw. 14,05%). Dies lässt sich offensichtlich mit den fehlenden Erfahrungswerten im Studium, mit der bisher fehlenden Auseinandersetzung und Beschäftigung mit dem Fachgebiet Orthopädie und der daraus resultierenden deutlichen Überforderung bei den Fragen zum Wissensstand erklären. Die hohen Werte der Befürwortung dieser Erhebung sind ebenso erfreulich wie auf die nochmals hingewiesene Rücklaufquote von 100%.

Bei Ablehnung der Evaluation gab es bei dieser Frage die Möglichkeit, in Form eines Zusatzes von „warum“ mit vier vorgedruckten Zeilen, einen Kommentar dazu abzugeben. Von dieser Möglichkeit, die Erhebung und den Fragebogen zu kommentieren, wurde jedoch wenig Gebrauch gemacht. Trotzdem erscheint es lohnend, die abgegebenen Kommentare an dieser Stelle zu diskutieren. Es zeigt sich, dass die Kommentare von Gruppe zu Gruppe sehr unterschiedlich, innerhalb der Gruppe aber inhaltlich geschlossen ausfielen. In den Gruppen 1 und 2 fanden sich Äußerungen wie „unsinnig, weil noch keine Ahnung von Orthopädie“ oder „eigentlich ja, aber bitte erst nach dem Kurs Orthopädie“. Dieser unangekündigte Test schien die Teilnehmer zu verunsichern. In den Gruppen 3, 4 und 8 wurden auch bei einer Befürwortung der Evaluation teilweise Kommentare abgegeben, wie z.B. „gut als Lernzielkontrolle“. Auch auf den Schwierigkeitsgrad der Fragen wurde Bezug genommen: „Offene Fragen sind ungewohnt und somit schwieriger“. In den Gruppen 6 und 7 äußerte sich in Form von Kommentaren wie „nicht nach Klausuren, da hat man keine Lust mehr“, „unsinnig“ und „sinnlos“ Unmut über die Befragung.

Die Akzeptanz der Befragung war sehr erfreulich. Wenngleich auch positive Kritik geäußert wurde, so verdeutlichen die teilweise von Unverständnis und Ablehnung gekennzeichneten Kommentare, wie ungewohnt und verunsichernd eine solche Untersuchung vor allem zum Wissensstand für die Befragten war. Ebenfalls muss berücksichtigt werden, dass die Durchführung der Befragung größtenteils nach Klausuren oder am Rande von Pflichtveranstaltungen stattfand. Die positiven Reaktionen sollen zu weiteren Untersuchungen auf diesem Gebiet ermuntern.

4.4.5) Diskussion der Ergebnisse: Vergleich MC-Fragen und Offene Fragen

Bei der Diskussion zum Vergleich der Ergebnisse für MC-Fragen und Offene Fragen muss beachtet werden, dass trotz der jeweils zu erreichenden Gesamtpunktzahl von 12 Punkten für jeden Fragenteil die beiden Fragenteile sich in Art der Bewertung und Umfang der Fragen unterscheiden und sich die Ergebnisse somit nicht direkt vergleichen lassen.

Hinsichtlich des Gesamtergebnisses liegen die Offenen Fragen mit einem Durchschnittswert aller Gruppen von 47,11% deutlich unter den Werten für die MC-Fragen mit einem Durchschnittswert von 60,13%. Dabei fällt jedoch auf, dass die MC-Fragen sowohl den niedrigsten pro Gruppe erreichten Mittelwert (30,63% in Gruppe 1) als auch den höchsten pro Gruppe erreichten Mittelwert (77,02% in Gruppe 3-Post) aufweisen, während die Offenen Fragen mit Werten von 32,83% (Gruppe 3-Prä) bis 59,33% (Gruppe 4-Post) zwischen den Extremwerten der MC-Fragen liegen (siehe Diagramm 15, S. 94). Nur im 1. klinischen und 2. klinischen Semester werden bessere Ergebnisse für die Offenen Fragen erreicht, was durch das gute Abschneiden bei Frage 23 zu erklären ist, in der auch ohne bisherige Auseinandersetzung und Beschäftigung mit dem Fachgebiet Orthopädie allein mit dem Wissen aus der Vorklinik und basierend auf grundlegenden Denkprozessen eine gewisse Punktzahl erreicht werden konnte. Für alle anderen Gruppen sind die Ergebnisse für die MC-Fragen besser als für die Offenen Fragen. Zu diesem Ergebnis kamen auch Newble et al. (55) und McCloskey und Holland (49), die als Grund den Hinweiseffekt und die 20%ige Ratewahrscheinlichkeit der Richtigerantwort von MC-Fragen anführen. In diese Begründung muss jedoch auch die Tatsache aufgenommen werden, dass das aktive Reproduzieren von Faktenwissen, wie es bei den Offenen Fragen gefordert wird, für die Studierenden eine größere Schwierigkeit darstellt als das Ankreuzen von vorgegebenen Antwortmöglichkeiten. Newble (55) konnte in seiner Studie aufzeigen, dass MC-Prüfungen das Wissen der Kandidaten überschätzen. Ob dies bei den vorliegenden Ergebnissen auch der Fall ist, lässt sich anhand der durchgeführten Untersuchung jedoch nicht aufzeigen. An dieser Stelle wäre eine zusätzliche Untersuchung erforderlich, die sich mit der Frage nach der kriterienbezogenen Validität (criterion validity) der beiden gestellten Fragemethoden beschäftigt. Da in der Prä-Post-Testung für beide Fragetypen die positiven Veränderungen des Fachwissens in ähnlichem Ausmaß zeigen, sind beide Testverfahren in der Lage, diese Wissensveränderungen zu messen.

Es wird die Tendenz deutlich, dass nach der Auseinandersetzung und Beschäftigung mit dem Fachgebiet Orthopädie in MC-Fragen mit weniger Wissen mehr Punkte zu erreichen sind als in Offenen Fragen. Dabei liegt die Erklärung nahe, dass MC-Fragen im Gegensatz zu Offenen Fragen vorgegebene Antwortmöglichkeiten enthalten und dass pro Frage eine Ratewahrscheinlichkeit von 20% für eine Richtigerantwort besteht.

Es muss ebenfalls darauf hingewiesen werden, dass die Studierenden durch die Art der vorgegebenen Staatsprüfungen auf die Beantwortung von MC-Fragen ausgerichtet sind und auch ihr Lernverhalten dementsprechend gestalten. Der Fragentyp der Offenen Fragen stellt eine sehr sinnvolle Wissensabfragemöglichkeit dar. Jedoch sind Offene Fragen für die Studierenden weitgehend ungewohnt, da die Studierenden im Laufe des Studiums kaum mit diesen konfrontiert werden, und auch eher im vorklinischen als im klinischen Studienabschnitt.

Einen wichtigen Einflussfaktor stellt sicherlich auch die Akzeptanz einer Prüfungsform dar, denn diese kann sich in Form einer unvollständigen Bearbeitung der Fragen äußern. In diesen Sachverhalt spielt auch die Tatsache hinein, dass das aktive Reproduzieren von Wissen einen aufwendigeren Prozess darstellt als die Auswahl vorformulierter Antwortmöglichkeiten. Hinzu kommt, dass den Offenen Fragen von den Studierenden weniger Bedeutung beigemessen wird. Wie bei einer fehlenden Akzeptanz einer Prüfungsform kann sich eine geringe zugemessene Bedeutung an einer lückenhaften und unvollständigen Bearbeitung dieser Fragen zeigen, wie es in der Korrektur vorgefunden wurde und sich von einem Auslassen der Frage mangels Wissen abgrenzen ließ.

Bei der Analyse der Veränderungen der Ergebnisse von Gruppe zu Gruppe fällt auf, dass für die Ergebnissteigerungen der Prä-Gruppen (Gruppen 3-Prä, 4-Prä und 8-Prä) zu den jeweiligen Post-Gruppen (Gruppen 3-Post, 4-Post und 8-Post) Werte gefunden wurden, die relativ dicht beieinander liegen und, abgesehen von der Ergebnissteigerung von Gruppe 3-Prä gegenüber Gruppe 2 bei den MC-Fragen, die höchsten positiven Veränderungen darstellen. Somit darf angenommen werden, dass die Studierenden während des Orthopädie-Praktikums einen erheblichen Teil ihres orthopädischen Fachwissens erwerben. Bei der Prä-Post-Testung haben sich für die MC-Fragen Ergebnissteigerungen um 14,88% bis 19,73% ergeben, bei den Offenen Fragen erfolgten positive Veränderungen im Bereich von 11,92% bis 23,23%. Dies bedeutet, dass sich der Wissenszuwachs in beiden Testformen (MC-Fragen und Offene Fragen)

annähernd gleich auswirkt, so dass beide Testformen geeignet erscheinen, diese Veränderungen zu messen.

Im Hinblick auf die Prüfungsmethoden steht fest, dass Offene Fragen den Wissensstand im Verlauf der medizinischen Ausbildung erfassen können. Der Vergleich der Ergebnisse der durchgeführten Studie mit der Literatur weist darauf hin, dass Offene Fragen einige Aspekte ärztlicher Kompetenz besser erfassen können als die MC-Fragen. Somit stellen Offenen Fragen grundsätzlich eine sinnvolle Ergänzung von MC-Prüfungen dar.

Wie in Kapitel 1.1.6 (siehe S. 31 ff) bereits ausführlich dargestellt wurde, weist die Literatur zum Thema „MC-Fragen versus Offene Fragen“ in die Richtung, dass Offene Fragen eine höhere Validität (construct validity) aufweisen als die MC-Fragen. Ein weiterer Vorteil der Offenen Fragen ist, dass sie nicht durch Hinweiseffekte und Ratewahrscheinlichkeit zur Verfälschung der Ergebnisse beitragen. Dadurch ergibt sich die Überlegung, ob der Preis eines höheren zeitlichen und finanziellen Aufwandes von diesen Vorteilen nicht aufgewogen werden kann. Der logische Rückschluss, eine Kombination der Anwendung dieser Fragetypen anzuwenden, wird gestützt durch die Ergebnisse derjenigen Studien, die den Einsatz mehrerer Testverfahren in einer Prüfung befürworten.

Bei Betrachtung der Ergebnisse aller Studien zum Thema „Vergleich verschiedener Prüfungsmethoden“ muss immer die folgende Fragestellung miteinbezogen werden:

Was macht ärztliche Kompetenz in all ihren Aspekten aus und auf welche Art sind diese Qualifikationen zu definieren und zu erfassen?

Somit sind weitere Untersuchungen zur Validität erforderlich, die sich vor allem mit der Frage nach der „construct validity“ beschäftigen. Bisher wurden einige Bestrebungen durchgeführt, die einen wichtigen Schritt in diese Richtung darstellen, indem sie auf der Grundlage einer Bestimmung ärztlicher Qualifikationen die entsprechenden Ausbildungsziele definieren. Ausgehend von diesen Ausbildungszielen erfolgt dann die Entwicklung adäquater Prüfungssysteme, die im Sinne einer Ergebnisevaluation die Umsetzung der Ausbildungsziele für Lehrende und Lernende sicherstellen sollen (19, 21, 22, 89). Jedoch konnten auch in diesen Studien bisher keine einheitlichen Ergebnisse erzielt werden.

4.4.6) Diskussion der Ergebnisse für die MC-Fragen im Einzelnen

Die Auswahl und Darstellung der Ergebnisse der einzelnen MC-Fragen erfolgte nach bestimmten inhaltlichen Kriterien (siehe Kapitel 2.2.2, S. 55 f), wodurch sich eine Einteilung in 5 Untergruppen (a bis e) ergibt.

a) Fragen zu typisch orthopädischen Symptomen

Es war zu erwarten, dass in Gruppe 1 und Gruppe 2 die schlechtesten Ergebnisse abgeliefert werden, da z.B. die Begriffe „Hammerfinger“ und „Hallux rigidus“ keinen Lehrgegenstand in vorklinischen Fächern oder im 1. klinischen Semester darstellen und als nicht bekannt angenommen werden dürfen. So verwundert es nicht, dass Gruppe 1 und Gruppe 2 für diese beiden Fragen teilweise auch Werte unterhalb der 20%igen Ratewahrscheinlichkeit aufweisen. Die Ergebnissteigerungen bei der Prä-Post-Testung (Gruppen 3, 4 und 8) um 28,79% (Gruppe 3), 30,36% (Gruppe 4) und 31,67% (Gruppe 8) geben einen deutlichen Hinweis darauf, dass während des Orthopädie-Praktikums die Wissensvermittlung stark auf typisch orthopädische Symptome abzielt.

b) Fragen zur Statik

Auffällig ist hier, dass MC-Frage Frage 12 (funktionelle Beinverkürzung) nur in den Gruppen 3, 4 und 5 besser beantwortet wurde, als die Offene Frage 24. Alle anderen Gruppen schnitten bei Frage 24 etwas besser ab. Als Erklärung darf hier angenommen werden, dass das Thema „funktionelle Beinverkürzung“ wohl Gegenstand der einwöchigen Einführung in die Orthopädie und des Orthopädie-Praktikums ist und dieser Wissenserwerb von den meisten Studierenden nach einem Semester wieder in Vergessenheit gerät. Die rückläufige Tendenz kann darauf hinweisen, dass dieses spezifische Wissen sich nicht stabilisiert. Dieser Sachverhalt stützt auch die Behauptung, dass die Wissensvermittlung der Blockpraktika größtenteils auf das Kurzzeitgedächtnis abzielt. Im Vergleich der Gruppe 8 (Universität Heidelberg) mit den Gruppen 3 und 4 (Universität Regensburg) fällt auf, dass sowohl im Prä-Test (32,09% gegenüber 48,87% bzw. 69,35%) als auch im Post-Test (52,50% gegenüber 69,32% bzw. 76,69%) die Gruppe 8 deutlich schlechter abschneidet. Somit darf angenommen werden, dass

an der Universität Regensburg eine grundlegendere Wissensvermittlung im Themengebiet „Statik“ stattfindet.

c) Fragen zur orthopädischen Diagnostik / zu orthopädischen Zeichen in der Diagnostik

Frage 14 (Alters-Osteoporose) stellt die Frage mit den schlechtesten Ergebnissen aller MC-Fragen dar. Die Überforderung aller Gruppen mit dieser Frage wird sowohl durch die Ergebnisse der Gruppen 1 und 2 verdeutlicht, die mit 8,11% bzw. 7,44% sogar weit unter der 20%igen Ratewahrscheinlichkeit liegen, als auch durch die Ergebnisse der Post-Gruppen (28,57% bis 45,45%). Auf Grund des guten Abschneidens bei den anderen Fragen zur orthopädischen Diagnostik ergibt sich insgesamt mit einem Mittelwert von 64,52% aller Gruppen noch ein zufriedenstellendes Ergebnis. Jedoch darf hiermit angeregt werden, das Thema „Osteoporose“ verstärkt in den Lehrplan aufzunehmen, da es auch häufig im medizinischen Alltag und in der medizinischen Praxis vertreten ist.

d) Fragen zu spezifisch-orthopädischen Krankheitsbildern

In dieser Fragegruppe bereitete besonders Frage 25 den Studierenden große Schwierigkeiten. Alle Gruppen erreichten hier weniger als 25% Richtigantworten. Auch für Frage 20, die ebenfalls das Krankheitsbild „Morbus Perthes“ behandelt, ergeben sich nur durchschnittliche Werte. Dieses schlechte Abschneiden lässt sich nicht erklären, da der „Morbus Perthes“ sowohl in der einwöchigen Einführung in die Orthopädie, als auch im Orthopädie-Praktikum Gegenstand des Lehrkataloges ist. Dabei fällt auf, dass bei Frage 20 die Antwort „B“ die mit Abstand häufigste Falschantwort repräsentiert. Als Distraktor fungierte hier das Krankheitsbild „Epiphysiolysis capitis femoris“, das jedoch nicht für ein fünfjähriges Mädchen mit Laufunlust und rascher Ermüdbarkeit typisch ist.

Das erheblich schlechtere Abschneiden von Gruppe 2 bei Frage 21 (7,44%) gegenüber Gruppe 1 (21,62%) lässt sich nicht anhand des Curriculums erklären. Auch das allgemein schlechte Abschneiden der Gruppen 1 und 2 bei dieser Frage war nicht unbedingt zu erwarten, da das Thema „Skoliose“ auch in der Vorklinik Gegenstand des Kursus der Anatomie ist. Einzig allein die Spezifizierung der Frage auf die „funktionelle Skoliose“ könnte das schlechte Ergebnis der Gruppen 1 und 2 rechtfertigen.

e) Frage zu anatomischen Begriffen

Es verwundert nicht, dass im Vergleich zu den anderen Offenen Fragen bei dieser Frage in den Gruppen 1 und 2 überdurchschnittlich gute Ergebnisse erzielt wurden. Der Begriff „Rotatorenmanschette“ darf nach Absolvierung des Kursus der Anatomie als bekannt vorausgesetzt werden. Allein in Gruppe 3-Prä findet sich mit 65,91% gegenüber einem Durchschnittswert aller Gruppen von 80,47% ein niedrigerer prozentualer Wert. Dieses Ergebnis ist nicht greifbar, da die muskulären Strukturen der Rotatorenmanschette im Lehrplan der Vorlesung zu „Schulter und obere Extremität“ einen festen Platz einnehmen und auch aus der Vorklinik noch als bekannt vorausgesetzt werden dürfen.

Bei Betrachtung der Mittelwerte aller Gruppen der Ergebnisse zum Wissensstand nach inhaltlichen Kriterien (siehe Diagramm 27, S. 112) lässt sich der Eindruck gewinnen, dass die Fragengruppe „anatomische Begriffe“ mit Abstand am besten bearbeitet wurde. Diese Einschätzung muss allerdings unter der Berücksichtigung erfolgen, dass bei dieser Fragengruppe nur eine Frage (Frage 23) zur Beantwortung vorhanden war, die als grundlegend für alle Studierenden jeden Ausbildungsstandes angesehen werden darf. Diese Voraussetzung wird durch den erreichten Durchschnittswert von 80,47% Richtigantworten definitiv bestätigt und muss auch erst einmal erreicht werden.

Der Durchschnittswert von 64,52% in der Fragengruppe „orthopädische Diagnostik“ lässt sich damit erklären, dass das Themengebiet Diagnostik nicht nur in der Orthopädie von großer Bedeutung ist, sondern auch in den anderen Fachgebieten, sei es im vorklinischen Abschnitt (Kursus Anatomie, Physiologie) als auch im klinischen Abschnitt (Radiologie, Chirurgie, ...), vermehrt Anwendung findet.

Die Durchschnittswerte für die Fragengruppen „orthopädische Symptome“ (54,07%) und „Statik“ (51,20%) entsprechen den Erwartungen. Die zugehörigen Fragen konnten teilweise auch ohne spezielles orthopädisches Wissen beantwortet werden, da einige Fragen auf Grundwissen oder statischen Grundregeln basieren und somit auch durch Herleiten zu lösen sind.

Die schlechtesten Durchschnittswerte für die Fragengruppe „spezifisch-orthopädische Krankheitsbilder“ liegen bereits in der Bezeichnung der Fragengruppe begründet, da in dieser Kategorie fast alle Fragen nur mit dem speziellen orthopädischen Wissen nach Absolvierung eines orthopädischen Kursus zu beantworten waren.

4.4.7) Diskussion der Ergebnisse: Vergleich Prä-Test und Post-Test der unterschiedlichen Universitäten

Für diesen Vergleich muss vorab angemerkt werden:

Trotz der ähnlichen Gruppengrößen von $n = 33$ (Gruppe 3), $n = 42$ (Gruppe 4) und $n = 30$ (Gruppe 8) gibt es auch diskussionsrelevante Unterschiede zwischen den Gruppen der Universität Regensburg und der Gruppe der Universität Heidelberg. So befinden sich in den Gruppen 3 (3. klinisches Semester) und 4 (4. klinisches Semester) der Universität Regensburg die jeweiligen Gruppenmitglieder geschlossen in einem Semester und somit auf dem gleichen Ausbildungsniveau, während in Gruppe 8 der Universität Heidelberg sich Gruppenmitglieder aus verschiedenen Semestern befinden (2. klinisches bis 6. klinisches Semester). Ebenfalls muss berücksichtigt werden, dass an der Universität Heidelberg (Gruppe 8) auf Grund des Modellstudiengangs „Heicumed“ das Orthopädie-Praktikum nach einem anderen Lehrsystem durchgeführt wird.

Bei den **MC-Fragen** fällt auf, dass die Gruppe der Universität Heidelberg im Prä-Test schlechter abschneidet als die Gruppen der Universität Regensburg. Das im Schnitt um 7,33% schlechtere Ergebnis lässt sich damit erklären, dass die Gruppen der Universität Regensburg zu Beginn des Semesters bereits eine einwöchige Einführung in die Orthopädie absolviert haben und somit ein grundlegendes Verständnis vom Fachgebiet Orthopädie erlangt haben.

Hinsichtlich der Ergebnissteigerung weist die Gruppe der Universität Heidelberg mit 19,73% gegenüber 16,92% bzw. 14,88% der Gruppen der Universität Regensburg einen höheren Wissenszuwachs auf. Dieser beruht jedoch auch auf einem niedrigeren Ausgangsniveau. Für die Post-Werte nach Absolvierung des Orthopädie-Praktikums finden sich in allen Gruppen ähnliche Endwerte, wobei die Gruppen der Universität Regensburg mit 77,02% und 76,98% etwas besser abschneiden als die Gruppe der Universität Heidelberg mit 73,06%.

Bei der Betrachtung der Ergebnisse nach inhaltlichen Kriterien lassen sich im interuniversitären Vergleich nach Absolvierung des Orthopädie-Praktikums folgende Tendenzen feststellen:

Bei den Fragen zu typisch orthopädischen Symptomen finden sich insgesamt leicht bessere Werte für die Gruppe der Universität Heidelberg. Deutlich besser allerdings schneiden die Gruppen der Universität Regensburg bei den Fragen zur Statik und zur orthopädischen Diagnostik ab. Bei den Fragen zu spezifisch orthopädischen Krankheitsbildern und zu anatomischen Begriffen lassen sich keine Tendenzen feststellen, da alle drei Gruppen bei diesen Fragen sehr ähnliche Werte aufweisen.

Im Folgenden wird eine detaillierte Betrachtung der einzelnen MC-Fragen vorgenommen, um anhand von orthopädischen Schlagworten Unterschiede zwischen den Lehrplänen der Universitäten Regensburg und Heidelberg aufzuzeigen:

Bei Frage 11 weist Gruppe 8 (Universität Heidelberg) einen auffällig hohen Wert im Prä-Test (60,00%) gegenüber den Gruppen 3 (21,21%) und 4 (16,67%) auf. Da bei einem solchen Ergebnis nicht von einer erstaunlichen Ratewahrscheinlichkeit ausgegangen werden kann, lässt sich dies nur damit erklären, dass der Begriff „Hammerfinger“ entweder bereits in einem Vorbereitungskurs oder in einem anderen Fachgebiet abgehandelt wurde, oder sich die Studenten der Universität Heidelberg bereits vor dem Orthopädie-Praktikum explizit auf derartige Themengebiete vorbereitet haben.

Die Gruppen der Universität Regensburg erreichen bei Frage 12 im Post-Test Werte von 81,82% bzw. 88,10%, während die Gruppe der Universität Heidelberg mit 46,67% deutlich schlechter liegt. Da sich auch im Prä-Test ein noch deutlicheres Missverhältnis zu Gunsten der Gruppen der Universität Regensburg zeigt, muss davon ausgegangen werden, dass der Begriff „funktionelle Beinverkürzung“ an der Universität Heidelberg weder in der Vorklinik, noch im Orthopädie-Praktikum Ansprache findet.

Bei Frage 13 („Hallux rigidus“) und bei Frage 17 („Scheiben-Meniskus“) finden sich im Post-Test leicht höhere Werte für die Gruppe der Universität Heidelberg, was o.g. Trend bestätigt, dass die Universität Heidelberg bei den Fragen zu typisch orthopädischen Symptomen etwas besser abschneidet.

Für die Fragen 14, 15 und 16 lassen sich keine auffälligen Unterschiede feststellen.

Deutlich bessere prozentuale Werte im interuniversitären Vergleich erreichen die Gruppen der Universität Regensburg bei Frage 18 („klinische Untersuchungen“). Dieses Ergebnis (84,85% bzw. 80,95% gegenüber 63,33%) lässt sich nicht erklären, da diese Frage auf medizinisches Grundwissen abzielt und auch ohne jegliches orthopädisches Fachwissen zu beantworten ist.

Für Frage 19 finden sich weder auffälligen Unterschiede noch Tendenzen.

Gruppe 8 (Universität Heidelberg) verzeichnet bei Frage 20 sowohl im Prä-Test als auch im Post-Test bessere Ergebnisse. Dies deutet darauf hin, dass das Krankheitsbild „Morbus Perthes“ im Lehrplan des Orthopädie-Praktikums an der Universität Heidelberg fester Bestandteil ist und auch sehr detailliert doziert wird.

Die höheren prozentualen Werte für Frage 21 („funktionelle Skoliose“) und Frage 22 („Morbus Scheuermann“) der Gruppen der Universität Regensburg sprechen für eine fundierte und grundlegende Lehre hinsichtlich spezifisch orthopädischer Krankheitsbilder. Jedoch sind die Werte der Gruppe der Universität Heidelberg für diese beiden Fragen ebenfalls befriedigend und zusammen mit dem guten Abschneiden bei Frage 20 darf ebenfalls von einer guten Vermittlung der Inhalte zu spezifisch orthopädischen Krankheitsbildern ausgegangen werden.

Bei den **Offenen Fragen** findet sich die größte Ergebnissteigerung mit 23,23% in Gruppe 3, die jedoch im Vergleich mit Gruppe 4 (11,92%) und Gruppe 8 (17,78%) auf einem erheblich niedrigeren Ausgangsniveau basiert. Für die Post-Werte nach Absolvierung des Orthopädie-Praktikums finden sich in allen Gruppen ähnliche Endwerte (56,02% bis 59,33%). Insgesamt variieren hier die Werte zwischen den drei Gruppen stärker, so dass sich das Feststellen einer Tendenz im interuniversitären Vergleich als nicht möglich gestaltet. Auch bei der detaillierten Betrachtung der einzelnen Offenen Fragen lassen sich keine auffälligen Unterschiede erkennen, wodurch sich auch kein Rückschluss auf die unterschiedlichen Lehrpläne der Universitäten Regensburg und Heidelberg ziehen lässt.

4.5) Diskussion der Ergebnisse von Prä- und Posttest im Blockpraktikum

Orthopädie

An dieser Stelle sollen die in Kapitel 3.6 (siehe S. 103 f) dargestellten Ergebnisse für die Prä- und Post-Testung der Teilnehmer des Orthopädie-Praktikums erörtert werden.

Das Resultat, dass bei der Beantwortung der MC-Fragen höhere prozentuale Werte erzielt wurden als bei den Offenen Fragen, entspricht den Erwartungen und lässt sich anhand folgender Sachverhalte erklären: MC-Fragen enthalten im Gegensatz zu Offenen Fragen vorgegebene Antwortmöglichkeiten und es besteht pro Frage eine Ratewahrscheinlichkeit von 20% für eine Richtigantwort.

Beim Vergleich der Gruppen fällt auf, dass Gruppe 3 und Gruppe 8 sowohl für die MC-Fragen als auch für die Offenen Fragen größere positive Veränderungen erzielen als Gruppe 4. Dies könnte darauf hinweisen, dass die Gruppen 3 und 8 besser auf die Post-Testung vorbereitet waren. Diese These ist jedoch für den Vergleich von Gruppe 4 mit Gruppe 3 auszuschließen, da beide Gruppen an der Universität Regensburg nach dem gleichen Themenkatalog unterrichtet wurden. Hinsichtlich des Vergleichs von Gruppe 4 mit Gruppe 8 wäre es sinnvoll, den Themenkatalog des Orthopädie-Praktikums beider Universitäten zu vergleichen. Da der Unterschied zwischen Gruppe 4 und Gruppe 3 deutlicher ausfällt als zwischen Gruppe 4 und 8, liegen für die niedrigeren positiven Veränderungen der Gruppe 4 vielmehr folgende Erklärungen nahe:

Da Gruppe 4 bei den MC-Fragen und vor allem bei den Offenen Fragen im Prä-Test höhere Ausgangswerte aufweist, stellt es für Gruppe 4 eine größere Schwierigkeit dar, eine annähernd gleiche prozentuale Ergebnissteigerung zu erzielen wie diejenige in Gruppe 3 und Gruppe 8.

Ebenfalls müssen an dieser Stelle die unterschiedlichen Bedingungen der Befragung in den drei Gruppen berücksichtigt werden: Während in Gruppe 3 der Post-Test am ersten Wochentag nach Absolvierung des Orthopädie-Praktikums am Rande eines Seminars und in Gruppe 8 am letzten Tag des Orthopädie-Praktikums erfolgte, fand in Gruppe 4 der Post-Test aus organisatorischen Gründen eine Woche nach Absolvierung des Orthopädie-Praktikums am Nachmittag nach einer Klausur statt. Somit kann davon ausgegangen werden, dass die Studierenden zu diesem Zeitpunkt bereits Einbußen an Konzentration und auch an Mitarbeitsbereitschaft aufgewiesen haben.

Insgesamt gesehen entsprechen die Ergebnisse für die MC-Fragen an den Universitäten Regensburg und Heidelberg den Erwartungen einer Leistung nach Auseinandersetzung und Beschäftigung mit einem Fachgebiet und weisen in ihrer Gesamtheit nur geringe Unterschiede auf. Da der Schwierigkeitsgrad der gestellten MC-Fragen dem Niveau der Fragen im Zweiten Abschnitt der Ärztlichen Prüfung nach neuer AO entspricht und die Werte im Post-Test der drei Gruppen zwischen 73,06% und 77,02% liegen, darf von einem zufrieden stellenden Ergebnis ausgegangen werden. In Bezug auf den Vergleich von Prä-Test und Post-Test lässt sich ein deutlicher Wissenszuwachs feststellen. Somit darf angenommen werden, dass den Studierenden an beiden Universitäten ein erheblicher Teil ihres orthopädischen Faktenwissens im Laufe des Orthopädie-Praktikums vermittelt wird.

Hinsichtlich der Offenen Fragen müssen die Ergebnisse als nicht zufrieden stellend angesehen werden. Auch unter Berücksichtigung der ungewohnten Prüfungsmethode und der Schwierigkeit des aktiven Reproduzierens von Wissen weisen vor allem die Werte im Post-Test (56,06% bis 59,33%) kein ausreichendes Wissen auf. Für einen Test nach direkter Auseinandersetzung und Beschäftigung mit einem Fachgebiet und auch hinsichtlich des Schwierigkeitsgrades der gestellten Offenen Fragen wären Ergebnisse im Post-Test von > 60% erwartungsgemäß und auch wünschenswert gewesen. In Bezug auf den Vergleich von Prä-Test und Post-Test lässt sich jedoch wie auch bei den MC-Fragen ein deutlicher Wissenszuwachs feststellen. Somit darf auch in Bezug auf die Offenen Fragen angenommen werden, dass den Studierenden an beiden Universitäten ein erheblicher Teil ihres orthopädischen Faktenwissens im Laufe des Orthopädie-Praktikums vermittelt wird.

Anhand der Prä-Post-Testung konnte ein definitiver Wissenszuwachs festgestellt werden. Diese Lernzielkontrolle am Ende des Praktikums stellt eine sinnvolle Rückmeldung für Lehrende und Lernende dar, vor allem auch hinsichtlich der Umsetzung und des Erwerbs von Ausbildungsinhalten. Daher ergibt sich die Überlegung, ob sie nicht grundsätzlich erfolgen sollte anstatt nur im Rahmen von speziellen Studien.

Jedoch sind zu diesem Thema noch weitere Untersuchungen erforderlich. Um einen direkten Vergleich zwischen MC-Fragen und Offenen Fragen zu ermöglichen, müssen identische Tests angewandt werden, die mit der „matched-pairs“-Technik entwickelt wurden. Um den sukzessiven Wissenszuwachs bzw. Wissensstand an einem identischen Untersuchungskollektiv über Jahre zu beobachten, könnte eine Langzeitstudie in Betracht gezogen werden. Einen Schritt weiter geht der Wunsch, die Veränderungen des Wissens auch nach dem Studium weiter beobachten zu können. Ein solches Vorhaben setzt eine ausreichende Akzeptanz der Befragung bei den Untersuchungsteilnehmern voraus.

Zum Schluss muss nochmals auf folgendes hingewiesen werden:

Die durchgeführte Untersuchung repräsentiert einen punktuellen Zustand und spiegelt nur ausschnittsweise das orthopädische Wissen in den verschiedenen Gruppen wider. Um die ermittelten und dargestellten Ergebnisse zu festigen, sind umfangreichere Untersuchungen während eines längeren Zeitraums nötig. Ebenso kann ein Fragebogen mit einem Umfang von 12 MC-Fragen und drei Offenen Fragen den Wissensstand auf unterschiedlichen Ausbildungsniveaus nur ansatzweise erfassen. Ein erweiterter Fragebogen nimmt allerdings mehr Zeit für die Bearbeitung in Anspruch.

Um bei einer Untersuchung dieser Art eine korrekte Durchführung und eine ausreichende Teilnahme zu garantieren, ist eine hinreichende Akzeptanz der jeweiligen Ausbildungsinstitute und beteiligten Personen unbedingt erforderlich. Voraussetzung dafür ist ein entsprechendes Bewusstsein, dass eine sorgfältige Evaluation der Evaluationsmethoden zur Sicherung der Qualität des Prüfungswesens unabdingbar ist.

4.6) Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der Zufriedenheit der Studierenden, dem Wissensstand im Fachbereich Orthopädie und dem Interesse am Fachbereich Orthopädie, wobei das Hauptaugenmerk auf dem Wissensstand liegt.

Hinsichtlich der **Zufriedenheit der Studierenden** konnte in dieser Untersuchung folgendes aufgezeigt werden: Die Mehrheit der Studierenden ist mit der Umsetzung der Lehre in den Vorlesungen und Praktika zufrieden. Hinsichtlich der Änderungen durch die neue AO empfindet sich der Großteil der Studierenden als nicht gut über die Änderungen informiert. Die Einführung des „Hammerexamens“ wird von der Mehrheit (74%) abgelehnt. Bevorzugte Prüfungsmethode sind MC-Fragen, jedoch sollen mündliche Prüfungen und Offene Fragen als Ergänzung dienen. Die Hälfte der Studierenden ist mit ihrer Situation als Medizinstudent zufrieden, jedoch ist auch jeder fünfte Student unzufrieden und vom Medizinstudium nicht überzeugt. Als meist genannte Änderungspunkte werden das zeitlich längere Dozieren von wichtigen Fächern wie Innere Medizin und Chirurgie, der Wunsch nach einer mehr auf Praxis orientierten Ausbildung und die Abschaffung des „Hammerexamens“ angegeben.

Die in dieser Arbeit dargestellte Untersuchung liefert Ergebnisse über den **Wissensstand im Fachbereich Orthopädie** in verschiedenen Ausbildungsabschnitten und ermöglicht den Vergleich von Multiple-Choice Fragen und Offenen Fragen in ihrer Anwendung am gleichen Untersuchungskollektiv. Im Einzelnen konnten die primär aufgestellten Hypothesen wie folgt beurteilt werden:

1) Ergebnisse für das 1. klinische und 2. klinische Semester

Die Hypothese, dass im 1. klinischen und 2. klinischen Semester die schlechtesten Ergebnisse zu finden sind, bestätigt sich bei den MC-Fragen und auch bei den Offenen Fragen. Die Ausnahme bildet Gruppe 3-Prä bei den Offenen Fragen mit einem schlechteren Ergebnis als das 1. klinische und 2. klinische Semester. Da dieses Ergebnis jedoch nicht erklärbar ist und auch die einzige Ausnahme bildet, ist die gestellte Hypothese in ihrer Grundaussage zu verifizieren.

2) Ergebnisse für das 3. klinische und 4. klinische Semester (Prä-Test)

Die Hypothese, dass in den Prä-Gruppen das orthopädische Wissen seinen Anfang nimmt und das Interesse am Fachgebiet Orthopädie steigt, findet in den Ergebnissen der vorliegenden Arbeit ihre Bestätigung und ist somit zu verifizieren.

3) Größter Wissenszuwachs im Laufe des Orthopädie-Praktikums

Diese Hypothese muss zweigeteilt betrachtet werden. Der größte prozentuale Wissenszuwachs findet für die MC-Fragen vom 2. klinischen Semester auf die Prä-Gruppen statt. Trotz der ebenfalls deutlichen Ergebnissteigerung im Laufe des Orthopädie-Praktikums für die MC-Fragen muss diese Hypothese für den Teilbereich MC-Fragen im engen Sinne falsifiziert werden, da dort nur die zweitgrößte Ergebnissteigerung anzufinden ist. Für den Teilbereich Offene Fragen findet der mit Abstand größte Wissenszuwachs im Laufe des Orthopädie-Praktikums statt, so dass hierfür die Hypothese als verifiziert anzusehen ist.

4) Ergebnisse für das 5. klinische und 6. klinische Semester

Die Hypothese, dass nach Absolvierung des Orthopädie-Praktikums ein Rückgang des orthopädischen Wissens erfolgt und sich das Wissen zum Ende des Studiums hin stagniert, ist anhand der Ergebnisse verifiziert.

5) Vergleich von MC-Fragen mit Offenen Fragen

Das Erreichen von besseren Werten bei der Beantwortung von MC-Fragen als bei den Offenen Fragen ist durch die vorliegenden Ergebnisse für die Gruppen mit bereits erworbenem orthopädischem Wissen bestätigt worden.

6) Zufriedenheit der Studierenden und Interesse am Fachgebiet Orthopädie

Die Hypothese eines im Laufe des Studiums zunehmend ambivalenten Verhaltens der Zufriedenheit der Studierenden und des Interesses am Fachgebiet Orthopädie muss falsifiziert werden. Vielmehr tritt für beide Verhaltensmuster eine Verschiebung der Mehrheiten von „zufrieden“ nach „unzufrieden“ und von „Interesse“ zu „wenig Interesse“ auf.

7) Einschätzung des Schwierigkeitsgrades der gestellten Fragen

Das hypothetisch zu erwartende Ergebnis der Verschiebung des Schwierigkeitsgrades von Kategorie „schwierig“ zu „leicht“ tritt nicht ein. Stattdessen ergibt sich eine Verschiebung von Kategorie „schwierig“ hin zu „angemessen“ in den Semestern mit dem Orthopädie-Praktikum im Lehrplan, während sich daraufhin die Einschätzung zurück in Kategorie „schwierig“ verlagert.

Insgesamt zeigt sich, dass die in Kapitel 1.2 aufgestellten Hypothesen zum Teil verifiziert wurden, zum Teil aber auch falsifiziert werden mussten. Dieses Ergebnis legt offensichtlich an den Tag, dass vor der Durchführung einer Untersuchung nur Erwartungen aufgestellt werden können, die auf bereits durchgeführten Studien (20, 41, 49, 55, 76) basieren, jedoch eine prospektive Arbeit immer wieder neue und teils unerwartete Ergebnisse hervorbringt. Sowohl in der Bestätigung bereits fundierter Ergebnisse als auch im Aufzeigen neuer und unerwarteter Ergebnisse liegt die Durchführung der vorliegenden Untersuchung bestätigt.

5.) Literaturverzeichnis

- (1) Anbar, M.: CAI – Computer Assisted Instruction – A way to avoid the pitfalls of multiple-choice behavior in medical practice. Medical Electronics, 18, 118-124, (1987)
- (2) Anbar, M.: Open-ended computer-based medical instruction – Avoiding the pitfalls of multiple-choice behavior. Resident & Staff Physician, 32, NO. 7, 91-95, (1986)
- (3) Ansprache des Bundesministers für Gesundheit, Horst Seehofer, zur Eröffnungsveranstaltung des 96. Deutschen Ärztetages in Dresden, 04.05.1993, (1993)
- (4) APA (American Psychological Association: Standards für Educational and Psychological Tests. APA, Washington D.C., (1974)
- (5) Approbationsordnung für Ärzte vom 28.10.1970. BGB1. I, Nr. 98, S. 1498, (1970)
- (6) Approbationsordnung für Ärzte vom 17.12.1997, (1997)
- (7) Approbationsordnung für Ärzte vom 27.06.2002, (2002)
- (8) Approbationsordnung für Ärzte, mit Kommentaren und praktischen Hinweisen von H.-P. Brauer, Th. Zickgraf. Deutscher Ärzteverlag, Köln, (1975)
- (9) Arzt im Praktikum, Notbehelf gekippt, Deutsches Ärzteblatt, Jg. 100, Heft 47, 21.11.2003, (2003)
- (10) Bestallungsordnung für Ärzte, mit einer Einführung von R. Rachold, 11.Auflage, Deutscher Ärzteverlag, Köln, (1968)

- (11) Bestallungsordnung für Ärzte vom 15.09.1953. BGB1., S. 1334, (1953)
- (12) Black, N.M.I. & Harden, R.M.: Providing feedback to students on clinical skills by using the Objective Structured Clinical Examination (OSCE). Medical Education, 20, 48-52, (1986)
- (13) Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwol, D.R.: The Taxonomy of Educational Objectives. New York: David McKay, (1956)
- (14) Bordage, G. & Zacks, R.: The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes, Medical Education, 18, S. 406-416, (1984)
- (15) Buckup, K.: Klinische Tests an Knochen, Gelenken und Muskeln – Untersuchungen - Zeichen - Phänomene, Thieme Verlag, 3. überarbeitete und erweiterte Auflage, (2005)
- (16) Buckwalter, J.A., Schumacher, R., Albright, J.P. & Cooper, R.R.: Use of an educational taxonomie for evaluation of cognitive performance, Journal of Medical Education, 56, S. 115-121, (1981)
- (17) Bundesministerium für Gesundheit, Internet-Seite des BMG: Die Ärztliche Ausbildung, <http://www.bmggesundheit.de/themen/berufe/voha/berufe/reform.htm>, (2006)
- (18) Cox, K.R.: How did you guess? Or, what do multiple-choice questions measure? Medical Journal of Australia, 1, 884-886, (1976)
- (19) Das Netzwerk: Positionspapier zur Reform des Prüfungswesens im Medizinstudium, Korrespondenzadresse: Prof. Dr. med. F. Eitel, Chirurgische Klinik Innenstadt LMU, 80336 München (1994)

- (20) Dissertation: Analyse des fachorthopädischen Wissensstandes in verschiedenen Ausbildungsabschnitten und Vergleich zweier schriftlicher Prüfungsmethoden; vorgelegt von Andrea Elmer, 1997, Orthopädische Klinik des St. Josef-Hospitals Bochum - Universitätsklinik – der Ruhr-Universität Bochum, Direktor: Prof. Dr. med. J. Krämer, (1997)
- (21) Eitel, F. et. al.: Verbesserung des Studentenunterrichts – Sicherung der Strukturqualität Medizinischer Versorgung. In: Habeck, Schagen, Wagner (1993) Reform der Ärzteausbildung. Blackwell Wissenschaft, 1993, (1993)
- (22) Eitel, F.: Sicherung der Strukturqualität in der Chirurgie – Ansätze zum Qualitätsmanagement im Gesundheitswesen. In: Maurer, P.C., von Sommoggy, S. (Hrsg.) Gefäßchirurgie im Fortschritt, Blackwell Wissenschaft, Berlin, (1995)
- (23) Elstein, A.S.: Beyond multiple-choice questions and essays: The need for a new way to assess clinical competence. *Academic Medicine*, 68, 244-249, (1993)
- (24) Feletti, G.I. & Engel, C.E.: The modified essay question for testing problem-solving-skills. *The Medical Journal of Australia*, 1980, 1: 79-80, (1980)
- (25) Frary, R.B.: Multiple-choice versus free-response: A simulation study. *Journal of Educational measurement* Vol. 22, NO. 1, pp 21-31, (1985)
- (26) Grifka, J.: Orthopädie und Unfallchirurgie in Frage und Antwort – Fragen und Fallgeschichten zur Vorbereitung auf mündliche Prüfungen während des Semesters und im Examen, Urban & Fischer Verlag, 4. Auflage, (2004)
- (27) Grifka, J.: Stellenwert der Lehre in der Mediziner Ausbildung, Arbeitstagung: „Innovation und Trends des Medizinstudiums im klinischen Teil“, Tübingen, 09.07.1993, herausgegeben von K.-H. Bichler und W. Mattauch, pmi Verlagsgruppe GmbH, Frankfurt, Moskau, Sennwald, Wien, (1994)

- (28) Haladyna, T.M.: Developing and Validating Multiple-Choice Test Items, Lawrence Erlbaum Associates, Third Edition, Mahwah, New Jersey, (2004)
- (29) Hakstian, A.R.: The effects of type of examination anticipated on test preparation and performance. *The Journal of Educational Research* 64: 319-324, (1971)
- (30) Harden, R.M.: A fresh look at clinical teaching. In: Bender W. et al. (ed.): Teaching and assessing clinical competence, p. 5-9, Boek Werk, Groningen, (1990)
- (31) Harden, R.M., Brown, R.A., Biran, L.A., Dallas Ross, W.P. & Wakeford, R.E.: Multiple-choice questions: to guess or not to guess. *Medical Education*, 10, 27-32, (1976)
- (32) Harden, R.M., Gleeson, F.A.: Assessment of clinical competence using objective structured examination. *Medical Education* 13, 41-54, (1979)
- (33) Harden, R.M., Stevenson, M., Downie, W.W. & Wilson, G.M.: Assessment of clinical competence using objective structured examination. *British Medical Journal* I, 447-451, (1975)
- (34) Hassard, T.H., Campbell, C., Klass, D., Kopelow, M., Schnabl, G.: An examination of the factor structure of clinical competence. In: Bender W. Eet al. (ed): Teaching and assessing clinical competence, p. 352-356, Boek Werk, Groningen, (1990)
- (35) Hodgkin, K., Knox, J.D.E.: Problem Centered Learning. London: Churchill Livingstone (1975)
- (36) Huerkamp, C.: Der Aufstieg der Ärzte im 19.Jahrhundert – vom gelehrten Stand zum professionellen Experten: Das Beispiel Preußens. *Kritische Studien zur Geschichtswissenschaft*, Bd. 68, Vandenhoeck & Ruprecht, Göttingen, (1985)

- (37) Irwin, W.G., Bamber, J.H.: The cognitive structure of modified essay question. *Medical Education*, 16, S. 326-331, (1982)
- (38) Kienle, G. & Kreysch, W.: Multiple-choice-Prüfung und ärztliche Berufsfähigkeit. *Deutsches Ärzteblatt*, 14, S. 838-841, (1978)
- (39) Klauer, K.J.: Kontenvalidität von Prüfungen. In: Westhoff, K. (Hrsg.) *Erstes Symposium zu Prüfungen in der Medizin: Multiple Choice. Möglichkeiten und Grenzen von Multiple-Choice-Prüfungen in der Medizin* (S. 78-88). Lengerich: Pabst, (1995)
- (40) Klauer, K.J.: *Kriteriumsorientierte Tests. Lehrbuch der Theorie und Praxis lehrzielorientierten Messens*, Göttingen, Toronto, Zürich: Hogrefe, (1987)
- (41) Klocke, R.: *Ein Beitrag zur Erfolgsmessung der medizinischen Ausbildung im Fach Orthopädie*. Diss., Universität Düsseldorf, (1980)
- (42) Kommission der Europäischen Gemeinschaften: *Bericht über die Beurteilungsmethoden in der medizinischen Grundausbildung*, Beratender Ausschuss für die ärztliche Ausbildung, III/D 1643/7/83-DE, Brüssel (1987)
- (43) Kurzfassung „Bericht der Sachverständigengruppe zu Fragen der Neuordnung des Medizinstudiums“ beim Bundesminister für Gesundheit, März 1993, (1993)
- (44) Leeder, S.R., Feletti, G.I., Engel, C.E.: Assessment – help or hurdle. *Programmed Learning and Education Technology*, 16, 309, (1979)
- (45) Lennox, B.: Marking Multiple-Choice Examinations. *British Journal of Medical Education*, 1, 203-211, (1967)

- (46) Levine, H.G., Mc Guire, C.H., Nattress, L.W.: The validity of multiple-choice achievement tests as measures of competence in medicine. *American Educational Research Journal*, Vol. 7, NO. 1. January 1970, S. 69-82, (1970)
- (47) Liertz, W., Paffrath, H. (Hrsg.): *Handbuch des Arztes*. Schwann Verlag, Düsseldorf, (1940)
- (48) McCarthy, W.H., M.B.B.S., F.R.A.C.S.: An assessment of the influence of cueing items in objective examinations. *Journal of Medical Education*, 10, 382-385, (1966)
- (49) McCloskey, D.I. & Holland, R.A.B.: A comparison of student performances in answering essay-type and multiple-choice questions. *Medical Education*, 10, 382-385, (1976)
- (50) McGuire, C.: A Process Approach to the Construction and Analysis of Medical Examinations, *Journal of Medical Education*, 38: 556-563 (1963)
- (51) Medizinische Hochschule Hannover, Rahmenbedingungen, <http://www.mh-hannover.de/73.html>, (2004)
- (52) Medizinstudium, Kritische Bewertung der Reform braucht Zeit, *Deutsches Ärzteblatt*, Jg.102, Heft 51-52, 26. Dezember 2005, (2005)
- (53) Murrhardter Kreis: *Das Arztbild der Zukunft*. 3., vollständig überarbeitete Auflage, Robert Bosch Stiftung, Bleicher Verlag (1994)
- (54) Narr, H.: *Ärztliches Berufrecht – Ausbildung, Weiterbildung, Berufsausübung*. Enke Verlag, Stuttgart, (1973)

- (55) Newble, D.I., Baxter, A. & Elmslie, R.G.: A comparison of multiple-choice tests and free-response tests in examinations of clinical competence. *Medical Education*, 13, S. 263-268, (1979)
- (56) Newble, D.I. & Entwistle N.J.: Learning styles and approaches: implications for medical education, *Medical Education*, 20, S. 162-175, (1986)
- (57) Newble, D.I., Swanson, D.B.: Psychometric characteristics of the objective structured clinical examination. *Medical Education* 22, 325-334, (1988)
- (58) Niederschrift der Besprechung über den Entwurf einer 7.Verordnung zur Änderung der Approbationsordnung am 13.07.1989 im BM JFFG, (1989)
- (59) Niethard, F, Pfeil, J.: *Duale Reihe Orthopädie*, Thieme Verlag, 5. korrigierte Auflage, (2005)
- (60) Nnodim, J.O.: Multiple-choice testing in anatomy. *Medical Education* 1992 Jul, 26(4), 301-309, (1992)
- (61) Norcini, J.J., Swanson, D.B., Grosso, L.J. & Webster, G.D. : Reliability, validity and efficiency of multiple choice questions and the patient management problem item formats in assessment of clinical competence. *Medical Education*, 19, S. 238-247, (1985)
- (62) Norman, G.R.: Problem-solving skills. Solving problems and problem-based learning, *Medical Education*, 22, S. 279-186 (1988)
- (63) Norman, G.R., Smith, E.K.M., Powles, A.C.P., Rooney, P.J., Henry, N.L. & Dodd, P.E.: Factors underlying performance on written tests of knowledge. *Medical Education*, 21, S. 297-304, (1987)

- (64) Odenbach, E.: Stellungnahme des Marburger Bundes zur Approbationsordnung. In: Medizinstudium und Approbationsordnung. Materialien der Bundesassistentenkonferenz, Bd. 5, Bonn, (1970)
- (65) Pabst, R.: How do medical students accept different types of assessment and how does this influence teaching and learning? Vortrag, Internationales Symposium "Prüfungsmethoden", Mainz, 29./30.05.1995, (1995)
- (66) Pauli, H.G., IAE (Institut für Ausbildungs- und Examensforschung) der medizinischen Fakultät Bern: Die Multiple-Choice-Methode in der Evaluation medizinischer Ausbildungsergebnisse. Schweizerische Ärztezeitung, Bd. 66, Heft 26, 1193-1201, (1985)
- (67) Planungsgruppe Medizin, Hardegg, W., Schäfer, M., Nelle, A., Richterich, A.: Gutachten über das Prüfungswesen in einer neuen Approbationsordnung für Ärzte (ÄAppO) anhand einer Untersuchung ausländischer Prüfungssysteme im Studiengang Medizin im Zusammenhang mit der Diskussion zur neuen Approbationsordnung für Ärzte. Erstellt im Auftrag des Bundesministeriums für Gesundheit. Medizinischer Fakultätentag, Heidelberg, Februar 1995, (1995)
- (68) Pressemitteilung der Bundesärztekammer, Ärztetag fordert Abschaffung des AiP, http://www.bundesaerztekammer.de/30/Aerztetag/105_DAET/20Pressemitt/200205297.html , (2002)
- (69) Reform des Medizinstudiums, Ein langer Weg, Deutsches Ärzteblatt, Jg. 99, Heft 18, 03.05.2002, (2002)
- (70) Reform des Medizinstudiums, Medizinische Fakultäten jetzt am Zug, Deutsches Ärzteblatt, Jg. 99, Heft 47, 22. November 2002, (2002)
- (71) Reznick, R. et. al.: An Objective Structured Clinical Examination for the Licentiate: Report of the Pilot Project of the Medical Council of Canada. Academic Medicine 1993, 67(8): 487-494, (1992)

- (72) Rothman, A.I. & Kerenyi, N.: The assessment of an examination in pathologie consisting of multiple-choice, practical and short essay questions, *Medical Education*, 14, 341-344 (1980)
- (73) Rössler H., Rütter, W.: *Orthopädie und Unfallchirurgie*, Urban & Fischer Verlag, 19.Auflage, (2005)
- (74) Schimmelpfenning, K.: Empfehlungen zur gesetzlichen Neugestaltung der Ärztlichen Vorprüfung und der Ärztlichen Prüfung. *Med. Ausbildung* 8:5-14 und *dtsh. Ärzteblatt* 89(5)B: 204-206, (1991)
- (75) Schumacher, C.F.: Reliability, Validity and Standard Setting. In: Hubbard, J.P., 9-71, *Measuring Medical Education*. Lea & Febiger, Philadelphia 1978, S. 59-71, (1978)
- (76) Schwartz, R.W., Donnelly, M.B., Sloan, D.A. & Young, B.: Knowledge gain in a problem-based surgery clerkship. *Academic Medicine*, 69, 148-151, (1994)
- (77) Siebente Verordnung zur Änderung der Approbationsordnung für Ärzte vom 21.12.1989, *BGB1 I*, Nr. 62, S. 2549, (1989)
- (78) Smith, R.M.: The Triple Jump Examinations as an Assessment Tool in the Problem-Based Medical Curriculum at the University of Hawaii. *Academic Medicine* 1993, 68(5): 366-372, (1993)
- (79) Stalenhoef-Halling, B.F., Van Der Vleuten, C.P.M., Jaspers, T.A.M., and Fiolet, J.F.B.M.: The feasibility, acceptability und reliability of open-ended questions. In: Bender, W. et al. (ed.): *Teaching and assessing clinical competence*, p.552-557, Boek Werk, Groningen, (1990)

- (80) Stöckel, S., Neitzke, G., Frewer, A., Lohff, B. (Abteilung Geschichte, Ethik und Philosophie der Medizin, Medizinische Hochschule Hannover): Innovation im Medizinstudium, Geschichte, Theorie und Ethik, Deutsches Ärzteblatt, Jg. 102, Heft 6, 11.02.2005, (2005)
- (81) Stork, G.: Neue Approbationsordnung, Änderungen der Approbationsordnung – die unendliche Geschichte, www.examensplaner.de/contents/artikel/dappo.php, medizinstudent.de GmbH, (2002)
- (82) Thorndike, R.L. (ed): Educational Measurement (second edition) (chapters 2-8). American Council on Education, Washington D.C., (1971)
- (83) TOP IV: Tätigkeitsbericht – Approbationsordnung, Langer Atem erforderlich, Deutsches Ärzteblatt, Jg. 101, Heft 22, 28.05.2004, (2004)
- (84) Trost, G.: Evaluation of examinations-methods, results, conclusions. Invited paper, presented at the International Symposium on Examinations in Medical Education and their influence on teaching and learning, Mainz, 29./30. 06. 1995, (1995)
- (85) Van den Busche, H.: Prüfungen in der ärztlichen Ausbildung – Funktionen, Möglichkeiten und Grenzen. In: Habeck, Schagen, Wagner (1993) Reform der Ärzteausbildung, S. 91-107, Blackwell Wissenschaft, (1993)
- (86) Wallace, W.F.M.: Student attitudes to examinations and how they are related to performance. British Journal of Medical Education, 5, 273-278, (1971)
- (87) Whitby, L.G.: Marking systems for multiple-choice examinations. Medical Education, 11, 216-220, (1977)
- (88) Wirsching, M.: Fünf Thesen zur Reform der ärztlichen Ausbildung. Dtsch. Ärztebl. 85:13-17, (1988)

- (89) Wissenschaftsrat: Leitlinien zur Reform der ärztlichen Ausbildung, Eigenverlag, Bremen, 03.07.1992, (1992)
- (90) Wülker, N.: Taschenlehrbuch – Orthopädie und Unfallchirurgie, Georg Thieme Verlag, Stuttgart, (2005)
- (91) Yang, J., Verheul, I.C., Verhelst, N., Essen Van E.: Opinions about multiple choice- and open-questions-tests and differences in study activities as a result of test expectation: report of a questionnaire. Tijdschrift voor Onderwijsresearch 10: 179-188, (1985)

6.) Tabellenverzeichnis

Tabelle 1:	Gruppenaufteilung für das Untersuchungskollektiv an der Universität Regensburg	- 65 -
Tabelle 2:	Gruppenaufteilung für das Untersuchungskollektiv an der Universität Heidelberg	- 66 -
Tabelle 3:	Gruppenaufteilung für das gesamte Untersuchungskollektiv	- 66 -
Tabelle 4:	Ergebnisse für Frage 1	- 68 -
Tabelle 5:	Ergebnisse für Frage 2	- 69 -
Tabelle 6:	Ergebnisse für Frage 3	- 71 -
Tabelle 7:	Ergebnisse für Frage 4	- 72 -
Tabelle 8:	Ergebnisse für Frage 5	- 74 -
Tabelle 9:	Ergebnisse für Frage 6	- 75 -
Tabelle 10:	Ergebnisse für Frage 7	- 77 -
Tabelle 11:	Ergebnisse für Frage 8	- 78 -
Tabelle 12:	Ergebnisse für Frage 9	- 80 -
Tabelle 13:	Ergebnisse für Frage 10	- 82 -
Tabelle 14:	Ergebnisse Fragenteil II (MC-Fragen) Richtigantworten und Standardabweichung (absolut und in Prozent)	- 84 -

Tabelle 15:	Statistische Analyse: p-Werte für die MC-Fragen.....	- 86 -
Tabelle 16:	Ergebnisse Fragenteil II (Offene Fragen) Erreichte Punktzahl und Standardabweichung (absolut und in Prozent)	- 88 -
Tabelle 17:	Statistische Analyse: p-Werte für die Offenen Fragen.....	- 90 -
Tabelle 18:	Ergebnisse für die Fragen 23, 24, 25 (Teil II, Offene Fragen) Erreichte Punktzahl (absolut und in Prozent).....	- 92 -
Tabelle 19:	Ergebnisse für Frage 26 (Teil III)	- 95 -
Tabelle 20:	Ergebnisse für Frage 27 (Teil III)	- 97 -
Tabelle 21:	Ergebnisse für Frage 28 (Teil III)	- 98 -
Tabelle 22:	Ergebnisse für Frage 29 (Teil III)	- 100 -
Tabelle 23:	Ergebnisse für Frage 30 (Teil III)	- 101 -
Tabelle 24:	Vergleich der Ergebnisse von Prä- und Post-Test (Gruppen 3, 4, 8)	- 103 -
Tabelle 25:	Ergebnisse für die Fragen 11, 13, 16 und 17.....	- 105 -
Tabelle 26:	Ergebnisse für die Fragen 12 und 24.....	- 107 -
Tabelle 27:	Ergebnisse für die Fragen 14, 15, 18 und 19.....	- 108 -
Tabelle 28:	Ergebnisse für die Fragen 20, 21, 22 und 25.....	- 110 -
Tabelle 29:	Ergebnisse für die Frage 23.....	- 111 -
Tabelle 30:	Ergebnisse für die MC-Fragen im Einzelnen	- 113 -

7.) Diagrammverzeichnis

Diagramm 1:	Zeitdiagramm der Änderungen der ärztlichen Approbationsordnung	- 11 -
Diagramm 2:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Wissensstandes zu den Änderungen durch die neue AO	- 69 -
Diagramm 3:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung der Änderung hinsichtlich der Anzahl der Staatsexamina durch die neue Approbationsordnung	- 70 -
Diagramm 4:	Mittelwerte aller Gruppen (Gruppen 1-8) zur bevorzugten Prüfungsmethode	- 72 -
Diagramm 5:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Vermittlung der Inhalte innerhalb der Praktika	- 73 -
Diagramm 6:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung einer ausreichenden Vorstellung von Patienten in den Vorlesungen	- 75 -
Diagramm 7:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung einer ausreichenden Vorstellung von Patienten in den Praktika.....	- 76 -
Diagramm 8:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Beurteilung der Strukturierung des Studiums in Bezug auf die Blockpraktika	- 78 -
Diagramm 9:	Mittelwerte der Gruppen 1-6 (Universität Regensburg) zur Beurteilung von ausreichend praktischem Unterricht	- 79 -
Diagramm 10:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Zufriedenheit mit Situation als Medizinstudent.....	- 81 -

Diagramm 11:	Mittelwerte aller Gruppen (Gruppen 1-8) zu Änderungspunkten an der aktuellen Studiensituation.....	- 83 -
Diagramm 12:	Mittelwerte der erreichten Punktwerte der einzelnen Gruppen für die MC-Fragen.....	- 85 -
Diagramm 13:	Mittelwerte der erreichten Punktwerte der einzelnen Gruppen für die Offenen Fragen.....	- 89 -
Diagramm 14:	Mittelwerte der erreichten Punkte der einzelnen Gruppen für die Offenen Fragen im Einzelnen.....	- 94 -
Diagramm 15:	Mittelwerte der erreichten Punktzahlen für MC-Fragen und Offene Fragen in Prozent.....	- 94 -
Diagramm 16:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Schwierigkeitsgrades der gestellten MC-Fragen	- 96 -
Diagramm 17:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Einschätzung des Schwierigkeitsgrades der gestellten Offenen Fragen	- 98 -
Diagramm 18:	Mittelwerte aller Gruppen (Gruppen 1-8) zum Interesse am Fachbereich Orthopädie	- 99 -
Diagramm 19:	Mittelwerte aller Gruppen (Gruppen 1-8) zum Berufswunsch Facharzt für Orthopädie / Unfallchirurgie	- 101 -
Diagramm 20:	Mittelwerte aller Gruppen (Gruppen 1-8) zur Befürwortung einer solchen Evaluation zum Thema Studienordnung und Fachwissenstand.....	- 102 -
Diagramm 21:	Vergleich der Ergebnisse von Prä-Test und Post-Test für die Gruppen 3, 4 und 8.....	- 104 -

Diagramm 22:	Mittelwerte der Ergebnisse für die Fragen zu typisch orthopädischen Symptomen.....	- 106 -
Diagramm 23:	Mittelwerte der Ergebnisse für die Fragen zur Statik	- 108 -
Diagramm 24:	Mittelwerte der Ergebnisse für die Fragen zur orthopädischen Diagnostik / zu orthopädischen Zeichen in der Diagnostik	- 109 -
Diagramm 25:	Mittelwerte der Ergebnisse für die Fragen zu spezifisch orthopädischen Krankheitsbildern	- 111 -
Diagramm 26:	Mittelwerte der Ergebnisse für die Frage zu anatomischen Begriffen	- 112 -
Diagramm 27:	Mittelwerte aller Gruppen (Gruppen 1-8) der Ergebnisse zum Wissensstand nach inhaltlichen Kriterien.....	- 112 -

8.) Abkürzungsverzeichnis

ÄAppO	Ärztliche Approbationsordnung
AiP	Arzt im Praktikum
AO	Approbationsordnung
AppOÄ	Approbationsordnung für Ärzte
BÄO	Bundesärzteordnung
BMG	Bundesministerium für Gesundheit
GTE	Geschichte, Theorie und Ethik der Medizin
Heicumed	Heidelberger Curriculum Medicinale
IMPP	Institut für Medizinische und Pharmazeutische Prüfungsfragen
MC	Multiple-Choice
MCQs	Multiple Choice Questions
MCIP	Münchner Curriculare Innovationsprojekt
MEQ	Modified-Essay-Question
MFT	Medizinischer Fakultätentag
NBME	National Board of Medical Examiners
OSCE	Objective Structured Clinical Examination
PMP-Fragetyp	Patient Management Problem Solving
PMPs	Patient Management Problems
PT	Progress Test
SIMP-Test	Simulation of Initial Medical Problem Solving
VO	Verordnung
z.B.	zum Beispiel
ZVS	Zentrale Vergabestelle für Studienplätze

9.) Anhang

Fragebogen

Lösungsschema für die MC-Fragen

Lösungsschema für die Offenen Fragen

Liebe Kommilitoninnen, liebe Kommilitonen,

der folgende Fragebogen soll im Rahmen meiner Promotion in der Orthopädie sowohl über die Zufriedenheit der Studenten mit der aktuellen Studiensituation, als auch über den Wissensstand im Fachbereich Orthopädie Auskunft geben. Deswegen bitte ich Euch, folgende Fragen ehrlich und nach bestem Wissen zu beantworten:

Fragen zur Zufriedenheit der Studenten im Fachbereich Medizin in Hinblick auf Inhalte, Struktur und Umsetzung durch die Dozenten unter Berücksichtigung der modifizierten Approbationsordnung

- | | |
|---|--|
| <u>Frage 1)</u> Wie schätzen Sie Ihren Wissensstand zu den Änderungen durch die neue Approbationsordnung ein? | sehr gut
gut
unentschieden
schlecht
sehr schlecht |
| <u>Frage 2)</u> Wie finden Sie die Änderung durch die neue Approbationsordnung in Hinblick auf die Anzahl der Staatsexamina („Hammerexamen“ am Ende)? | sehr gut
gut
unentschieden
schlecht
sehr schlecht |
| <u>Frage 3)</u> Was bevorzugen Sie in der Prüfungssituation? | MC-Fragen
Offene Fragen
mündliche Prüfung |
| <u>Frage 4)</u> Wie finden Sie die Vermittlung der Inhalte innerhalb der Praktika? | sehr gut
gut
unentschieden
schlecht
sehr schlecht |
| <u>Frage 5)</u> Werden Ihrer Meinung nach in den Vorlesungen genügend Patienten vorgestellt? | trifft völlig zu
trifft zu
unentschieden
trifft nicht zu
trifft gar nicht zu |

Frage 6) Werden Ihrer Meinung nach in den Praktika genügend Patienten vorgestellt?

trifft völlig zu
trifft zu
unentschieden
trifft nicht zu
trifft gar nicht zu

Frage 7) Wie finden Sie die Strukturierung des Studiums in Bezug auf die Blockpraktika?

sehr gut
gut
unentschieden
schlecht
sehr schlecht

Frage 8) Halten Sie es für ausreichend, im gesamten Studium 4 Wochen Chirurgie
5 Wochen Innere Medizin
2 Wochen Orthopädie
praktisch unterrichtet zu werden?

trifft völlig zu
trifft zu
unentschieden
trifft nicht zu
trifft gar nicht zu

Frage 9) Sind Sie mit Ihrer Situation als Medizinstudent zufrieden und von Ihrem Studium überzeugt?

sehr zufrieden / absolut überzeugt
zufrieden / überzeugt
unentschieden
nicht zufrieden / nicht ganz überzeugt
gar nicht zufrieden/Zweifel am Studium

Frage 10) Was würden Sie an der aktuellen Studiensituation ändern (Stichpunkte)?

Fragen zum Wissensstand der Studenten im Fachbereich Orthopädie

Frage 11) Bei einem jungen Volleyballspieler hängt nach einem Aufprall des Balls auf den gestreckten Mittelfinger das Endglied herab & kann nicht aktiv gestreckt werden. Es handelt sich um einen

- A Schwanenhalsfinger
- B Hammerfinger
- C Knopflochdeformität
- D Kamptodaktylie
- E Klinodaktylie

Frage 12) Eine funktionelle Beinverkürzung rechts kann bedingt sein durch

- A Kniegelenksbeugekontraktur links
- B Spitzfußkontraktur rechts
- C kompensatorischen Beckenschiefstand bei rechtskonvexer LWS-Skoliose
- D Adduktionskontraktur des rechten Hüftgelenks
- E Hüftgelenksdysplasie links

Frage 13) Welche Aussage trifft nicht zu?

Der Hallux rigidus

- A beruht auf einer Arthrose des Großzehengrundgelenks
- B beginnt mit einer Beugehemmung im Großzehengrundgelenk
- C behindert das Abrollen des Fußes
- D wird konservativ durch eine starre Einlage behandelt
- E kann eine Indikation zur operativen Behandlung darstellen

Frage 14) Zu den typischen statischen Veränderungen bei Alters-Osteoporose gehören

1. eine vermehrte Brustkyphose
2. Genua valga
3. Beckenschiefstand mit Beinlängendifferenz
4. eine Skoliose der Wirbelsäule mit Rippenbuckelbildung
5. Bauchvorwölbung

- A nur 1 ist richtig
- B nur 1 und 5 sind richtig
- C nur 2 und 5 sind richtig
- D nur 1, 4 und 5 sind richtig
- E alle Aussagen 1 – 5 sind richtig

Frage 15) Was bezeichnet man mit dem Begriff „schmerzhafter Bogen“?

- A die Unterarmzerrung beim klassischen Bogenschießen
- B die letzten Grade der Kniebeugung beim Meniskusschaden
- C schmerzhafte Schulterabduktion zwischen 30° und 50°
- D schmerzhafte Schulterabduktion zwischen 70° und 130°
- E Bewegungsschmerz bei Ellbogenstreckung

- Frage 16) Was ist typisch für pseudoradikuläre Beschwerden?
- A klarer bandförmiger Schmerz im Verlauf einer Nervenwurzel am Bein
 - B isolierte Fußheberschwäche
 - C diffuser Schmerz im Bein, u.U. in beide Beine strahlend
 - D isolierter Druckschmerz im Rückenbereich
 - E Schmerzverstärkung bei sog. Würfelposition (90° Beugung in Hüft- und Kniegelenk)
- Frage 17) Welche der folgenden Aussagen trifft zu?
- A Rupturen des Außenmeniskusvorderhorns sind die häufigste klinische Meniskuläsion
 - B Der M. Osgood-Schlatter erfordert regelhaft eine operative Intervention
 - C Der sog. Scheiben-Meniskus betrifft fast ausschließlich den Innenmeniskus
 - D Ein Streckdefizit am Kniegelenk führt zu einem erhöhten Patella-Anpressdruck und prädisponiert zur Femuropatellararthrose
 - E angeborene Knie-Luxationen müssen notfallmäßig operiert werden
- Frage 18) Ein 13-jähriger Junge berichtet in der Sprechstunde über ziehende Schmerzen im Oberschenkel unter sportlicher Belastung. Welche weiteren Untersuchungen führen Sie nach der klinischen Untersuchung direkt in Ihrer Klinik/Ambulanz/Praxis sinnvollerweise durch?
- 1. Sonographie
 - 2. Röntgen
 - 3. Kernspintomographie
 - 4. Computertomographie
 - 5. Szintigraphie
- A nur 1 und 2 sind richtig
 - B nur 1, 2 und 3 sind richtig
 - C nur 1, 2 und 4 sind richtig
 - D nur 1, 2 und 5 sind richtig
 - E alle Aussagen 1 – 5 sind richtig
- Frage 19) Eine Streckhemmung im Kniegelenk weist hin auf eine(n)
- A Außenbandriss
 - B Innenbandabriss
 - C Meniskuseinklemmung
 - D Riss des vorderen Kreuzbandes
 - E Riss des hinteren Kreuzbandes
- Frage 20) Ein 5-jähriges Mädchen mag seit 3 Monaten im Kindergarten nicht mehr gerne mit herumlaufen. Die Eltern berichten, dass sie beim Spazieren gehen rasch über Müdigkeit klage und getragen werden will. Woran denken Sie?
- A Psychische Erkrankung, z.B. Angstzustände
 - B Epiphysiolysis capitis femoris
 - C Leistenhernie
 - D Morbus Perthes
 - E Fibromyalgie-Syndrom

Frage 21) Kennzeichnend für eine funktionelle Skoliose ist:

- A Ausgleichbarkeit zur Gegenseite
- B Wirbelkörper torsion
- C Atrophie des M. erector spinae
- D Progredienz der Krümmung
- E Dominanz des weiblichen Geschlechts

Frage 22) Welches sind typische Kennzeichen für einen Morbus Scheuermann (Adoleszentenkyphose)?

- 1. Kyphose der LWS
 - 2. Kyphose der BWS
 - 3. Einbuchtungen der Grund- und Deckplatten
 - 4. mäßiggradige Knickplattfüße
 - 5. Beschwerdesymptomatik zwischen dem 8. und 12. Lebensjahr
- A nur 1 und 2 sind richtig
 - B nur 2 und 3 sind richtig
 - C nur 1 und 3 sind richtig
 - D nur 3 und 5 sind richtig
 - E nur 1, 2, 4 und 5 sind richtig

Frage 23) Nennen Sie die muskulären Strukturen der Rotatorenmanschette:

Frage 24) Nennen Sie 4 Ursachen, die zu einem Beckenschiefstand führen können:

Frage 25) Nennen Sie 4 Frühzeichen, die bei einem Perthes gelten:

Fragen zur Bewertung des Schwierigkeitsgrades der gestellten Fragen und zum Interesse am Fachbereich Orthopädie

Frage 26) Wie beurteilen Sie den Schwierigkeitsgrad der gestellten MC-Fragen (Fragen 11-22)?

sehr leicht
leicht
angemessen
schwierig
sehr schwierig

Frage 27) Wie beurteilen Sie den Schwierigkeitsgrad der gestellten Offenen Fragen (Fragen 23-25)?

sehr leicht
leicht
angemessen
schwierig
sehr schwierig

Frage 28) Wie schätzen Sie Ihr Interesse am Fachbereich Orthopädie ein?

Sehr großes Interesse vorhanden
Interesse vorhanden
unentschieden
wenig Interesse vorhanden
gar kein Interesse vorhanden

Frage 29) Können Sie es sich vorstellen, später den Beruf als Facharzt für Orthopädie auszuüben?

strebe den Beruf als Orthopäde/in an
kann ich mir vorstellen
unentschieden
kann ich mir nicht vorstellen
kommt für mich sicher nicht in Frage

Frage 30) Befürworten Sie die Durchführung solcher Evaluationen zum Thema Studienordnung und Fachwissenstand?

ja
nein
warum nicht?

Vielen Dank für die Mitarbeit!

Lösungsschema

Antworten zu Teil II des Fragebogens:

Wissensstand der Studenten im Fachbereich Orthopädie

MC-Fragen

Von den fünf Antwortmöglichkeiten jeder Frage war jeweils nur eine Antwort richtig. Für jede richtige Antwort wurde 1 Punkt vergeben, Abzüge für falsch beantwortete Fragen wurden nicht vollzogen.

Frage 11: Antwort B

Frage 12: Antwort D

Frage 13: Antwort B

Frage 14: Antwort B

Frage 15: Antwort D

Frage 16: Antwort C

Frage 17: Antwort D

Frage 18: Antwort A

Frage 19: Antwort C

Frage 20: Antwort D

Frage 21: Antwort A

Frage 22: Antwort B

Lösungsschema

Antworten zu Teil II des Fragebogens:

Wissensstand der Studenten im Fachbereich Orthopädie

Offene Fragen:

In Frage 23, 24 und 25 waren jeweils maximal 4 Punkte zu erreichen. Falschantworten hatten keinen Punktabzug bzw. keine Punktabzüge zur Folge.

Frage 23: Musculus supraspinatus
Musculus infraspinatus
Musculus subscapularis
Musculus teres minor

(Die Antwort „Ligamentum coracohumerale“ geht nicht in die Wertung mit ein, da in der Fragestellung explizit nach den muskulären Strukturen der Rotatorenmanschette gefragt ist.)

Frage 24: Hüftabduktionskontraktur
Hüftadduktionskontraktur
Kniebeugekontraktur
Spitzfuß

Ebenfalls mit einem Punkt werden folgende Aussagen gewertet:

Fraktur, Skoliose, Hüftdysplasie, Beinlängendifferenz und
Gluteusatrophie

Frage 25: Belastungsabhängige Schmerzen in der Hüfte
Schmerzen im Knie
Schmerzhafte Innenrotation im Hüftgelenk
Schmerzhafte Abduktion im Hüftgelenk

Ebenfalls mit einem Punkt werden folgende Aussagen gewertet:

Schonhinken, Gehschwäche und positives Drehmann-Zeichen

10.) Danksagung

Herrn Professor Dr. med. J. Grifka danke ich herzlich für die Möglichkeit, die vorliegende Untersuchung an der von ihm geleiteten Klinik und Poliklinik für Orthopädie der Universität Regensburg durchführen zu dürfen. Er begleitete mich mit fachlicher Kompetenz und engagierter Kritik vom ersten Tag bis zur Fertigstellung der Arbeit.

Herrn Professor Dr. med. V. Ewerbeck danke ich sehr für die Bereitschaft, dass die Untersuchung auch an der von ihm geleiteten Klinik und Poliklinik für Orthopädie der Universität Heidelberg stattfinden konnte.

Mein besonderer Dank gilt dem Betreuer der Untersuchung und der vorliegenden Arbeit, Herrn Priv.-Doz. Dr. med. C. Lüring. Die zügige Fertigstellung der Arbeit gelang vor allem durch den engen und intensiven Kontakt mit ihm. Die Arbeit profitierte wesentlich von seinem Engagement und seinen fachlichen Anregungen. Ebenso verdanke ich seinem Ideenreichtum einen deutlichen Zugewinn an sprachlicher Vielfalt.

Herrn Dr. med. F. Zeifang und Herrn Dr. med. C. Rau danke ich für die Betreuung der Untersuchung an der Klinik und Poliklinik für Orthopädie der Universität Heidelberg. Sie ermöglichten mir die Befragung der Studierenden zu Beginn des Orthopädie-Praktikums und führten diese selbst am Ende des Orthopädie-Praktikums durch.

Wichtigen Anteil am Gelingen der Arbeit tragen auch meine Eltern. Beide standen mir jederzeit zur Seite und sprachen mir immer wieder Mut zu.

11.) Lebenslauf

Zur Person

Name: Bredl
Vorname(n): Klaus Peter
Familienstand: ledig
Staatsangehörigkeit: deutsch
Geburtsdatum: 22.04.1982
Geburtsort: Regensburg

Ausbildung

April 2005 Studium der Humanmedizin an der Universität Regensburg
- Mai 2009 Kl. Abschnitt, Staatsexamen
 Abschluss: 3,0

Oktober 2006 Studium der Betriebswirtschaftslehre an der Universität
- März 2009 Regensburg, Bachelor of Science, B.sc. (BWL)
 Abschluss: 2,7
Bachelorarbeit: Theorie und Anwendung des strategischen
 Benchmarking im Gesundheitssektor
 Note: 1,3

Januar 2008 Intensivstudium Gesundheitsökonomie an der European
- Oktober 2008 Business School Östlich-Winkel
 Gesundheitsökonom (ebs)
 Abschluss: 1,8

Oktober 2002 Studium der Humanmedizin an der Otto-von-Guericke Universität
- März 2005 Magdeburg Vorklinischer Abschnitt

1999 – 2001 Goethe-Gymnasium Regensburg, Abitur
1992 – 1999 Albrecht-Altdorfer-Gymnasium Regensburg
1988 – 1992 Grundschule Tegernheim

Auszeichnungen

Stipendium e-fellows

Promotion

Lehre und Wissensvermittlung in der Orthopädie unter Berücksichtigung der modifizierten Approbationsordnung – eine Befragung an 506 Studenten unterschiedlicher Semester -

Orthopädische Universitätsklinik Regensburg, Prof. Dr. med. J. Grifka

Berufserfahrung

Oktober 2008 – Januar 2009	Praktisches Jahr: Universitätsklinikum Regensburg Innere Medizin, Chefarzt Prof. Dr. med. G. Riegger
Juni 2008 – Oktober 2008	Praktisches Jahr: Krankenhaus der Barmherzigen Brüder Regensburg Chirurgie, Chefarzt Prof. Dr. med. R. Neugebauer
April 2008 – Juni 2008	Praktisches Jahr: Orthopädische Universitätsklinik Regensburg Orthopädie, Chefarzt Prof. Dr. med. J. Grifka
Februar 2008 – April 2008	Unterassistent Schulthess Klinik Zürich Orthopädie, Chefarzt Dr. med. U. Munzinger
August 2007 – November 2007	Praktikum Oberender & Partner - Unternehmensberatung im Gesundheitswesen, Prof. Dr. Dr. h.c. P. Oberender Schwerpunkte: Strategische Beratung von stationären Leistungserbringern und Pharma- & Medizinprodukteindustrie, Innovationsfinanzierung

August 2006 - September 2006	Famulatur Dr. med. H. Buresch, Facharzt für Innere Medizin Praxis für Innere Krankheiten und hausärztliche Besuche
März 2006 - April 2006	Famulatur Prof. Dr. med. V. Ewerbeck Stiftung Orthopädische Universitätsklinik Heidelberg
Februar 2006 - März 2006	Famulatur Prof. Dr. med. J. Grifka Orthopädische Universitätsklinik Regensburg
August 2005 - September 2005	Famulatur Chefärztin Dr. med. R. Eisenmann-Klein Plastische-, Hand- und Wiederherstellungschirurgie Krankenhaus St. Josef Regensburg
Oktober 2001 - Dezember 2001	Krankenpflegedienst, Universitätsklinikum Regensburg Herz-Thorax-Chirurgie, Dermatologie

Sprachkenntnisse

Deutsch	(Muttersprache)
Englisch	(fließend)
Französisch	(Konversationsniveau)